

Dissertation zur Erlangung des naturwissenschaftlichen Doktorgrades der Bayerischen  
Julius-Maximilians-Universität Würzburg



*Implementierung von verbesserten  
Landoberflächenparametern und -prozessen in das  
hochaufgelöste Klimamodell REMO*

**vorgelegt von:  
Katrín Ziegler  
aus Nordheim am Main**

**Würzburg 2021**





Eingereicht am: 29. Juni 2021

Erstgutachter: Prof. Dr. Heiko Paeth

Zweitgutachterin: Prof. Dr. Daniela Jacob

Drittgutachter: Prof. Dr. Hubert Job

Mentor\*in: Prof. Dr. Daniela Jacob, Prof. Dr. Christopher Conrad

Tag des Kolloquiums: 14. März 2022

Prüfer\*in: Prof. Dr. Heiko Paeth, Prof. Dr. Daniela Jacob, Prof. Dr. Christopher Conrad



*„Never Give Up. Try one more time!“*

Raghav Pal



# Inhaltsverzeichnis

<b>Abbildungsverzeichnis</b>	<b>IV</b>
<b>Tabellenverzeichnis</b>	<b>VIII</b>
<b>Symbolverzeichnis</b>	<b>IX</b>
<b>Abkürzungsverzeichnis</b>	<b>X</b>
<b>Zusammenfassung</b>	<b>1</b>
<b>Abstract</b>	<b>4</b>
<b>1 Überblick über das Themenfeld Klimamodelle sowie Hinführung zum Thema und Aufbau der Arbeit</b>	<b>6</b>
1.1 Allgemeine Informationen zu Klimamodellen . . . . .	6
1.2 Projektbeschreibung und Aufbau dieser Arbeit . . . . .	8
<b>2 Theorie und Grundlagen der Modellierung von Landoberflächenprozessen</b>	<b>10</b>
2.1 Übersicht der Energieflüsse . . . . .	10
2.2 Definition verschiedener Bodenparameter und -prozesse . . . . .	14
<b>3 Das regionale Klimamodell REMO</b>	<b>22</b>
3.1 Entwicklung und Aufbau von REMO . . . . .	22
3.2 Technische Einführung und Modellaufbau . . . . .	24
3.2.1 Grundgerüst des Modells . . . . .	24
3.2.2 Ablauf eines Modelllaufs . . . . .	27
3.2.2.1 Präprozessor . . . . .	27
3.2.2.2 Modelllauf . . . . .	29
3.2.2.3 Postprozessierung . . . . .	30
3.2.2.4 Kalt- und Warmstart . . . . .	31
3.3 Beschreibung verschiedener Parametrisierungen für Prozesse in REMO . . . . .	35
3.4 Benötigte Eingangsdaten für REMO, deren Auflösung und Aufgaben . . . . .	41
3.4.1 Antriebsdaten . . . . .	41
3.4.2 Bisherige externe Oberflächendaten . . . . .	41
<b>4 Aufbereitung und Vergleich der neuen Oberflächendatensätze und Validationsdaten</b>	<b>45</b>
4.1 Statistische und programmiertechnische Methoden zur Datenaufbereitung und -analyse . . . . .	45
4.1.1 Verwendete Interpolationen . . . . .	45
4.1.2 Programmiertechnische Änderungen . . . . .	49

4.2	Analysmethoden zum Vergleich verschiedener Datensätze . . . . .	53
4.3	Beschreibung, Aufbereitung und Einbau der neuen Oberflächendaten . . . . .	55
4.3.1	Neue Topographie-Datensätze . . . . .	55
4.3.2	Neue Boden-Datensätze . . . . .	60
4.4	Vergleich der bisherigen und neuen Datensätze . . . . .	63
4.4.1	Topographie-Daten . . . . .	64
4.4.2	Bodendaten . . . . .	78
4.5	Beobachtungs- und Validierungsdaten . . . . .	86
4.5.1	Klimatologische Datensätze . . . . .	86
4.5.1.1	E-OBS . . . . .	86
4.5.1.2	CRU . . . . .	87
4.5.1.3	ERA-Interim . . . . .	87
4.5.1.4	ERA5 . . . . .	87
4.5.1.5	DWD . . . . .	88
4.5.2	Aufbereitung der Datensätze . . . . .	89
4.5.3	Validationsdatensätze für die Bodenparameter . . . . .	91
<b>5</b>	<b>Implementierung neuer Parametrisierungen und Prozesse sowie daraus resultierende Veränderungen im Programmcode für die Bodenparameter</b>	<b>93</b>
5.1	Neue Berechnung der abgeleiteten Bodenparameter . . . . .	93
5.1.1	Pedotransferfunktion für Bodenparameter . . . . .	94
5.1.2	Ausweitung der Bodenparameter in die Tiefe . . . . .	110
5.1.3	Verbesserung und alternative Formeln für die Wärmeeigenschaften des Bodens . . . . .	111
5.2	Neues Bodentemperaturschema und Anpassung des Bodenwärmestroms . . . . .	122
<b>6</b>	<b>Auswirkungen der neuen Datensätze und Prozesse</b>	<b>130</b>
6.1	Statistische Methoden zur Datenanalyse . . . . .	130
6.1.1	Univariate Kenngrößen . . . . .	130
6.1.2	Statistische Tests . . . . .	133
6.1.3	Verschiedene Indizes zur Feststellung der Güte . . . . .	140
6.1.4	Zusammenhänge von Variablen . . . . .	142
6.1.5	Weitere Gütemaße . . . . .	144
6.2	Übersicht der Modellexperimente . . . . .	147
6.3	Auswirkungen auf die Eingangsdaten . . . . .	152
6.3.1	Topographie-Daten . . . . .	152
6.3.2	Bodendaten . . . . .	156
6.4	Auswirkungen auf die abgeleiteten Parameter . . . . .	162
6.5	Auswirkungen auf klimatische Variablen . . . . .	169
6.5.1	Vergleich zwischen den verschiedenen Modellversionen . . . . .	169
6.5.1.1	Ergebnisse der Hypothesentests . . . . .	169
6.5.1.2	Veränderungen in ausgewählten Variablen . . . . .	175
6.5.2	Vergleich der Validationsdatensätze untereinander und mit den Modelldaten	204



6.5.2.1	Analyseergebnisse der kurzen Modellläufe der drei Experimentpläne	207
6.5.2.2	Analyseergebnisse der langen Modellläufe . . . . .	217
<b>7</b>	<b>Diskussion der Änderungen und Ergebnisse</b>	<b>248</b>
7.1	Probleme beim Vergleich zwischen Validations- und Modelldaten . . . . .	248
7.2	Einschätzung der Modellperformanz durch eingeführte Änderungen, Anregungen und weitere Erkenntnisse für die Modellentwicklung . . . . .	250
<b>8</b>	<b>Fazit und Ausblick</b>	<b>257</b>
	<b>Literaturverzeichnis</b>	<b>259</b>
	<b>Anhang</b>	<b>286</b>
	<b>Glossar</b>	<b>321</b>

# Abbildungsverzeichnis

Abb. 1:	Schaubild zur Entwicklung der Klimamodellkomponenten in den letzten Jahrzehnten . . . . .	7
Abb. 2:	Schematische Übersicht des Strahlungshaushaltes der Erde . . . . .	12
Abb. 3:	Darstellung eines regulären und rotierten Modelluntersuchungsgebiets anhand des EUR-44-Testgebiets . . . . .	25
Abb. 4:	Schematische Darstellung des Arakawa-C-Gitters . . . . .	25
Abb. 5:	Klassischer Workflow beim Ausführen eines REMO-Modelllaufs . . . . .	28
Abb. 6:	Neuer Workflow beim Ausführen eines REMO-Modelllaufs . . . . .	29
Abb. 7:	Kalt- und Warmstart im Vergleich - Flächenmittel des EUR-44 Testgebiets für zwei unterschiedliche Variablen . . . . .	33
Abb. 8:	Vergleich des Flächenmittels des EUR-44 Testgebiets für unterschiedliche lange Spin-Up Zeiten . . . . .	34
Abb. 9:	Schematisches Beispiel einer Landoberflächeneinteilung eines Pixels in REMO	35
Abb. 10:	Bodenschichten in REMO . . . . .	36
Abb. 11:	Schemata des bisherigen und des neuen Suchalgorithmus zur Interpolation des Bodendatensatzes in der BodLibKit . . . . .	46
Abb. 12:	Graphische Darstellung und Herleitung der Berechnung des Flächeninhaltes eines Dreieckes über die Determinante . . . . .	51
Abb. 13:	Weltkarte der EGM96 Geoid Höhe . . . . .	59
Abb. 14:	Schaubild der Gravitationsmodelle . . . . .	59
Abb. 15:	Geographischer Vergleich der Topographie-Datensätze in GER-1 km . . . . .	65
Abb. 16:	Statistischer Vergleich der Topographie-Datensätze in GER-1 km . . . . .	66
Abb. 17:	Beispiele von Fehlwerten im TANDEM-X Datensatz . . . . .	69
Abb. 18:	Vergleich der topographischen Höhe (FIB) von GTOPO- und ASTER-Datensatz	70
Abb. 19:	Statistischer Vergleich von GTOPO und ASTER für FIB in 0.44° Auflösung	72
Abb. 20:	Vergleich der Varianz der topographischen Höhe (VAROR) von GTOPO- und ASTER-Daten . . . . .	73
Abb. 21:	Statistischer Vergleich von GTOPO und ASTER für VAROR in 0.44° Auflösung	74
Abb. 22:	Vergleich der Rauigkeitslänge (AZ0) von GTOPO- und ASTER-Datensatz	75
Abb. 23:	Statistischer Vergleich von GTOPO und ASTER für AZ0 in 0.44° Auflösung	76
Abb. 24:	Topographievergleich von ASTER, GIMP und ArcticDEM für Grönland . .	77
Abb. 25:	Anteilmäßige Texturklassenverteilung der neuen Bodendatensätze in deren Originalauflösung . . . . .	79
Abb. 26:	Vergleich des Sandanteils des ersten Layers (SAND1) von FAOn- und SoilGrids-Datensatz . . . . .	80
Abb. 27:	Statistischer Vergleich von FAOn und Soilgrid für SAND in 0.44° Auflösung	81
Abb. 28:	Vergleich des Tonanteils des ersten Layers (CLAY1) von FAOn- und SoilGrids-Datensatz . . . . .	82

Abb. 29: Statistischer Vergleich von FAOn und Soilgrid für CLAY in 0.44° Auflösung	83
Abb. 30: Statistischer Vergleich von FAOn und Soilgrid für FORG in 0.44° Auflösung	84
Abb. 31: Vergleich des organischen Anteils des ersten Layers (FORG1) von FAOn- und SoilGrids-Datensatz . . . . .	85
Abb. 32: Schaubild zur Herleitung der Koordinaten für das Texturdreieck . . . . .	98
Abb. 33: Textur-Triangle mit den dazugehörigen 12 USDA-Texturklassen . . . . .	100
Abb. 34: Analyse der in REMO verwendeten sieben Bodenparameter abhängig vom jeweils zugewiesenen Texturklassenmittelwert für Sand- und Tonanteil . . .	103
Abb. 35: Vergleich der Werte der Bodenparameter anhand des USDA-Triangles für verschiedene Methoden . . . . .	105
Abb. 36: Mittelwert über die Landpunkte des EUR-44-Gebiets für die Variablen Sand-, Ton- und organischen Anteil für die originalen Tiefen sowie für die REMO-Schichten für HWS- und SoilGrids-Datensatz . . . . .	110
Abb. 37: Vergleich verschiedener Methoden zur Berechnung der Wärmeleitfähigkeit unter verschiedenen Bodenbedingungen . . . . .	115
Abb. 38: Vergleich der Wärmeleitfähigkeit nach Lawrence & Slater (2008) in verschiedenen Böden . . . . .	117
Abb. 39: Vergleich der volumetrischen Wärmekapazität nach Lawrence & Slater (2008) in verschiedenen Böden . . . . .	121
Abb. 40: Differenzensterne zur Darstellung der räumlichen und zeitlichen Zusammenhänge bei verschiedenen Lösungsansätze . . . . .	123
Abb. 41: Auswirkung der veränderten Parametrisierung und des neuen Temperaturschemas auf die Oberflächentemperatur anhand zweier Beispielpixel . . . . .	128
Abb. 42: Schaubild zur Erklärung der Ableitung von Windgeschwindigkeit und -richtung aus der u- und v-Komponente . . . . .	132
Abb. 43: Schematischer Aufbau einer Vierfeldertafel und die jeweiligen Indizes . . . . .	146
Abb. 44: Experimentplan 1: Übersicht der durchgeführten Läufe zum Vergleich verschiedener Topographie- und Bodendatensätze . . . . .	147
Abb. 45: Experimentplan 2: Übersicht der durchgeführten Läufe zum Vergleich der verschiedenen Modellschalter . . . . .	149
Abb. 46: Experimentplan 3: Übersicht der durchgeführten Läufe zum Vergleich der verschiedenen Parametrisierungen und kleiner Änderungen . . . . .	150
Abb. 47: Topographie-Daten im Vergleich für EUR-44 . . . . .	153
Abb. 48: Topographie-Daten im Vergleich für GER-11 . . . . .	155
Abb. 49: Sandanteil der neuen Bodendatensätze im Vergleich für EUR-44 . . . . .	157
Abb. 50: Sandanteil der neuen Bodendatensätze im Vergleich für GER-11 . . . . .	157
Abb. 51: Tonanteil der neuen Bodendatensätze im Vergleich für EUR-44 . . . . .	159
Abb. 52: Tonanteil der neuen Bodendatensätze im Vergleich für GER-11 . . . . .	159
Abb. 53: Organischer Anteil der neuen Bodendatensätze im Vergleich für EUR-44 . .	160
Abb. 54: Organischer Anteil der neuen Bodendatensätze im Vergleich für GER-11 . .	161
Abb. 55: Räumlicher Vergleich der Amplitude der Wärmeleitfähigkeit von Validationsdaten und unterschiedlichen Modellläufen in GER-11 . . . . .	164

Abb. 56: Räumlicher Vergleich des Porenvolumens von Validationsdaten und unterschiedlichen Modellläufen in GER-11 . . . . .	166
Abb. 57: Räumlicher Vergleich der Feldkapazität von Validationsdaten und unterschiedlichen Modellläufen in GER-11 . . . . .	167
Abb. 58: Erklärte Varianz der Faktoren der pixelweisen 2 x 2 ANOVA . . . . .	174
Abb. 59: Mittlere Windgeschwindigkeit und -richtung der ERA5-Daten für beide Untersuchungsgebiete für das Jahr 2000 . . . . .	177
Abb. 60: Zeitreihe des räumlichen Mittelwerts der Windgeschwindigkeit und der mittleren Windrichtung von GER-11 für ERA5 und Modelllauf id01 für das Jahr 2000 . . . . .	178
Abb. 61: Unterschied der Windrichtung und -geschwindigkeit von Modelllauf id01 und ERA5 abhängig von der Windrichtung . . . . .	180
Abb. 62: Differenzkarten für Windgeschwindigkeit und -richtung zwischen Modelllauf id01 und ERA5 . . . . .	182
Abb. 63: Unterschied der Windrichtung und -geschwindigkeit von Modelllauf id01_AFT und id01 für EUR-44 . . . . .	185
Abb. 64: Unterschied der Windrichtung und -geschwindigkeit von Modelllauf id01_AFT und id01 für GER-11 . . . . .	186
Abb. 65: Vergleich des latenten Wärmeflusses der verschiedenen Modellläufe von Experimentplan 1 mit den ERA5-Validationsdaten für EUR-44 und GER-11 .	189
Abb. 66: Mittlere jährliche Temperatur- und Wärmeflussvariablen der ERA5-Daten für GER-11 . . . . .	191
Abb. 67: Differenz der mittleren jährlichen Temperatur- und Wärmeflussvariablen zwischen dem bisherigen Modelllauf id01 und den ERA5-Daten für GER-11	192
Abb. 68: Differenz verschiedener Modellläufe von Experimentplan 1 und verschiedener Variablen zu Modelllauf id01 für GER-11 . . . . .	195
Abb. 69: Vergleich des latenten und sensiblen Wärmeflusses sowie der mittleren Temperatur der verschiedenen Modellläufe von Experimentplan 2 mit den ERA5-Validationsdaten für GER-11 . . . . .	197
Abb. 70: Zeitliches und räumliches Mittel der Bodentemperaturen im Vergleich für EUR-44 . . . . .	200
Abb. 71: Zeitliches und räumliches Mittel der Bodentemperaturen im Vergleich für GER-11 . . . . .	202
Abb. 72: Bodentemperaturprofile der Modellläufe des Experimentplan 2 im Vergleich für die Monatsmittelwerte des Jahres 2000 . . . . .	203
Abb. 73: Kernfunktionsdichteschätzung der klimatologischen Validationsdatensätze pro Jahreszeit für EUR-44 . . . . .	205
Abb. 74: Übersicht des 1., 2. und 3. Quartils des jeweiligen RMSE und Korrelationskoeffizienten der Modellläufe des ersten Experimentplans im Vergleich zu verschiedenen Validationsdaten für RR, T2M, T2MAX und T2MIN pro Jahr und Jahreszeit . . . . .	208

Abb. 75: Übersicht des 1., 2. und 3. Quartils des jeweiligen RMSE und Korrelationskoeffizienten der Modellläufe des zweiten Experimentplans im Vergleich zu den EOBSV21-Daten für RR, T2M, T2MAX und T2MIN pro Jahr und Jahreszeit	209
Abb. 76: Übersicht des 1., 2. und 3. Quartils des jeweiligen RMSE und Korrelationskoeffizienten einiger Modellläufe des dritten Experimentplans im Vergleich zu den EOBSV21-Daten für RR, T2M, T2MAX und T2MIN pro Jahr und Jahreszeit . . . . .	210
Abb. 77: Darstellung der Overall-Accuracy verschiedener Variablen pro Modelllauf des Experimentplans 3 für EUR-44 und GER-11 . . . . .	216
Abb. 78: Durchschnittliche Temperatur- und Niederschlagsanomalien der Jahre 1950-2018 im Vergleich zur Referenzperiode 1961-1990 laut EOBSV19 in 0.25° Auflösung . . . . .	219
Abb. 79: Übersicht des 1., 2. und 3. Quartils des jeweiligen RMSE und Korrelationskoeffizienten der vier Warmstartläufe im Vergleich zu den Validationsdaten für RR, T2M, T2MAX und T2MIN pro Jahr und Jahreszeit . . . . .	221
Abb. 80: Bodentemperaturprofile für Modell- und Validationsdaten im Vergleich für Monatsmittelwerte des Jahres 2018 . . . . .	227
Abb. 81: Boxplot der mittleren Bodentemperaturen pro Bodenschicht und Jahreszeit für den Beispielpixel für verschiedene Beobachtungs- und Modelldaten der Jahre 2000-2018 . . . . .	230
Abb. 82: Räumlicher Vergleich der jährlichen und saisonalen Bodentemperaturen zwischen den Validationsdaten und unterschiedlichen Modellläufen für Deutschland	232
Abb. 83: Vergleich der Zeitreihen des Flächenmittels vom GER-11-Gebiet täglicher Werte von Juni bis August 2003 für die verschiedenen Warmstartläufe und ERA5 . . . . .	236
Abb. 84: Vergleich zwischen Validationsdaten und Modellläufen anhand verschiedener Kenntage in Deutschland für den Sommer 2003 . . . . .	238
Abb. 85: Räumlicher Vergleich der mittleren Minimumtemperatur zwischen Validationsdaten und Modellläufen für Juni bis August 2003 . . . . .	240
Abb. 86: Vergleich der Zeitreihen des Flächenmittels vom GER-11-Gebiet täglicher Werte von November 2005 bis März 2006 für die verschiedenen Warmstartläufe und ERA5 . . . . .	242
Abb. 87: Vergleich zwischen Validationsdaten und Modellläufen anhand verschiedener Kenntage in Deutschland für den Winter 2005/2006 . . . . .	243
Abb. 88: Räumlicher Vergleich der mittleren Maximumtemperatur zwischen Validationsdaten und Modellläufen für November 2005 bis März 2006 . . . . .	245
Abb. 89: Räumlicher Vergleich der Anzahl der Frosttage zwischen Validationsdaten und Modellläufen für November 2005 bis März 2006 . . . . .	246

# Tabellenverzeichnis

Tab. 1: Volumetrische Wärmekapazität von verschiedenen Materialien . . . . .	15
Tab. 2: Wärmeleitfähigkeit von verschiedenen Materialien . . . . .	16
Tab. 3: Bodenparameter für die jeweilige Bodentextur . . . . .	40
Tab. 4: Übersicht der bisher verwendeten Datensätze in der BodLibKit . . . . .	42
Tab. 5: Reklassifikation der FAO-Bodentexturen in der Bodenbibliothek von REMO	43
Tab. 6: Variablen der bisherigen Bodenbibliothek . . . . .	44
Tab. 7: Allgemeine Informationen zu den verwendeten topographische Datensätzen .	56
Tab. 8: Eigenschaften der verwendeten topographische Datensätzen im Vergleich . .	58
Tab. 9: Allgemeine Informationen zu den verwendeten Boden-Datensätzen . . . . .	62
Tab. 10: Statistische Kennwerte der fünf verschiedenen Topographie-Datensätze für GER-1 km im Vergleich . . . . .	66
Tab. 11: Vergleich der verschiedenen Bodenparameter von REMO mit Literaturwerten pro Texturklasse . . . . .	96
Tab. 12: Verschiedene Berechnungsmethoden der Wärmeleitfähigkeit des Bodens im Vergleich . . . . .	113
Tab. 13: Übersicht der möglichen Effekte bei einer 2 x 2 ANOVA . . . . .	139
Tab. 14: Übersicht der Bodenparameteränderungen . . . . .	163
Tab. 15: Anzahl der Variablen pro Modelllauf des Experimentplan 1 mit signifikanten Unterschieden für GER-11 . . . . .	170
Tab. 16: Übersicht der durchschnittlichen Overall-Accuracy-Werte für Experimentplan 1 für beide Untersuchungsgebiete . . . . .	212
Tab. 17: Übersicht der durchschnittlichen Overall-Accuracy-Werte für Experimentplan 2 für beide Untersuchungsgebiete . . . . .	213
Tab. 18: Übersicht der Qualitätsindizes verschiedener Variablen für die langen Warm- startläufe . . . . .	223
Tab. 19: Übersicht und Definition der verwendeten Klimaindizes . . . . .	234

# Symbolverzeichnis

$\Psi$	Porosität
$c_a$	volumetrische Wärmekapazität von Luft in $\frac{J}{m^3K}$
$c_i$	volumetrische Wärmekapazität von Eis in $\frac{J}{m^3K}$
$c_o$	volumetrische Wärmekapazität von organischem Material im Boden in $\frac{J}{m^3K}$
$c_{sol}$	volumetrische Wärmekapazität des Feststoffanteils im Boden in $\frac{J}{m^3K}$
$c_s$	spezifische Wärmekapazität in $\frac{J}{kgK}$
$c_v$	volumetrische Wärmekapazität in $\frac{J}{m^3K}$
$c_w$	volumetrische Wärmekapazität von Wasser in $\frac{J}{m^3K}$
$\alpha$	Temperaturleitfähigkeit, thermische Diffusivität des Bodens in $\frac{m^2}{s}$
$\alpha_{ice}$	Temperaturleitfähigkeit von Eis in $\frac{m^2}{s}$
$\kappa$	Kappa-Koeffizient
$K_e$	Kerstenfaktor, Kerstenzahl
$\lambda$	Wärmeleitfähigkeit des Bodens in $\frac{W}{mK}$
$\lambda_a$	Wärmeleitfähigkeit von Luft (20 °C) in $\frac{W}{mK}$
$\lambda_B$	Wärmeleitfähigkeit von Basalt in $\frac{W}{mK}$
$\lambda_C$	Wärmeleitfähigkeit von Ton in $\frac{W}{mK}$
$\lambda_{dry}$	Wärmeleitfähigkeit von trockenem Boden in $\frac{W}{mK}$
$\lambda_i$	Wärmeleitfähigkeit von Eis (0 °C) in $\frac{W}{mK}$
$\lambda_o$	Wärmeleitfähigkeit von organischem Material in $\frac{W}{mK}$
$\lambda_Q$	Wärmeleitfähigkeit von Quarz in $\frac{W}{mK}$
$\lambda_{sat}$	Wärmeleitfähigkeit von gesättigtem Boden in $\frac{W}{mK}$
$\lambda_w$	Wärmeleitfähigkeit von Wasser in $\frac{W}{mK}$
$\Gamma$	Temperaturgradient (hier: Abnahme der Temperatur mit zunehmender Höhe)
$\phi$	Longitude des Rotationspols bei der Koordinatentransformation
$e$	Porenziffer, Porenzahl
$\rho$	Dichte in $\frac{kg}{m^3}$
$\rho_{ice}$	Dichte von Eis in $\frac{kg}{m^3}$
$\rho_{soil}$	Dichte des Bodens in $\frac{kg}{m^3}$
$\rho_{water}$	Dichte von Wasser in $\frac{kg}{m^3}$
$\theta$	Latitude des Rotationspols bei der Koordinatentransformation
$\theta$	Wassergehalt des Bodens in $\frac{m^3}{m^3}$
$\theta_r$	Wassergehalt von trockenem Boden in $\frac{m^3}{m^3}$
$\theta_s$	Wassergehalt von gesättigtem Boden in $\frac{m^3}{m^3}$ , entspricht dem Porenvolumen

# Abkürzungsverzeichnis

<b>AA</b>	Average Accuracy
<b>ACC</b>	Anomaly Correlation Coefficient
<b>ADW</b>	Winkelabstandsgewichtete Interpolation
<b>AFR</b>	afrikanische CORDEX-Region
<b>ALOS</b>	Advanced Land Observing Satellite
<b>ANN</b>	Artificial Neural Networks
<b>ANOVA</b>	Analysis of Variance
<b>Antarctic RAMP2</b>	Radarsat Antarctic Mapping Project DEM Version 2
<b>ASI</b>	Italien Space Agency
<b>ASTER</b>	Advanced Spaceborne Thermal Emission and Reflection Radiometer
<b>AUS</b>	australische CORDEX-Region
<b>AV</b>	Added Value
<b>AVHRR</b>	Advanced Very High Resolution Radiometer
<b>AW3D30</b>	ALOS World 3D 30m
<b>BALTEX</b>	Baltic Sea Experiment
<b>BATS</b>	Biosphere-Atmosphere Transfer Scheme
<b>BodLibKit</b>	Bodenbibliothek
<b>BPRC</b>	Byrd Polar Research Center
<b>CABLE</b>	Community Atmosphere Biosphere Land Exchange
<b>CDD</b>	höchste Anzahl der Tage in eine Trockenperiode
<b>CDO</b>	Climate Data Operators
<b>CE</b>	Commission Error
<b>CEDA</b>	Centre for Environmental Data Analysis
<b>CFD</b>	höchste Anzahl der Tage in einer Frostperiode
<b>CFSR</b>	Climate Forecast System Reanalysis
<b>CLASS</b>	Canadian Land Surface Scheme
<b>CLM</b>	Community Land Model
<b>CoLm</b>	Common Land Model
<b>CORDEX</b>	Coordinated Regional Climate Downscaling Experiment
<b>Corine</b>	Coordination of Information on the Environment
<b>COSMO</b>	Consortium for SmallScale Modeling
<b>CPU</b>	Zentrale Recheneinheit (engl. Central Processing Unit)
<b>CRU</b>	Climatic Research Unit
<b>CSU</b>	höchste Anzahl der Tage in einer Sommerperiode
<b>CWD</b>	höchste Anzahl der Tage in einer Regenperiode
<b>DEM</b>	Digital Elevation Model
<b>DGM</b>	digitales Geländemodell
<b>DKRZ</b>	Deutsches Klimarechenzentrum



<b>DLR</b>	Deutsches Luft- und Raumfahrtzentrum
<b>DM</b>	Deutschland-Modell
<b>DOE</b>	Department of Energy
<b>DWD</b>	Deutscher Wetterdienst
<b>E-OBS</b>	European daily high-resolution gridded dataset
<b>EAS</b>	ostasiatische CORDEX-Region
<b>ECHAM</b>	ECMWF Hamburg Model
<b>ECMWF</b>	European Center for Medium-Range Weather Forecasts
<b>EGM</b>	Earth Gravitational Model
<b>EM</b>	Europa-Modell
<b>EOBSV19</b>	E-OBS Version v19.0e
<b>EOBSV21</b>	E-OBS Version v21.0e
<b>ERA-40</b>	40 Jahre ECMWF Reanalysis
<b>ERA-Interim</b>	ECMWF interim Reanalysis
<b>ERA5</b>	ECMWF Reanalysis 5th Generation
<b>EROS</b>	U.S. Geological Survey Center for Earth Resources Observation and Science
<b>ESDB</b>	Europäische Bodendatenbank
<b>ESM</b>	Erdsystemmodell
<b>ETCCDI</b>	Expert Team on Climate Change Detection and Indices
<b>EUR-44</b>	europäische CORDEX-Region in 0.44° Auflösung
<b>FAO</b>	Food and Agricultural Organization of the United Nations
<b>FD</b>	Anzahl der Frosttage
<b>FOR</b>	False Omission Rate
<b>FPAR</b>	Anteil der von der Vegetation absorbierten photosynthetisch aktiven Strahlung an der theoretisch zur Photosynthese verfügbaren Strahlung (engl. Fraction of absorbed Photosynthetic Active Radiation)
<b>FPR</b>	False Positive Rate
<b>GCM</b>	globales Klimamodell
<b>GER-11</b>	definiertes Untersuchungsgebiet für Deutschland in 0.11° Auflösung
<b>GERICS</b>	Deutsches Klimaservicezentrum (engl. Climate Service Center Germany)
<b>GIMP</b>	Greenland Ice Mapping Project
<b>GKSS</b>	Gesellschaft für Kernenergieverwertung in Schiffbau und Schifffahrt
<b>GLAS</b>	Geoscience Laser Altimeter System
<b>GLCCD</b>	Global Land Cover Characterization Dataset
<b>GLOBE</b>	Global Land One-km Base Elevation
<b>GSDE</b>	Global Soil Dataset for Earth System Models
<b>GSFC</b>	Goddard Space Flight Center
<b>HiHydroSoil</b>	High Resolution Soil Map of Hydraulic Properties
<b>HIRAM</b>	High Resolution Atmospheric Model
<b>HIRLAM</b>	High Resolution Limited Area Model

<b>HTESSEL</b>	Hydrology-Tiled ECMWF Scheme for Surface Exchanges over Land
<b>HWSD</b>	Harmonized World Soil Database
<b>HZG</b>	Helmholtz-Zentrum Geesthacht
<b>IceSat</b>	Ice, Cloud and land Elevation Satellite
<b>IDW</b>	Inverse Distance Weighted
<b>IGBP</b>	International Geosphere-Biosphere Programme
<b>InSAR</b>	Radarinferometrie
<b>IO</b>	Einlesen und Herausschreiben (engl. Input/Output) der Daten
<b>IQR</b>	Interquartilsabstand
<b>ISBA</b>	Interaction Soil-Biosphere-Atmosphere
<b>JAXA</b>	Japan Aerospace Exploration Agency
<b>JMA</b>	Japan Meteorological Agency
<b>JRA-25</b>	25 Jahre Japanese Reanalysis der JMA
<b>JULES</b>	Joint UK Land Environment Simulator
<b>KDE</b>	Kernel Density Estimation
<b>KS</b>	Kolmogoroff-Smirnoff
<b>LAI</b>	Blattflächenindex (engl. Leaf Area Index)
<b>LBC</b>	Randbedingungen (engl. Lateral Boundary Conditions)
<b>LiDAR</b>	Light Detection And Ranging
<b>LPJmL</b>	Lund-Potsdam-Jena managed Land
<b>LSM</b>	Landoberflächenmodelle (engl. Land Surface Models)
<b>MA</b>	Mean Accuracy
<b>MERIT DEM</b>	Multi-Error-Removed Improved-Terrain DEM
<b>MERRA</b>	Modern-Era Retrospective Analysis for Research and Applications
<b>METI</b>	Ministry of Economy, Trade, and Industry
<b>MODIS</b>	Moderate-resolution Imaging Spectroradiometer
<b>MPI-M</b>	Max-Planck-Institut für Meteorologie
<b>NASA</b>	National Aeronautics and Space Administration
<b>NCAR</b>	National Center for Atmospheric Research
<b>NCEP</b>	National Centers for Environmental Prediction
<b>NDVI</b>	Normalized Difference Vegetation Index
<b>nFK</b>	nutzbare Feldkapazität
<b>NGA</b>	National Geospatial-Intelligence Agency
<b>NNSE</b>	Normalized Nash-Sutcliffe Effizienz
<b>NOAA</b>	National Oceanic and Atmospheric Administration
<b>NPV</b>	Negative Predicted Value
<b>NSE</b>	Nash-Sutcliffe Effizienz
<b>OA</b>	Overall Accuracy
<b>OE</b>	Omission Error
<b>ORCHIDEE</b>	Organizing Carbon and Hydrology In Dynamic Ecosystems
<b>PA</b>	Producer Accuracy

<b>PBIAS</b>	prozentualer Bias
<b>PiP</b>	Point-in-Polygon
<b>PP</b>	Bodendruck
<b>PPR</b>	Positive Predicted Rate
<b>PRISM</b>	Panchromatic Remote-sensing Instrument for Stereo Mapping
<b>PRM</b>	Purdue Regional Model
<b>PTF</b>	Pedotransferfunktion
<b>PWP</b>	permanenter Welkepunkt
<b>RCM</b>	regionales Klimamodell
<b>REMO</b>	Regional Modell
<b>RMSE</b>	Root Mean Square Error
<b>RoT</b>	„rules of thumb“
<b>RR</b>	Niederschlag
<b>RR1mm</b>	Anzahl der Regentage
<b>RR20mm</b>	Anzahl der Starkregentage
<b>rRMSE</b>	relative Root Mean Square Error
<b>RSR</b>	RMSE-observations Standard deviation Ratio
<b>SAM</b>	südamerikanische CORDEX-Region
<b>SEWAB</b>	Surface Energy and Water Balance
<b>SiB</b>	Simple Biosphere Model
<b>SOC</b>	Soil Organic Content
<b>SOM</b>	Soil Organic Matter
<b>SOTER</b>	Soil and Terrain
<b>SRTM</b>	Shuttle Radar Topography Mission
<b>SU</b>	Anzahl der Sommertage
<b>SURFEX</b>	Surface Externalisée
<b>TanDEM-X</b>	TerraSAR-X add-on for Digital Elevation Measurements
<b>TESSEL</b>	Tiled ECMWF Scheme for Surface Exchanges over Land
<b>TG</b>	mittlere Temperatur
<b>TN</b>	Minimaltemperatur
<b>TNR</b>	True Negative Rate
<b>TPR</b>	True Positive Rate
<b>TR</b>	Anzahl der Tropennächte
<b>TX</b>	Maximaltemperatur
<b>UA</b>	User Accuracy
<b>UM</b>	Met Office Unified Model
<b>USDA</b>	United States Department of Agriculture
<b>USGS</b>	United States Geological Survey
<b>WaSiM</b>	Water Balance Simulation Model
<b>WAVA</b>	pflanzenverfügbares Bodenwasser (engl. plant-available soil water)
<b>WCRP</b>	World Climate Research Program

<b>WISE</b>	World Inventory of Soil Emission Potential
<b>WRC</b>	Water Retention Curve
<b>WRF</b>	Weather Research and Forecasting

# Zusammenfassung

Das Ziel dieser Arbeit war neue Eingangsdaten für die Landoberflächenbeschreibung des regionalen Klimamodells REMO zu finden und ins Modell zu integrieren, um die Vorhersagequalität des Modells zu verbessern. Die neuen Daten wurden so in das Modell eingebaut, dass die bisherigen Daten weiterhin als Option verfügbar sind. Dadurch kann überprüft werden, ob und in welchem Umfang sich die von jedem Klimamodell benötigten Rahmendaten auf Modellergebnisse auswirken. Im Zuge der Arbeit wurden viele unterschiedliche Daten und Methoden zur Generierung neuer Parameter miteinander verglichen, denn neben dem Ersetzen der konstanten Eingangswerte für verschiedene Oberflächenparameter und den damit verbundenen Änderungen wurden als zusätzliche Verbesserung auch Veränderungen an der Parametrisierung des Bodens speziell in Hinblick auf die Bodentemperaturen in REMO vorgenommen. Im Rahmen dieser Arbeit wurden die durch die verschiedenen Änderungen ausgelösten Auswirkungen für das CORDEX-Gebiet EUR-44 mit einer Auflösung von ca. 50 km und für das in dem darin eingebetteten neu definierten Deutschlandgebiet GER-11 mit einer Auflösung von ca. 12 km getestet sowie alle Änderungen anhand von verschiedenen Beobachtungsdatensätzen validiert.

Die vorgenommenen Arbeiten gliederten sich in drei Hauptteile. Der erste Teil bestand in dem vom eigentlichen Klimamodell unabhängigen Vergleich der verschiedenen Eingangsdaten auf unterschiedlichen Auflösungen und deren Performanz in allen Teilen der Erde, wobei ein besonderer Fokus auf der Qualität in den späteren Modellgebieten lag. Unter Berücksichtigung der Faktoren, wie einer globalen Verfügbarkeit der Daten, einer verbesserten räumlichen Auflösung und einer kostenlosen Nutzung der Daten sowie verschiedener Validationsergebnissen von anderen Studien, wurden in dieser Arbeit vier neue Topographiedatensätze (SRTM, ALOS, TANDEM und ASTER) und drei neue Bodendatensätze (FAOn, Soilgrid und HWSD) für die Verwendung im Präprozess von REMO aufbereitet und miteinander sowie mit den bisher in REMO verwendeten Daten verglichen. Auf Grundlage dieser Vergleichsstudien schieden bei den Topographiedaten die verwendeten Datensatz-Versionen von SRTM, ALOS und TANDEM für die in dieser Arbeit durchgeführten REMO-Läufe aus. Bei den neuen Bodendatensätzen wurde ausgenutzt, dass diese verschiedene Bodeneigenschaften für unterschiedliche Tiefen als Karten zur Verfügung stellen. In REMO wurden bisher alle benötigten Bodenparameter abhängig von fünf verschiedenen Bodentexturklassen und einer zusätzlichen Torfklasse ausgewiesen und als konstant über die gesamte Modellbodensäule (bis ca. 10 m) angenommen. Im zweiten Teil wurden auf Basis der im ersten Teil ausgewählten neuen Datensätze und den neu verfügbaren Bodenvariablen verschiedene Sensitivitätsstudien über das Beispieljahr 2000 durchgeführt. Dabei wurden verschiedene neue Parametrisierungen für die bisher aus der Textur abgeleiteten Bodenvariablen und die Parametrisierung von weiteren hydrologischen und thermalen Bodeneigenschaften verglichen. Ferner wurde aufgrund der neuen nicht über die Tiefe konstanten Bodeneigenschaften eine neue numerische Methode zur Berechnung der Bodentemperaturen der fünf Schichten in REMO getestet, welche wiederum andere Anpassungen erforderte. Der Test und die Auswahl der verschiedenen Datensatz- und Parametrisierungsversionen auf die Modellperformanz wurde

in drei Experimentpläne unterteilt. Im ersten Plan wurden die Auswirkungen der ausgewählten Topographie- und Bodendatensätze überprüft. Der zweite Plan behandelte die Unterschiede der verschiedenen Parametrisierungsarten der Bodenvariablen hinsichtlich der verwendeten Variablen zur Berechnung der Bodeneigenschaften, der über die Tiefe variablen oder konstanten Eigenschaften und der verwendeten Berechnungsmethode der Bodentemperaturänderungen. Durch die Erkenntnisse aus diesen beiden Experimentplänen, die für beide Untersuchungsgebiete durchgeführt wurden, ergaben sich im dritten Plan weitere Parametrisierungsänderungen. Alle Änderungen dieses dritten Experimentplans wurden sukzessiv getestet, sodass der paarweise Vergleich von zwei aufeinanderfolgenden Modellläufen die Auswirkungen der Neuerung im jeweils zweiten Lauf widerspiegelt. Der letzte Teil der Arbeit bestand aus der Analyse von fünf längeren Modellläufen (2000-2018), die zur Überprüfung der Ergebnisse aus den Sensitivitätsstudien sowie zur Einschätzung der Performanz in weiteren teilweise extremen atmosphärischen Bedingungen durchgeführt wurden. Hierfür wurden die bisherige Modellversion von REMO (id01) für die beiden Untersuchungsgebiete EUR-44 und GER-11 als Referenzläufe, zwei aufgrund der Vergleichsergebnisse von Experimentplan 3 selektierte Modellversionen (id06 und id15a für GER-11) sowie die finale Version (id18a für GER-11), die alle vorgenommenen Änderungen dieser Arbeit enthält, ausgewählt.

Es stellte sich heraus, dass sowohl die neuen Topographiedaten als auch die neuen Bodendaten große Differenzen zu den bisherigen Daten in REMO haben. Zudem änderten sich die von diesen konstanten Eingangsdaten abgeleiteten Hilfsvariablen je nach verwendeter Parametrisierung sehr deutlich. Dies war besonders gut anhand der Bodenparameter zu erkennen. Sowohl die räumliche Verteilung als auch der Wertebereich der verschiedenen Modellversionen unterschieden sich stark. Eine Einschätzung der Qualität der resultierenden Parameter wurde jedoch dadurch erschwert, dass auch die verschiedenen zur Validierung herangezogenen Bodendatensätze für diese Parameter deutlich voneinander abweichen. Die finale Modellversion id18a ähnelte trotz der umfassenden Änderungen in den meisten Variablen den Ergebnissen der bisherigen REMO-Version. Je nach zeitlicher und räumlicher Aggregation sowie unterschiedlichen Regionen und Jahreszeiten wurden leichte Verbesserungen, aber auch leichte Verschlechterungen im Vergleich zu den klimatologischen Validationsdaten festgestellt. Größere Veränderungen im Vergleich zur bisherigen Modellversion konnten in den tieferen Bodenschichten aufgezeigt werden, welche allerdings aufgrund von fehlenden Validationsdaten nicht beurteilt werden konnten. Für alle 2 m-Temperaturen konnte eine tendenzielle leichte Erwärmung im Vergleich zum bisherigen Modelllauf beobachtet werden, was sich einerseits negativ auf die ohnehin durchschnittlich zu hohe Minimumtemperatur, aber andererseits positiv auf die bisher zu niedrige Maximumtemperatur des Modells in den betrachteten Gebieten auswirkte. Im Niederschlagssignal und in den 10 m-Windvariablen konnten keine signifikanten Änderungen nachgewiesen werden, obwohl die neue Topographie an manchen Stellen im Modellgebiet deutlich von der bisherigen abweicht. Des Weiteren variierte das Ranking der verschiedenen Modellversionen jeweils nach dem angewendeten Qualitätsindex.

Um diese Ergebnisse besser einordnen zu können, muss berücksichtigt werden, dass die neuen Daten für Modellgebiete mit 50 bzw. 12 km räumlicher Auflösung und der damit verbundenen

---

hydrostatischen Modellversion getestet wurden. Zudem sind vor allem in Fall der Topographie die bisher enthaltenen GTOPO-Daten (1 km Auflösung) für die Aggregation auf diese gröbere Modellauflösung geeignet. Die bisherigen Bodendaten stoßen jedoch mit 50 km Auflösung bereits an ihre Grenzen. Zusätzlich ist zu beachten, dass nicht nur die Mittelwerte dieser Daten, sondern auch deren Subgrid-Variabilität als Variablen im Modell für verschiedene Parametrisierungen verwendet werden. Daher ist es essentiell, dass die Eingangsdaten eine deutlich höhere Auflösung bereitstellen als die zur Modellierung definierte Auflösung. Für lokale Klimasimulationen mit Auflösungen im niedrigen Kilometerbereich spielen auch die Vertikalbewegungen (nicht-hydrostatische Modellversion) eine wichtige Rolle, die stark von der Topographie sowie deren horizontaler und vertikaler Änderungsrate beeinflusst werden, was die in dieser Arbeit eingebauten wesentlich höher aufgelösten Daten für die zukünftige Weiterentwicklung von REMO wertvoll machen kann.

# Abstract

The main aim of this work was to find new input data sets for the land surface description of the regional climate model REMO and to integrate them into the model in order to improve the predictive quality of the model. The new data sets have been incorporated into the model in such a way that the previous data are still available as an option for the model run. This allows to check whether and to what extent the boundary data required by each climate model have an impact on the model results. In this study comparisons of many different data sets and methods for generating new parameters are included. In addition to replacing the constant input values for different surface parameters and the associated changes, changes were also made for the parameterization of the soil, especially with regard to the soil temperatures in REMO. The effects of different changes which were made in this study were analysed for the CORDEX region EUR-44 with a resolution of 50 km and for a newly defined German area GER-11 with a resolution of 12 km. All changes were validated with different observational data sets.

The work process was divided into three main parts. The first part was independent of the actual climate model and included the comparison of different input data sets at different resolutions and their performance in all parts of the world. Taking into account factors such as global availability of the data, improved spatial resolution and free use of the data, as well as various validation results from other studies, four new topography data sets (SRTM, ALOS, TANDEM and ASTER) and three new soil data sets (FAOn, Soilgrid and HWSD) were processed for the usage by REMO and compared with each other and with the data sets previously used in REMO. Based on these comparative studies of the topographical data sets the SRTM, ALOS and TANDEM data set versions were excluded from the further usage in REMO in this study. For the new soil data sets the fact that they provide different soil properties for different depths as maps has been taken advantage of. In the previous REMO versions, all required soil parameters so far have been determined depending on five different soil texture classes with an additional peat class and assumed to be constant over the entire model soil column (up to approximately 10 m). In the second part, several sensitivity studies were tested for the year 2000 based on the new data sets selected in the first part of the analysis and on the new available soil variables. Different new parameterizations for soil variables previously derived from the soil texture now based on the sand, clay and organic content of the soil as well as new parameterizations of further hydrological and thermal properties of soil were compared. In addition, due to the new non-constant soil properties, a new numerical method for calculating the soil temperatures of the five layers in the model was tested, which in turn necessitated further adjustments. The testing and selection of the different data sets and parameterization versions for the model according to performance was divided into three experimental plans. In the first plan, the effects of the selected topography and soil data sets were examined. The second plan dealt with the differences between the different types of parameterization of the soil variables in terms of the variables used to calculate the properties, the properties variable or constant over depth, and the method used to calculate the changes in soil temperature. The findings of these two experimental plans, which were carried out



---

for both study areas, led to further parameterization changes in the third plan. All changes in this third experimental plan were tested successively, so the pairwise comparison of two consecutive model runs reflects the impact of the innovation in the second run. The final part of the analysis consists of five longer model runs (2000-2018), which were carried out to review the results of the sensitivity studies and to assess the performance under other, sometimes extreme, atmospheric conditions. For this purpose, the previous model version of REMO (id01) for the two study areas (EUR-44 and GER-11) served as reference runs. Two new model versions (GER-11 of id06 and id15a) were selected on the basis of the comparison results of the third experimental plan and the final version (GER-11 of id18a) which contains all changes made in this work was also chosen for a detailed analysis.

Taken together the results show that both the new topography data and the new soil data differ crucially from the previous data sets in REMO. In addition, the auxiliary variables derived from these constant input data change significantly depending on the parameterization used, especially for the soil parameters. Both the spatial distribution and the range of values of the different model versions differ greatly. However, a quality assessment of the parameterization is difficult because different soil data sets used for the validation of the parameters also differ significantly. The final model version (id18a) is similar to the results of the previous REMO version in most variables, despite the extensive changes of the input data and parametrizations. Depending on temporal and spatial aggregation as well as different regions and seasons, slight improvements have been observed, but also slight deterioration compared to the climatological validation data. In the deeper soil layers larger changes could be identified compared to the previous model version, which could not be assessed due to a lack of validation data. Overall, there was also a slight warming of all 2 m temperatures compared to the previous model run, which on the one hand has a negative effect on the already too high minimum temperature, but on the other hand has a positive effect on the previously too low maximum temperature of the model in the study areas. No significant changes could be detected in the precipitation signal and in the 10 m wind variables, although the new topography differs significantly from the previous topography at some points in the test area. Furthermore, the ranking of the different model versions varied according to the quality index applied.

To evaluate the results it has to be considered that the new data were tested for model regions with 50 and 12 km spatial resolution and the associated hydrostatic model version. The so far already included data are suitable for aggregation to this coarser model resolution, especially in the case of topography (GTOPO with 1 km resolution). However, the previous soil data already reach their limits with 50 km resolution. In addition, it should be noted that not only the mean values of these data, but also their subgrid variability are used as variables in the model for different parameterizations. Therefore, it is essential that the input data provide a significantly higher resolution than the resolution defined for modeling. Vertical fluxes (non-hydrostatic model version) play an important role in local climate simulations with resolutions in the low kilometre range, which are strongly influenced by the topography and its horizontal and vertical change rate, which may make the much higher resolution data incorporated in this work valuable for the future development of REMO.

# 1 Überblick über das Themenfeld Klimamodelle sowie Hinführung zum Thema und Aufbau der Arbeit

## 1.1 Allgemeine Informationen zu Klimamodellen

Die Entwicklung des Klimas ist seit langer Zeit ein wichtiger Gegenstand der Forschung. Bereits seit Ende des 19. Jahrhunderts werden weltweit Messdaten erhoben und Änderungen im Klima nachgewiesen (Le Treut et al., 2007). Um die Frage zu klären, wie sich das Klima in Zukunft nicht nur im Hinblick auf durchschnittliche Temperatur- und Niederschlagswerte, sondern auch hinsichtlich der Häufigkeit und Dauer von Extremereignissen wie Hochwasser oder Dürre entwickeln wird, müssen Klimamodelle herangezogen werden. Anders als Wettervorhersagemodelle, deren Ziel möglichst gute Vorhersagen für das Wetter an einem bestimmten Ort innerhalb der nächsten Tage sind, sind Klimamodelle auf die Erforschung langer Zeitreihen ausgelegt, insbesondere auf deren mittleren Zustand und dessen Variabilität. Trotzdem sind viele Klimamodelle auf der Basis von Wettermodellen entwickelt worden (Le Treut et al., 2007; National Research Council, 2012).

Klimamodelle sind vereinfachte mathematische Abbildungen des komplexen irdischen Klimasystems. Um die Vergangenheit rekonstruieren und sich auf zukünftige Veränderungen des Klimas vorbereiten zu können, ist die Weiterentwicklung von existierenden Klimamodellen sehr wichtig. Wechselwirkungen und Rückkopplungen zwischen den einzelnen Komponenten spielen eine wichtige Rolle im Klimasystem. Ziel der Klimamodellentwicklung ist es, möglichst viele dieser relevanten Prozesse in ein Modell zu integrieren, um die Effekte aufeinander besser untersuchen zu können. Sind die wichtigen Komponenten wie Atmosphäre, Lithosphäre, Pedosphäre, Biosphäre, Hydrosphäre, Kryosphäre und Anthroposphäre enthalten, wird von einem Erdsystemmodell gesprochen. Wie Abbildung 1 zeigt, sind im Lauf der letzten Jahrzehnte immer neue Modellkomponenten zum anfänglichen Atmosphärenmodell dazugekommen. Teile von spezialisierten Modellen wie Ozeanzirkulationsmodellen, Landoberflächenmodellen oder Kohlenstoffkreislaufmodellen wurden nach und nach mit den Klimamodellen gekoppelt (R. T. Watson & Core Writing Team, 2001). Eine ausführliche Beschreibung der Geschichte der Klimamodelle kann bei Randall et al. (2007) und Edwards (2011) nachgelesen werden.

Die globalen Klimamodelle (GCMs) sind allerdings wegen ihrer geringen räumlichen Auflösung (ca. 200 km Gitterbreite) für viele Fragestellungen ungeeignet (Davin et al., 2011). Kleinräumige Effekte, die z. B. durch variable Orographie oder spezielle Landoberflächen bedingt werden, können darin nicht dargestellt werden (Y. Wang et al., 2004). Deshalb gibt es verschiedene regionale Klimamodelle (RCMs), die für Ausschnitte der Erdoberfläche höher aufgelöste Daten liefern können. Globale Klimamodelle können nämlich nicht mit höheren Auflösungen gerechnet werden, weil die Modellierung trotz der sich immer weiter entwickelnden Technik mit schnelleren

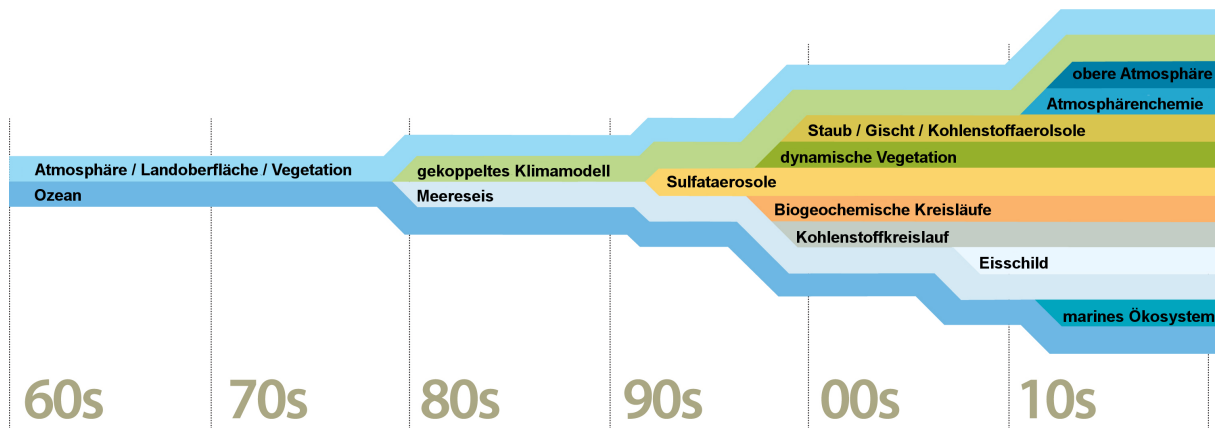


Abb. 1: Schaubild zur Entwicklung der Klimamodellkomponenten in den letzten Jahrzehnten (verändert nach University Corporation for Atmospheric Research, 2011)

Rechnern und höherer Speicherkapazität immer noch limitiert ist. Um die benötigten Klimadaten für die ganze Welt in ausreichender Auflösung bereitstellen zu können, hilft man sich daher mit dem sogenannten „Downscaling“ (Regionalisierung). Dadurch ist es möglich, aus gröber aufgelösten Daten eine sehr viel höhere Auflösung für einzelne Teilgebiete zu erreichen (Jacob et al., 2017).

Es existieren statistische, dynamische sowie Mischformen dieser Downscaling-Methoden. Beim statistischen Downscaling wird ein mathematischer Zusammenhang zwischen zwei Datensätzen erzeugt, um anschließend diesen Zusammenhang auf einen weiteren Datensatz anzuwenden. Wie der Zusammenhang ermittelt und welche Kombination von Modell- und Beobachtungsdaten benötigt wird, ist von der angewendeten Methode abhängig. Der Nachteil des statistischen Downscalings ist, dass die Qualität des Downscalings stark von den verwendeten Trainingsdaten abhängt und dass die Annahme getroffen werden muss, dass sich dieser zugrundeliegende statistische Zusammenhang über die Zeit nicht ändert. Des Weiteren können statistische Methoden nur in den Teilen des Wertebereichs gut performen, in denen sie auch trainiert und kalibriert wurden. Eine Übersicht der verschiedenen Methoden sowie deren Vor- und Nachteile liefern beispielsweise Maraun & Widmann (2018).

Dem gegenüber steht das erwähnte dynamische Downscaling, das auf physikalischen Grundlagen basiert. Die Informationen eines gröber aufgelösten Modellgitters werden von einem höher aufgelösten Modell für eine ausgewählte Teilregion verwendet (Giorgi, 1990). Das in dieser Arbeit verwendete deutsche Regional Modell (REMO) nutzt beispielsweise dieses dynamische Downscaling. Die Einbettung von einem höher aufgelösten, kleineren Modellgebiet in ein gröber aufgelöstes, größeres Modellgebiet - Nesting genannt - erfolgt numerisch gesehen am Rand des Modellgebiets (Giorgi, 2019). Die Daten des Globalmodells werden dazu genutzt, ein höher aufgelöstes Regionalmodell anzutreiben, indem sie die äußeren Rahmenbedingungen vorgeben. Da das Downscaling einen Auflösungsfaktor (Reduktionsverhältnis der horizontalen Auflösung von Antriebs- und Modelldaten) von 12 nicht übersteigen sollte, werden durch das Nesting horizontale Auflösungen von 10–50 km erreicht (Giorgi, 2006; Giorgi, 2019). Durch sogenannte Mehrfachnestungen, deren Einfluss auf Modellergebnisse von z. B. Brisson et al. (2016) oder

Matte et al. (2016) untersucht wurde, können jedoch noch höhere Modellauflösungen erreicht werden.

Um die physikalischen Prozesse im irdischen Klimasystem beschreiben zu können, sind viele teilweise nichtlineare Differentialgleichungen zu lösen. Oft existieren keine analytischen Lösungen für diese Gleichungssysteme, weshalb die Lösungen numerisch approximiert werden müssen. Außerdem müssen in allen Modellen aufgrund der begrenzten Rechenkapazität räumliche und zeitliche Diskretisierungen vorgenommen werden, bei denen der Raum abhängig von der Modellart in unterschiedlich große Volumeneinheiten und dementsprechende Zeitschritte unterteilt wird. Je höher die räumliche Auflösung, desto höher muss auch die zeitliche Frequenz der Neuberechnung der Variablen im Modell sein, sodass der Rechenaufwand für höhere Auflösungen mehr als linear ansteigt. Der Mehrwert an Genauigkeit und Qualität, der durch das Betreiben eines neueren Modells im Vergleich zum alten Modell erreicht wird, wird Added Value (AV) genannt. Er umfasst das Abwägen der Kosten und Nutzen, die durch das Betreiben eines höher aufgelösten Modells entstehen. Die Definition und die Abschätzung der Größe dieses AV ist Gegenstand von verschiedenen Studien (z. B. Di Luca et al., 2015; Rummukainen, 2016 oder Lloyd et al., 2020). Trotz der immer höher aufgelösten Modelle finden manche wichtige physikalische Prozesse auf noch kleineren Skalen als auf der jeweiligen Modellauflösung statt. Sie müssen durch eine vereinfachte Abbildung auf der entsprechenden Modellskala parametrisiert werden. Diese Parametrisierungen zählen jedoch zu den Hauptfehlerquellen der Modelle (Prein et al., 2015) und müssen daher sorgfältig ausgewählt werden.

Die ständige Weiterentwicklung von Klimamodellen, das Verbessern der Parametrisierungen und das Erforschen der Wechselwirkungen verschiedener Modellkomponenten auf den unterschiedlichen Skalen sind die Herausforderungen der Klimamodellierung (Giorgi, 2019). Je besser und genauer Klimamodelle die Zukunft abbilden können, desto leichter können Maßnahmen und Vorkehrungen zur Eindämmung sowie Anpassungen im Bezug auf den Klimawandel getroffen werden. Außerdem erlauben diese Modelle, dass verschiedene Testszenarien auf ihre jeweiligen Auswirkungen auf das Klima analysiert werden können. Die Daten der Klimamodelle sind somit ein äußerst wichtiges Werkzeug, das Entscheidungsträgern unterschiedlicher Branchen wie zum Beispiel aus Politik oder Land- und Forstwirtschaft zur Planung zur Verfügung steht.

## 1.2 Projektbeschreibung und Aufbau dieser Arbeit

Im Rahmen des Kooperationsprojekts REMOLAND vom Deutschen Klimaservicezentrum (GERICS) und der Universität Würzburg, welchem diese Arbeit zugrunde liegt, soll das Klimamodell REMO einen weiteren Schritt hin zu einem regionalen Erdsystemmodell entwickelt werden. Es wird eine Verbesserung hinsichtlich der räumlichen Auflösung der Eingangsdaten stattfinden, um künftig besser aufgelöste Klimamodellläufe starten zu können. Parallel wird die von Wilhelm et al. (2011) ausgearbeitete Erweiterung von REMO um eine dynamische Vegetationsbedeckung (REMO-iMOVE) an die aktuelle REMO-Version 2015 gekoppelt (Pollinger et al., 2020). Zusätzlich wird das neue bodenhydrologische Schema von Hagemann & Stacke (2015) (REMO-5Layer) auf-

bauend auf den Änderungen, die im Rahmen des Projekts durchgeführt werden, weiter verbessert (Abel et al., 2020).

Das Hauptziel dieser Arbeit ist, neue räumlich und inhaltlich höher aufgelöste Datensätze in REMO einzubauen. Mit der Intention der zunehmend höheren Modellauflösungen werden auch höhere Anforderungen an die Eingangsdaten der Modelle gestellt. Diese Eingangsdaten, die als Oberflächendaten die unteren Rahmenbedingungen des Modells (engl. boundary conditions) darstellen, haben einen großen Einfluss auf die Regionalmodelle (z. B. Henderson-Sellers & Dickinson, 1993; Dickinson, 1995). Eine gute Repräsentation der tatsächlichen Oberflächenbedingungen, die den gesamten Energieaustausch zwischen Oberfläche und Atmosphäre steuern, ist daher noch wichtiger für höhere Auflösungen. Durch systematisches Testen wird in dieser Arbeit auf den bisherigen Modellauflösungen untersucht, wie diese neuen konstanten Eingangsdaten die Modellergebnisse bereits auf dieser Raumskala beeinflussen. Basierend auf den neuen Datensätzen werden in weiteren Schritten neue Bodenparametrisierungen und deren Auswirkungen in Sensitivitätsstudien getestet.

In den nachfolgenden Kapiteln werden zunächst einige theoretische Grundlagen der Landoberflächenprozesse erläutert und die für diese Arbeit wichtigen Variablen definiert. Anschließend wird das regionale Klimamodell REMO mit seinen allgemeinen Eckdaten und den Modellkomponenten ausführlich vorgestellt (vgl. Kapitel 3). Danach werden in Kapitel 4 die neuen Eingangsdaten, deren Aufbereitung und Weiterverarbeitung sowie die zur Validation benutzten Beobachtungsdaten erklärt. Im darauffolgenden Kapitel 5 werden die vorgenommenen Änderungen am Modellcode dargestellt. Der Fokus liegt hierbei auf der Herleitung und dem Vergleich der Kennziffern und der geänderten Parametrisierungen des Modellbodens. In Kapitel 6 folgt zunächst eine methodische Einführung der verwendeten Analysen. Danach werden die Ergebnisse und Auswirkungen der Modelländerungen zum einen im Vergleich zur bisherigen Modellversion und zum anderen im Vergleich zu Beobachtungsdaten anhand verschiedener Variablen gezeigt. Abschließend werden in Kapitel 7 und 8 die Ergebnisse diskutiert und ein Ausblick gegeben. Da in dieser Arbeit viele unterschiedliche Forschungsfelder und Teilbereiche enthalten sind, werden der jeweilige Forschungsstand und die Ergebnisse der Literaturrecherche im jeweiligen Kapitel erörtert.

## 2 Theorie und Grundlagen der Modellierung von Landoberflächenprozessen

In diesem Kapitel werden die grundlegenden Zusammenhänge, die bei der Modellierung des Systems Atmosphäre-Boden eine Rolle spielen, erläutert. Außerdem werden die benötigten Parameter definiert und deren Verwendungszweck sowie verschiedene Ansätze von Parametrisierungen dargestellt. Bei einer Parametrisierung handelt es sich um eine Annäherung durch Formeln mit empirischen Parametern an subskalige Prozesse, die nicht auf der Raumskala des jeweiligen Modells abgebildet werden (Rummukainen, 2010).

Bei der Entwicklung der Klimamodelle hin zu Erdsystemmodellen werden nach und nach neue Prozesse und deren Wechselwirkungen mit bereits bestehenden Systemen im Modell eingebaut (vgl. Abb. 1) (Kawamiya et al., 2020). Aufgrund der komplexen Landoberfläche-Atmosphären-Interaktion ist eine gute Abbildung des Austausches von Energie, Momentum und Feuchtigkeit besonders wichtig für eine realistischere Darstellung des Klimas (Y. Wang et al., 2004). Eine gute Richtlinie der umfänglichen Prozesse dieser Wechselwirkung und deren Parametrisierung liefern hydrologische Modelle und Landoberflächenmodelle (engl. Land Surface Models) (LSM). Sie modellieren die Wasser- und Energieflüsse zwischen der Landoberfläche und der Atmosphäre auf unterschiedlichen räumlichen Skalen (Niu & Zeng, 2012; Tafasca et al., 2020). Pitman (2003) erklärt in seinem Review-Artikel zur Entwicklung der LSMs anhand vieler Beispielstudien, dass diese Prozesse nicht nur auf kleinen zeitlichen und räumlichen Skalen einen großen Effekt auf die Atmosphäre haben, sondern auch für Klimamodelle auf regionaler Skala einen wichtigen Einfluss auf die Interaktion zwischen Oberfläche und Atmosphäre haben. Die Parametrisierung dieser Landoberflächenprozesse hat sich von einfachen Oberflächenenergiebilanzmodellen (vgl. Manabe, 1969) über mit der Atmosphäre interagierende Modelle (vgl. Sellers et al., 1997) zu eigenen komplexen Modellen mit vielen verschiedenen simulierten Prozessen wie beispielsweise des Carbonkreislaufs oder dynamischer Vegetation entwickelt. Diese Einteilung wird auch als 1., 2. und 3. Modellgeneration an LSMs bezeichnet. Die Kopplung der 3. Modellgeneration LSMs mit Klimamodellen erfordert eine multidisziplinäre Zusammenarbeit und ist Gegenstand vieler laufenden Studien und Projekte. Weitere Informationen zur Geschichte und Entwicklung von LSMs sowie ein Überblick verschiedener Parametrisierungen sind unter anderem bei Z.-L. Yang (2004) zu finden. In den folgenden Abschnitten werden die Grundlagen an Prozessen und einige Parameter für Klimamodelle erläutert.

### 2.1 Übersicht der Energieflüsse

Die kurz- und langwelligen Strahlungsflüsse sowie die Strahlungsbilanz gehören zu den wichtigsten Faktoren in einem Klimamodell, denn sie beschreiben die Energiequellen und -senken im System und bilden die auf verschiedenen zeitlichen Skalen stattfindenden Energieflusszyklen ab (Kothe

et al., 2011). Aus der Annahme, dass alle Energieflüsse im Modell nur vertikal stattfinden, lässt sich die Oberflächenenergiebilanzgleichung 2.1a aufstellen.

$$R_n = R_{s,n} + R_{l,n} = (R_{s,d} + R_{s,u}) + (R_{l,d} + R_{l,u}) \quad (2.1a)$$

$$R_{s,n} = R_{s,d} + R_{s,u} = (1 - \alpha) \cdot R_{s,d} \quad (2.1b)$$

$$\alpha = \frac{R_{s,u}}{R_{s,d}} \quad (2.1c)$$

$$R_{l,n} = R_{l,d} + R_{l,u} \quad (2.1d)$$

mit:

$R_{s,d}$  = eingehende kurzwellige solare Strahlung (positiv), Globalstrahlung

$R_{s,u}$  = ausgehende kurzwellige solare Strahlung (negativ), Reflexstrahlung

$R_{l,d}$  = eingehende langwellige thermale Strahlung (positiv), Gegenstrahlung

$R_{l,u}$  = ausgehende langwellige thermale Strahlung (negativ), Ausstrahlung

$R_{s,n}$  = kurzwellige solare Nettostrahlung

$R_{l,n}$  = langwellige thermale Nettostrahlung

$\alpha$  = Albedo

Der Strahlungshaushalt der Erde ist in Abbildung 2 schematisch dargestellt. Ein Teil der kurzwelligen Strahlung der Sonne, die oberhalb der Atmosphäre ca.  $342 \frac{\text{W}}{\text{m}^2}$  beträgt, trifft, nachdem ein Teil an Wolken und Aerosolen reflektiert wurde, auf die Erdoberfläche. Diese Strahlung  $R_{s,d}$  wird auch Globalstrahlung genannt und setzt sich aus direkter und diffuser Strahlung zusammen. Die direkte Strahlung ist der Anteil, der ungehindert durch die Atmosphäre direkt auf den Erdboden trifft, während die indirekte (auch diffuse Strahlung genannt) zunächst durch Aerosole und Wolken reflektiert wurde, bevor sie am Boden ankommt. Anschließend wird ein Teil dieser kurzwelligen Strahlung in die Atmosphäre zurück emittiert ( $R_{s,u}$ ). Die Summe (bzw. Differenz wegen des negativen  $R_{s,u}$ -Wertes) zwischen eingehender und ausfallender Strahlung wird Nettosolarstrahlung an der Oberfläche (vgl. Formel 2.1b) genannt. Der Anteil an reflektierter Strahlung im Vergleich zur einfallenden Strahlung wird (effektive) Albedo genannt (vgl. Formel 2.1c). Diese Größe drückt das Reflexionsvermögen der Oberfläche aus und ist von Faktoren wie Zenitwinkel der einfallenden Strahlung, Vegetationstyp an der Oberfläche, Schneebedeckung und Bodenfeuchte abhängig. Je höher die Albedo, desto mehr Strahlung wird zurück emittiert. Der nicht reflektierte Teil der Solarstrahlung wird vom Boden absorbiert und in langwellige Wärmestrahlung umgewandelt. Beim langwelligen Summanden in Formel 2.1a ( $R_{l,d}$ ) fällt infrarote Wärmestrahlung von Wolken, Aerosolen und anderen emittierten infrarot Strahlungen ein. Ein anderer Teil dieser Strahlung wird als langwellige Wärmestrahlung ( $R_{l,u}$ ) in die Atmosphäre abgegeben. Diese langwellige emittierte Strahlung wird als Ausstrahlung

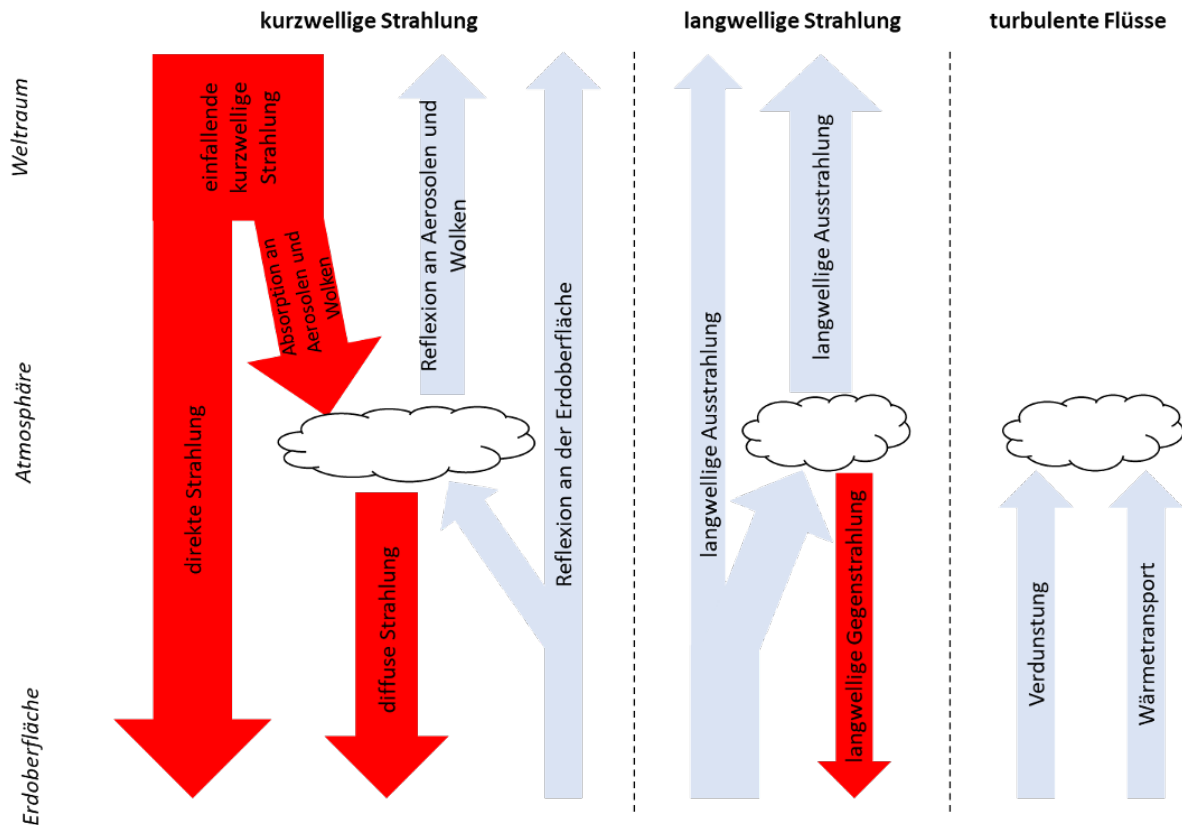


Abb. 2: Schematische Übersicht des Strahlungshaushaltes der Erde (verändert nach: Fortak, 1982, Kiehl & Trenberth, 1997 und Pitman, 2003)  
 rote Pfeile: Energieeintrag (positiver Wert)  
 blaue Pfeile: Energieverlust (negativer Wert)

bezeichnet. Sie ist für gewöhnlich größer als die Gegenstrahlung. Die langwellige Nettostrahlung ( $R_{l,n}$ ), welche als Summe (bzw. Differenz wegen des negativen  $R_{l,u}$ -Wertes) von einfallender und emittierter Strahlung definiert ist (vgl. Formel 2.1d), repräsentiert die Energiemenge des Wärmeverlustes der Oberfläche und hängt von der Oberflächentemperatur, der Feuchtigkeit und der Wolkenbedeckung ab (Qin et al., 2002; R. Li et al., 2016). Die langwelligeren Strahlungen beziehen sich auf die elektromagnetische Energie, die von der Erdoberfläche und der Atmosphäre emittiert wird, und werden mit dem Stefan-Boltzmann-Gesetz (vgl. Formel 2.2) bestimmt. Sie hängen daher direkt von der Temperatur und der (effektiven) Emissivität  $\varepsilon$  ab.

$$R_l = \varepsilon \cdot \sigma \cdot T^4 \quad (2.2)$$

mit:

$\sigma = 5,67 \cdot 10^{-8}$  = Stefan-Boltzmann-Konstante

$\varepsilon$  = Emissivität der Landoberfläche

$T$  = Oberflächentemperatur in  $K$

Die Energiemenge  $R_n$  kann gleichzeitig als Summe der drei Energieströme latente Wärmestromdichte (Verdunstung), sensible (oder fühlbare) Wärmestromdichte und Bodenwärmestrom



dargestellt werden (vgl. Formel 2.3). Insgesamt sind alle Summanden der Formel 2.3 von der Oberflächentemperatur abhängig, sodass diese der wichtigste Faktor zum Lösen der Gleichung ist (vgl. Qin et al., 2002). Die einzelnen Größen besitzen einen Tages- und Jahresgang und variieren je nach Region (vgl. Bonan, 2019).

$$R_n = \lambda E + H + G \quad (2.3)$$

mit:

$R_n$  = Nettostrahlungsbilanz am Boden

$\lambda E$  = latente Wärmestromdichte

$H$  = sensible Wärmestromdichte

$G$  = Bodenwärmestrom

Die Einheit aller Summanden in Formel 2.1a und 2.3 ist  $\frac{W}{m^2}$ , denn mit ihnen wird die Energie pro Fläche und Zeit ( $\frac{J}{s \cdot m^2}$ ) ausgedrückt. Diese überschüssige Energiemenge steht zum Erwärmen verschiedener Systeme wie Luft, Vegetation oder Boden zur Verfügung oder kann zum Aggregatzustandswechsel von Wasser genutzt werden. Die Vorzeichen der jeweiligen Summanden geben die Richtung des Energieflusses an. Während positive Vorzeichen einen Energieeintrag kennzeichnen, bedeutet ein negativer Wert, dass Energie entzogen wird (vgl. Abb. 2).

Die latente Energie ist die Energiemenge in der Luft, die zum Ändern des Aggregatzustandes entzogen oder freigesetzt wird. Es findet dabei keine Temperaturänderung statt. Sie ist direkt abhängig vom Wassergehalt, denn um Wasser zu verdunsten, wird Energie benötigt, um die Wasserstoffbrücken zwischen den Wassermolekülen aufzubrechen. Bei der Kondensation von Wasserdampf wird diese in der Luft „gespeicherte“ Energiemenge der sensiblen Wärme zugefügt. Das heißt, die im Wasserdampf gespeicherte latente Energiemenge wird freigesetzt und erhöht die Temperatur der Umgebungsluft. Die sensible Wärme hingegen ist die Menge an Wärme, die zur Erwärmung eines Mediums benötigt wird, sodass dessen Temperatur ansteigt. Insgesamt stellt die Summe aus sensibler und latenter Wärme den Energiegehalt der Luft dar (Bonan, 2019). Der Quotient von sensibler und latenter Wärme wird Bowen-Verhältnis (Bowen Ratio) genannt (Bowen, 1926).

Der ebenfalls in Gleichung 2.3 vorkommende Bodenwärmestrom  $G$  ist abhängig vom Temperaturgradienten und der thermalen Leitfähigkeit des Bodens. In fast allen Modellen wird er nur vertikal modelliert, da der horizontale Fluss einen vergleichsweise sehr geringen Einfluss hat. Dadurch lässt sich dieser Wärmestrom mit Formel 2.4 ausdrücken.

$$G = -\lambda \frac{T_1 - T_s}{\frac{\Delta z_1}{2}} \quad (2.4)$$

mit:

$T_s$  = Oberflächentemperatur

$T_1$  = Temperatur der obersten Bodenschicht

$\Delta z_1$  = Schichtdicke der ersten Bodenschicht

$\lambda$  = thermale Konduktivität des Bodens

Die thermale Konduktivität  $\lambda$  ist von der jeweiligen Bodenzusammensetzung abhängig und wird in Kapitel 2.2 genauer vorgestellt. Die Tatsache, dass  $G$  genauer berechnet werden kann, je kleiner  $\Delta z$  gewählt wird, führt dazu, dass in vielen Modellen mehrere Bodenschichten verwendet werden (Niu & Zeng, 2012).

## 2.2 Definition verschiedener Bodenparameter und -prozesse

Wie wichtig der Boden für die Modellierung des Klimas ist, fassen Fatichi et al. (2020) aus ihren zitierten Quellen zusammen. Über den Boden werden die atmosphärischen und hydrologischen Prozesse mit der Biosphäre und dem biochemischen Kreislauf verbunden. Bodeneigenschaften und deren Änderung beeinflussen das lokale Klima über Änderungen der Oberflächenevaporation, der Wärmeflüsse, der Extremtemperaturen sowie des gesamten hydrologischen Kreislaufs stark. Fatichi et al. (2020) merken jedoch an, dass das Nachbilden der Bodenprozesse auf den unterschiedlichen Modellskalen eine große Herausforderung ist, denn zur Berechnung der thermalen Flüsse sowie der hydrologischen Prozesse im Boden werden spezielle Bodenparameter im Modell benötigt. Diese Bodeneigenschaften beeinflussen die Geschwindigkeit und die Größe der Veränderungen von beteiligten Variablen (Sauer & Horton, 2005). Allerdings liegen für viele dieser benötigten Parameter keine direkten Messwerte oder Datensätze vor, weswegen man sie approximieren muss. Nachfolgend werden zum einen die für die jeweiligen Parametrisierungen verwendeten thermischen und hydrologischen Kennziffern sowie deren mögliche Berechnung vorgestellt und zum anderen ein kurzer Überblick über die thermischen Bodenprozesse gegeben.

### Thermische Parameter

Boden setzt sich aus unterschiedlichen Materialien zusammen. Um die Parameter des Bodens besser berechnen und abschätzen zu können, werden die meisten der nachfolgenden Eigenschaften abhängig von der Zusammensetzung aus den drei Hauptbestandteilen Feststoff, Wasser und Luft betrachtet. Dabei wird beim Wasseranteil teilweise zwischen gefrorenem und ungefrorenem Anteil und beim Feststoffanteil die Zusammensetzung aus organischen Material und sandigen bzw. tonigen Mineralien unterschieden, da diese Komponenten sehr unterschiedliche physikalische Eigenschaften aufweisen (Bonan, 2019).

**Volumetrische Wärmekapazität:** Diese Unterscheidung der Komponenten ist beispielsweise bei der volumetrischen Wärmekapazität  $c_v$  sehr wichtig. Die Wärmekapazität ist definiert als die Menge an Energie, die benötigt wird, um einen Kubikmeter Boden um 1 °C zu erwärmen. Sie wird in der Einheit  $\frac{\text{J}}{\text{m}^3 \text{K}}$  angegeben. Die volumetrische Wärmekapazität ist das Produkt der Dichte  $\rho$  und der spezifischen Wärmekapazität  $c_s$  eines Stoffes. Um die volumetrische Wärmekapazität eines Bodens zu berechnen, wird die Summe der einzelnen volumetrischen Wärmekapazitäten

der Bodenkomponenten mit dem jeweiligen Volumenanteil multipliziert (vgl. Formel 2.5) (de Vries, 1963).

$$c_v = f_{sol} \cdot c_{sol} + f_w \cdot c_w + f_i \cdot c_i + f_a \cdot c_a \quad (2.5)$$

mit:

$c_{sol}$  = volumetrische Wärmekapazität der Festanteile im Boden

$c_w = 4.18 \frac{MJ}{m^3 K}$  = volumetrische Wärmekapazität von Wasser

$c_i = 1.9 \frac{MJ}{m^3 K}$  = volumetrische Wärmekapazität von Eis

$c_a = 0.0012 \frac{MJ}{m^3 K}$  = volumetrische Wärmekapazität von Luft

$f_{sol}$  = Feststoffanteil im Boden

$f_w$  = flüssiger Anteil des Wasseranteils im Boden

$f_i$  = Eisanteil des Wasseranteils im Boden

$f_a$  = Luftanteil im Boden

In Tabelle 1 sind die Kapazitäten der wichtigsten Bestandteile angegeben. Es fällt auf, dass es einen Unterschied zwischen Wasser und Eis und einen noch größeren Unterschied zwischen Wasser und Luft gibt. Das bedeutet, je mehr Wasser im Boden enthalten ist, desto höher ist die volumetrische Wärmekapazität. Aufgrund dieser kleineren Wärmekapazität erwärmt sich trockener Boden daher schneller als feuchter Boden.

Tab. 1: Volumetrische Wärmekapazität von verschiedenen Materialien (nach: van Wijk & de Vries, 1963)

Material	Symbol	Wert in $\frac{MJ}{m^3 K}$
Bodenmineralien	$c_{sol}$	1.9
bodenorganisches Material	$c_o$	2.5
Wasser	$c_w$	4.18
Eis (0 °C)	$c_i$	1.9
Luft	$c_a$	0.0012

**Thermische Leitfähigkeit:** Ein weiterer wichtiger thermischer Bodenparameter ist die thermische Leitfähigkeit, die auch Wärmeleitfähigkeit genannt wird. Sie wird oft mit  $\lambda$  abgekürzt und beschreibt den Wärmefluss (Energienmenge pro Sekunde) durch eine 1 m dicke Schicht eines Materials, wenn eine Temperaturdifferenz von 1 K zwischen den beiden Seiten der Schicht vorherrscht. Sie wird daher mit der Einheit  $\frac{W}{mK}$  angegeben. Insgesamt ist die Wärmeleitfähigkeit bei sandigen Böden aufgrund des größeren Quarzanteils höher. Außerdem ist die Leitfähigkeit von trockenem Boden geringer, da Wasser eine viel höhere Leitfähigkeit besitzt als Luft (vgl. Tab. 2). Die Leitfähigkeit von Boden erreicht aber schon vor der Sättigung annähernd ihr Maximum, da sich ab einem kritischen Punkt keine weiteren Wasserstoffbrücken zwischen den Bodenpartikeln bilden und die Leitfähigkeit danach nur noch minimal ansteigt. Zudem hängt die Leitfähigkeit von der Temperatur des Bodens ab (Tarnawski & Gori, 2002; Eggert, 2011).

Tab. 2: Wärmeleitfähigkeit von verschiedenen Materialien (nach de Vries, 1975 und Bachmann, 2014)

Material	Symbol	Wert in $\frac{W}{mK}$
Basalt	$\lambda_B$	2.20
Quarz	$\lambda_Q$	8.80
Ton	$\lambda_C$	2.92
organisches Material	$\lambda_o$	0.25 bzw. 0.2-0.7
Wasser	$\lambda_w$	0.57
Eis (0°C)	$\lambda_i$	2.18
Luft (20°C)	$\lambda_a$	0.025

**Thermische Diffusivität:** Die thermische Diffusivität  $\alpha$  ist der dritte wichtige Parameter zur Beschreibung der thermalen Eigenschaften einer Substanz. Sie wird auch Temperaturleitfähigkeit genannt und stellt die Eindringgeschwindigkeit der Wärme in der Einheit  $\frac{m^2}{s}$  dar. Sie kann durch die Wärmeleitfähigkeit und -kapazität berechnet werden (vgl. Formel 2.6). Aufgrund der Tatsache, dass die Wärmeleitfähigkeit nicht so stark vom Wassergehalt abhängt wie die Wärmekapazität, wird die maximale thermische Diffusivität bereits bei einem mittleren Bodenwassergehalt erreicht (Ochsner et al., 2001; Bonan, 2019).

$$\alpha = \frac{\lambda}{\rho \cdot c_s} = \frac{\lambda}{c_v} \quad (2.6)$$

mit:

$\alpha$  = thermische Diffusivität eines Stoffes in  $\frac{m^2}{s}$

$\lambda$  = thermische Konduktivität eines Stoffes in  $\frac{W}{mK}$

$\rho$  = Dichte eines Stoffes in  $\frac{kg}{m^3}$

$c_s$  = spezifische Wärmekapazität eines Stoffes in  $\frac{J}{kgK}$

$c_v$  = volumetrische Wärmekapazität eines Stoffes in  $\frac{J}{m^3K}$

### Hydrologische Parameter

Neben den thermischen Eigenschaften sind auch die hydrologischen Eigenschaften ein weiterer wichtiger Faktor bei der Modellierung des Zusammenspiels zwischen den Systemen Boden und Atmosphäre sowie bei der Steuerung der Prozesse innerhalb des Bodens. Dabei beeinflussen diese nicht nur die hydrologischen Prozesse, sondern steuern indirekt auch die Geschwindigkeit und die Menge der thermalen Energieflüsse im Boden. Nachfolgend werden die wichtigsten Kennziffern, die auch in dieser Arbeit analysiert und verändert werden, definiert und erläutert.

**Porenvolumen und Porenziffer:** Das Porenvolumen, auch Porosität oder „void space“ bzw. Hohlraumvolumen genannt, stellt den volumetrischen Anteil von Boden dar, der nicht mit Feststoff gefüllt ist. Dieses Volumen kann mit Wasser oder Luft gefüllt sein (vgl. Formel 2.7). Als Einheit wird meistens  $\frac{m^3}{m^3}$  oder Prozent verwendet. Das Porenvolumen wird in dieser Arbeit als PORVOL abgekürzt und in der Literatur meist mit  $n$  oder  $\epsilon$  bezeichnet. Sandige Böden haben

ein geringeres Porenvolumen als feinkörnigere Böden.

$$PORVOL = \frac{V_{space}}{V_{ges}} = \frac{V_w + V_i + V_a}{V_w + V_i + V_a + V_{sol}} \quad (2.7)$$

$$e = \frac{V_{space}}{V_{sol}} \quad (2.8)$$

mit:

$V_{space}$  = Hohlraumvolumen

$V_{sol}$  = Volumen der Feststoffe

$V_w$  = Volumen vom flüssigen Wasser

$V_i$  = Volumen von Eis

$V_a$  = Volumen der Luft

$V_{ges}$  = Gesamtvolumen

Die sogenannte Porenziffer oder -zahl  $e$  (vgl. Formel 2.8) zeigt hingegen das Verhältnis des Hohlraumvolumens zum Feststoffvolumen an. Dadurch ergeben sich folgende Umrechnungsmöglichkeiten:

$$\epsilon = \frac{e}{e + 1} \quad \text{und} \quad e = \frac{\epsilon}{1 - \epsilon}$$

Der Wert des Porenvolumens stellt gleichzeitig auch den Wassergehalt eines vollständig gesättigten Bodens dar und wird deshalb in der Literatur oft mit  $\theta_s$  oder  $\theta_{sat}$  bezeichnet. Vor allem bei der Darstellung der sogenannten Water Retention Curve (WRC) (auch Wasserretentionscharakteristiken, Retentionskurve oder Wasserspannungscharakteristik genannt), welche den Zusammenhang von einem Matrixpotential  $h$  und einem volumetrischen Wassergehalt  $\theta$  beschreibt, wird  $\theta_s$  als Kurven beschreibender Parameter benötigt. Ausgehend von im Labor ermittelbarer Größen kann das Porenvolumen mit Formel 2.9 berechnet werden.

$$PORVOL = 1 - \frac{p_b}{p_m} \quad (\text{optional} \cdot 100 \text{ für } \%) \quad (2.9)$$

mit:

$p_b$  = Trockenrohdichte des Bodens (auch Lagerungsdichte oder engl. bulk density genannt)

= Masse der festen Bestandteile pro Gesamtbodenvolumen in  $(\frac{kg}{m^3})$

$p_m$  = spezifische Dichte des Bodens (engl. particle density)

= Masse Boden pro Festsubstanzvolumen in  $(\frac{kg}{m^3})$

**Feldkapazität:** Ähnlich dem Porenvolumen bezeichnet die Feldkapazität einen speziellen Wassergehalt des Bodens. Der Begriff Feldkapazität (in REMO FCAP) wurde von Veihmeyer & Hendrickson (1931) wie folgt beschrieben: „the amount of water held in the soil after the excess gravitational water has drained away and after the rate of downward movement of water has

materially decreased (...).“ In Amelung et al. (2018) wird die Feldkapazität als der Wassergehalt definiert, bei dem ein Gleichgewicht in der Verteilung des Bodenwassers nach ca. 1-2 Tagen vermutet werden kann. Noch einfacher drückt es die Bodenkundliche Kartieranleitung (vgl. Finnern, 1994) aus, indem sie die Feldkapazität als „die Wassermenge, die ein Boden gegen die Schwerkraft zurückhalten kann“, bezeichnet, was der Definition von Veihmeyer & Hendrickson (1931) gleichkommt. Formal wird die Feldkapazität häufig als der Wassergehalt bei 33 kPa Matrixpotential definiert, wobei in der Literatur auch Werte von 10 kPa vor allem bei sandig lehmigen Böden angegeben sind. Zudem finden sich auch Angaben als „Saugspannung“ pF, die den dekadischen Logarithmus des negativen Matrixpotentials beschreibt (van Vlaardingen & Smit, 2009). Hier werden Werte von 1.8 (für grundwassernahe Böden) und 2.5 (für grundwasserferne Böden) für die Feldkapazität benutzt. Eng mit diesem Parameter verknüpft ist die Luftkapazität, die den volumetrischen Gehalt an Luft bei Feldkapazität, also die Differenz von Porenvolumen und Feldkapazität, ausdrückt (Finnern, 1994).

**Permanenter Welkepunkt:** Der permanente Welkepunkt (PWP) ist je nach Studie als der Wassergehalt oder das Matrixpotential definiert, ab dem eine Pflanze zu welken beginnen, weil zu wenig Wasser im Boden verfügbar ist. Eine häufig verwendete Indikatorpflanze ist dabei die Sonnenblume (Briggs & Shantz, 1912; Veihmeyer & Hendrickson, 1928). In dieser Arbeit wird PWP, in Anlehnung an die bisherige Definition von REMO, als der Wassergehalt festgelegt, bei dem das entsprechende Matrixpotential erreicht wird. Manchmal wird dieser Wassergehalt daher analog zur Feldkapazität als Welkepunkt(wasser)kapazität bezeichnet. In der Literatur wird zumeist ein Matrixpotential von 1500 kPa bzw. eine Saugspannungswert von 4.2 für PWP verwendet. Die Differenz von Feldkapazität und permanentem Welkepunkt wird als pflanzenverfügbares Bodenwasser (engl. plant-available soil water, WAVA) oder nutzbare Feldkapazität (nFK) bezeichnet (Blanc & Milton L., 1957; Durner & Iden, 2015; Leenaars et al., 2015; Kushwa et al., 2016; Amelung et al., 2018; Stahr et al., 2020).

**Umrechnung der Wassergehalteinheiten:** Der Bodenwassergehalt sowie die eben vorgestellten spezifischen Wassergehalte PORVOL, FCAP und PWP können entweder gravimetrisch ( $\theta_g$  in  $\frac{g_{Wasser}}{g_{Boden}}$ ) oder volumetrisch ( $\theta_v$  in  $\frac{m^3_{Wasser}}{m^3_{Boden}}$ ) angegeben werden. Zur Umrechnung von gravimetrisch in volumetrisch wird die Trockenrohddichte ( $\rho_b$ ) benötigt (vgl. Formel 2.10) (Bonan, 2019).

$$\theta_v = \theta_g \cdot \frac{\rho_b}{\rho_w} \quad (2.10)$$

mit:

$\theta_v$  = volumetrischer Wassergehalt in  $\frac{cm^3}{cm^3}$  bzw.  $\frac{m^3}{m^3}$

$\theta_g$  = gravimetrischer Wassergehalt in  $\frac{g}{g}$  bzw.  $\frac{kg}{kg}$

$\rho_b$  = Trockenrohddichte des Bodens in  $\frac{g}{cm^3}$  bzw.  $\frac{kg}{m^3}$

$\rho_w = 1 \frac{g}{cm^3} = 1000 \frac{kg}{m^3}$  = Dichte von Wasser

### Thermische Bodenprozesse

Die Bodentemperatur ist gekennzeichnet von Tages- und Jahresgängen. Da die Wärmezufuhr in den Boden hauptsächlich durch die Sonneneinstrahlung stattfindet, sind an der Oberfläche die Bodentemperaturen in den Mittagsstunden im Sommer am höchsten. Je tiefer jedoch eine Bodenschicht liegt, desto mehr kehrt sich der Jahresgang im Vergleich zum Jahresgang der Lufttemperatur um. Das heißt, in tieferen Bodenschichten werden die höchsten Temperaturen erst im Herbst oder Winter erreicht. Außerdem nimmt die Amplitude der Bodentemperatur mit zunehmender Tiefe ab. Die Bodentemperatur steht zudem in Wechselwirkung mit dem Wassergehalt und dessen Verteilung im Boden. Der Wärmetransport in einem porösen Medium (wie dem Boden) wird auch durch den Temperaturgradienten bestimmt.

Im Allgemeinen existieren drei verschiedene Arten von Wärmetransport:

- Konduktion: Wärmetransfer zwischen Teilchen unterschiedlicher Temperatur (Wärmeleitung)
- Konvektion: Wärmetransfer zwischen verschiedenen Teilen des Mediums durch sich bewegendes Fluidum oder Gase (Wärmeströmung)
- Strahlung: Wärmetransfer zwischen kalten und warmen Regionen im Medium mittels elektromagnetischer Wellen (Infrarotstrahlung)

Bei der Wärmeleitung wird die Wärme mittels schwingender Atome durch das Medium transportiert, ohne dass sich die Teilchen selbst von der Stelle bewegen. Der Wärmestrom hingegen beschreibt den Wärmetransport in einem strömenden Medium. Daher ist die Wärmeströmung bei Festkörpern nicht möglich, da diese ortsgebunden sind. Beide Wärmeübertragungsarten sind abhängig vom jeweiligen Material. Der Wärmetransfer aufgrund von Strahlung ist bei der Modellierung der Bodenprozesse nur in großporigen trockenen Böden bei hohen Temperaturen relevant und wird daher in den Modellen meist vernachlässigt (de Vries & Afgan, 1975; Hubrechts & Feyen, 1996).

Veränderungen der Wärmemenge  $F$  durch Wärmeleitung kann vom ersten Gesetz der Thermodynamik und dem Fourierschen Gesetz abgeleitet werden. Daraus ergibt sich folgende Gleichung 2.11. Der Wärmefluss ist dem Temperaturgradienten entgegengesetzt, d. h. die Energie bewegt sich immer vom wärmeren zum kälteren Bereich. Die Geschwindigkeit oder Stärke des Flusses ist dabei sowohl von der Temperaturdifferenz als auch von der materialabhängigen thermischen Konduktivität  $\lambda$  abhängig.

$$F(\text{Raum}, \text{Zeit}) = -\lambda \cdot \nabla T(\text{Raum}, \text{Zeit}) \quad (2.11)$$

mit:

$F$  = Wärmefluss(dichte) in Abhängigkeit von Raum und Zeit in  $\frac{W}{m^2}$

$\lambda$  = thermische Konduktivität in  $\frac{W}{mK}$

$\nabla T$  = Temperaturgradient in  $\frac{K}{m}$

Da der Wärmefluss in die horizontalen Richtungen aufgrund der Modellauflösung um ein Vielfaches kleiner ist, wird in regionalen Klimamodellen meist nur der vertikale Wärmetransport abgebildet. Dadurch vereinfacht sich Formel 2.11 zu Formel 2.12.

$$F = -\lambda \cdot \frac{\delta T}{\delta z} \quad (2.12)$$

mit:

$F$  = vertikaler Wärmefluss in  $\frac{W}{m^2}$

$\lambda$  = thermische Konduktivität in  $\frac{W}{mK}$

$\delta T$  = Temperaturdifferenz in  $K$

$\delta z$  = Schichtdicke in  $m$

Aufgrund der Energieerhaltung folgt für die Änderung des Wärmeflusses mit der Tiefe, dass diese mit dem Produkt aus volumetrischer Wärmekapazität und Temperaturänderung pro Zeit gleichzusetzen ist (vgl. Formel 2.13).

$$\rho \cdot c_s \cdot \frac{\delta T}{\delta t} = -\frac{\delta F}{\delta z} = \frac{\delta}{\delta z} \left( \lambda \cdot \frac{\delta T}{\delta z} \right) \quad (2.13)$$

mit:

$\rho$  = Bodendichte in  $\frac{kg}{m^3}$

$c_s$  = spezifische Wärmekapazität des Bodens in  $\frac{J}{kgK}$

$\delta T$  = Temperaturdifferenz in  $K$

$\delta t$  = Zeitschritt in  $s$

$F$  = vertikaler Wärmefluss in  $\frac{W}{m^2}$

$\delta z$  = Schichtdicke in  $m$

$\lambda$  = thermische Konduktivität in  $\frac{W}{mK}$

Bei eindimensionalen vertikalen Wärmeflüssen in einem Boden, bei dem sich die Wärmeleitfähigkeit nicht mit der Bodentiefe ändert, kann damit die Temperaturänderung pro Zeit mit der vereinfachten Formel 2.14 ausgedrückt werden (Bonan, 2019).

$$\frac{\delta T}{\delta t} = \frac{\lambda}{\rho \cdot c_s} \cdot \frac{\delta^2 T}{\delta^2 z} = \alpha \cdot \frac{\delta^2 T}{\delta^2 z} \quad (2.14)$$

mit:

$\delta T$  = Temperaturdifferenz in  $K$

$\delta t$  = Zeitschritt in  $s$

$\lambda$  = thermische Konduktivität in  $\frac{W}{mK}$

$\rho$  = Bodendichte in  $\frac{kg}{m^3}$

$c_s$  = spezifische Wärmekapazität des Bodens in  $\frac{J}{kgK}$

$\delta z$  = Schichtdicke in  $m$

$\alpha$  = thermische Diffusivität in  $\frac{m^2}{s}$



Veränderungen der Wärmemenge im Boden erfolgen durch Wärmeleitung, Konvektion, Änderungen im Wassergehalt und Phasenübergänge. Weil die beiden Prozesse Wärmeleitung und -transport für den Boden nicht getrennt werden können, liegen empirische Werte als „effektive Wärmeleitfähigkeit“ vor (Trinks, 2010). Die Wärmeleitfähigkeit gilt als eine der wichtigsten Bodeneigenschaften in LSMs. Sie wird zur Steuerung des Wärmeaustausches benötigt und beeinflusst neben den Temperaturen auch das Wasserbudget (R. Li et al., 2019). Die Wärmeleitfähigkeit kann auf unterschiedliche Weise berechnet werden und hängt von vielen Einflussfaktoren wie dem Bodentyp und damit Bodendichte, -porosität, -zusammensetzung sowie von Wassergehalt und Temperatur ab (Balland & Arp, 2005). Dong et al. (2015) unterscheiden beispielsweise nach Mixing-Modellen sowie empirischen und mathematischen Modellen. Die erste Gruppe umfasst alle Modelle, bei denen die thermale Leitfähigkeit anhand der Werte der einzelnen Komponenten und in Abhängigkeit der Verteilung approximiert wird. Bei den empirischen Modellen wird die Konduktivität abhängig von der Sättigung durch die Angabe von Leitfähigkeitswerten für den trockenen und gesättigten Zustand ermittelt. Die mathematischen Modelle basieren auf den Konduktivitätswerten und volumetrischen Anteilen der einzelnen Komponenten. Dies ist jedoch nicht die einzige Klassifikationsmöglichkeit, sodass eine Einteilung und Übersicht der Vielzahl an Methoden zur Berechnung der Wärmeleitfähigkeit schwierig ist. Weit verbreitet ist das sogenannte „de Vries-Modell“ (de Vries, 1963) oder das Johansen-Modell (Johansen, 1975). Die darin verwendeten Parameter wurden jedoch in vielen Studien an die jeweiligen Bedingungen in den Untersuchungsgebieten angepasst, sodass von diesen Approximationen mehrere Abwandlungen existieren. Diese und weitere unterschiedliche Methoden zur Berechnung der thermischen Konduktivität werden in Kapitel 5.1.3 genauer erläutert.

Die Qualität und Genauigkeit der Parametrisierung aller Bodenparametern hängt nicht nur von den eingehenden Parametern selbst ab, die meist erst auf die im Untersuchungsgebiet vorhandenen bodentypischen Bedingungen kalibriert werden, sondern auch von vielen anderen Faktoren. An et al. (2016) fanden beispielsweise heraus, dass verschiedene Berechnungsmethoden die Wärmekonduktivität bei vier unterschiedlichen Wetterbedingungen unterschiedlich gut abschätzen konnten. Eine unter allen Wetterbedingungen gut abschneidende Methode wurde nicht gefunden. Weitere Einflussfaktoren auf die Bodenwärmeflüsse und auch auf die Bodentemperatur wie z. B. die Art der Bodenbedeckung oder die Bodenfarbe, die über die Albedo die Energieflüsse ebenfalls stark verändern können, werden in manchen Modellen ebenfalls berücksichtigt (Bonan et al., 2002; Onwuka, 2018). Welche mathematischen Methoden es gibt, die Gleichung 2.14 zu lösen, wird in Kapitel 5.2 aufgezeigt.

## 3 Das regionale Klimamodell REMO

Um die Änderungen, die im Rahmen des REMOLAND Projekts am regionalen Klimamodell REMO durchgeführt wurden, besser verstehen zu können, werden in diesem Kapitel allgemeine Informationen zum Modell, dessen Aufbau und Entwicklung, die einzelnen Schritte zum Durchführen eines Modelllaufs und die in REMO2015 standardmäßig verwendeten Parametrisierungen näher erklärt. Außerdem werden die für REMO benötigten Datensätze und deren Verwendungszwecke vorgestellt.

### 3.1 Entwicklung und Aufbau von REMO

Das regionale Klimamodell REMO (Jacob & Podzun, 1997; Jacob, 2001) ist eine im Rahmen des internationalen Baltic Sea Experiment (BALTEX) (Raschke et al., 1998) durchgeführte Weiterentwicklung des zur numerischen Wettervorhersage verwendeten Europa-Modell (EM) des Deutschen Wetterdiensts (DWD) (Majewski, 1991). Seit 1993 entwickeln der DWD, das Helmholtz-Zentrum Geesthacht-Zentrum für Material- und Küstenforschung (HZG, bis 2010 GKSS), das Deutsches Klimarechenzentrum (DKRZ) und das Max-Planck-Institut für Meteorologie (MPI-M) REMO als Wettervorhersage- und als Klimamodell. Deshalb gibt es für REMO zwei verschiedene physikalische Parametrisierungen je nachdem wie das Modell betrieben wird. Zum einen gibt es die ursprüngliche originale EM-Parametrisierung vom DWD und zum anderen die Parametrisierung vom GCM ECHAM4 (Roeckner et al., 1996). Mittels der Parametrisierungen können subskalige Prozesse, die auch in Regionalmodellen aufgrund ihrer Kleinräumigkeit nicht direkt als physikalische Gleichung mit einfließen können, trotzdem berücksichtigt werden.

In mehreren Studien wurden viele Neuerungen an dieser ECHAM4-Parametrisierung zur Verbesserung von REMO untersucht und teilweise in die Standardversion übernommen, wie zum Beispiel:

- neue Landoberflächenklassifikation (Hagemann, 2002)
- neue Schneeparametrisierung (Semmler, 2002)
- neue Landfraktionsklasse: Eis (Semmler et al., 2004)
- Veränderungen bei der Wolkenparametrisierung (Pfeifer, 2006)
- Veränderungen in der Gletschermaske (Kotlarski, 2007)
- Jahresgang für Vegetationsindizes (Rechid, 2009)
- Änderung der Bodenparameter für Afrika (Eggert, 2011)
- Einbau von Aerosolprozessen (REMO-HAM) (Pietikäinen et al., 2012)
- Albedoänderungen durch ein neues Modell für Landnutzungsänderungen (Preuschmann, 2012)
- Änderung der Wurzeltiefe und des kritischen Wassergehalts in Südamerika (Remedio, 2013)
- neue Parametrisierung der thermischen Bodenparameter und Erweiterung um eine 6. Bodenschicht für Permafrostmodellierung (Blome, 2014)

- neue thermale Bodenparameter für Asien (Kumar et al., 2014)
- Einbau von Teilen der dynamischen Vegetation von JSBACH (REMO-iMOVE) (Wilhelm et al., 2014)
- Corine statt GLCCD als Landnutzungsklassendatensatz (Y. Gao et al., 2015)
- Einbau einer neuen Parametrisierung für Seen (Flake) (Pietikäinen et al., 2017)

REMO wurde bisher als hydrostatisches Modell mit horizontalen Auflösungen von 10–50 km betrieben (Jacob et al., 2014). Bei dieser hydrostatischen Version wird davon ausgegangen, dass sich die Atmosphäre in einer Schichtung befindet, bei der sich die Schwerkraft und der geostrophische Auftrieb, der aus Dichteunterschieden in der Luft erzeugt wird, im Gleichgewicht befinden und gegenseitig auflösen, sodass keine Vertikalbewegung der Teilchen modelliert wird. Diese Annahme kann bis zu einer Auflösung von ca. 10 km verwendet werden. In höheren Auflösungen liegt diese spezielle hydrostatische Schichtung nicht mehr vor und die Vertikalprozesse in der Atmosphäre spielen eine so wichtige Rolle, dass nicht-hydrostatisch gerechnet werden sollte (Dobler et al., 2017). Eine nicht-hydrostatische Version von REMO (REMO-NH) befindet sich in der Entwicklung (Göttel, 2009; Remke & Jacob, 2017).

In vielen Studien wurde das Klimamodell REMO bereits im Vergleich zu anderen regionalen Modellen, zur Analyse von verschiedenen regionalen Klimaänderungen oder zur Analyse der Auswirkungen z. B. von Landnutzungsänderungen auf unterschiedlichen räumlichen Skalen und für unterschiedliche Zeiträume benutzt. Als regionales Klimamodell wird REMO auf kontinentaler Skala verwendet, wie beispielsweise für Südamerika (Sturm et al., 2007; Remedio, 2013), für Europa (Semmler & Jacob, 2004; Jacob et al., 2007), für Asien (Kumar et al., 2015), für Afrika (Hänsler et al., 2016), für Zentralafrika (Fotso-Nguemo, Vondou, Pokam et al., 2017; Fotso-Nguemo, Vondou, Tchawoua et al., 2017; Vondou & Haensler, 2017), für Südwestafrika (Hänsler, 2011) oder für Westafrika (Paeth, 2011). Aber auch auf Länderebene, wie z. B. in einer Studie zur Niederschlagsmodellierung in Indonesien (Aldrian et al., 2005) oder China (Cui et al., 2008) sowie für Änderungen durch Aufforstung oder Entwaldung in Ungarn (Gálos, 2010), wurde mit REMO gerechnet. Auf kleinen Raumskalen unter anderem für eine Analyse der Auswirkung des Waldumbaus in der Lüneburger Heide (Petersen & Rechid, 2014) oder für eine Analyse der Grundwasserneubildung in Niedersachsen (Herrmann et al., 2017) wurde REMO ebenfalls schon angewendet. Neben den typischen Klimavariablen Temperatur und Niederschlag werden auch Windregime (Kunz et al., 2010; Larsén et al., 2010) oder verschiedene Komponenten des Wasserkreislaufs (Karstens et al., 1996, Rechid et al., 2003) mit REMO analysiert. Weitere Studien und Veröffentlichungen in Bezug auf REMO sind auf der Modell-Webseite von GERICS (<https://www.remo-rcm.de/059965/index.php.en>) zu finden. Außerdem ist REMO ein Teil des Coordinated Regional Climate Downscaling Experiment (CORDEX) des World Climate Research Program (WCRP) und wird als Teil von Modell-Ensemble-Läufen eingesetzt. Das regionale Modell wird weltweit eingesetzt, um das jeweilige regionale Klima zu untersuchen und mit verschiedenen Szenarien zu prognostizieren (vgl. Jacob et al., 2012; Teichmann et al., 2013).

## 3.2 Technische Einführung und Modellaufbau

In den nachfolgenden Unterkapiteln werden die grundlegenden Eigenschaften von REMO und die verwendeten physikalischen Grundlagen des Modells kurz erläutert.

### 3.2.1 Grundgerüst des Modells

**Rotation:** Aufgrund der besseren numerischen Verarbeitung wird bei der Berechnung der klimatischen Variablen in REMO ein rotiertes Kugelkoordinatensystem verwendet. Die geographischen Longitude- und Latitude-Koordinaten des ursprünglichen Untersuchungsgebiets werden so rotiert (vgl. Formeln im Anhang A), dass der „neue Äquator“ durch die Mitte des Modellgebiets verläuft. Dadurch werden die Modellpixel annähernd quadratisch und für die Modellgleichungen lassen sich numerische Lösungen einfacher finden (Sieck, 2013). Da das Untersuchungsgebiet im rotierten System als Rechteck mittels der gewünschten Auflösung und der Anzahl an Gitterpunkten gewählt wird, ähnelt der zurück auf reguläre Koordinaten rotierte Modellausschnitt der Form eines Kreissektors, was als kurvilineare Projektion bezeichnet wird. Abbildung 3 verdeutlicht dies anhand des in dieser Arbeit verwendeten Untersuchungsgebiets EUR-44. Der Nordpol wird von  $0^\circ\text{E}$  und  $90^\circ\text{N}$  auf  $-162^\circ\text{E}$  und  $39.25^\circ\text{N}$  rotiert. Dadurch wird Europa an den Äquator rotiert (vgl. rechte obere Teilabbildung in Abbildung 3). Würde man das EUR-44-Gebiet in diesen rotierten Koordinaten auf einer Karte mit regulären Koordinaten einzeichnen, würde das Gebiet in der rechten unteren Teilabbildung von Abbildung 3 angezeigt werden. Durch die Entrotation (vgl. Anhang A) der vom Untersuchungsgebiet abgedeckten Koordinaten erhält man die Fläche, die in der linken unteren Teilabbildung dargestellt ist.

**Modellgitter:** Als dreidimensionales atmosphärisches Zirkulationsmodell wird das jeweilige Untersuchungsgebiet in REMO sowohl in vertikale Schichten als auch in horizontale Pixel unterteilt. In vertikaler Richtung werden die einzelnen Schichten durch ein sogenanntes „hybrid vertikales Koordinatensystem“ (Simmons & Burridge, 1981) anhand des Bodendrucks berechnet, sodass die Schichtgrenzen am Boden der Orographie folgen ( $\sigma$ -System), während sie in den oberen Atmosphärenschichten weitestgehend horizontal verlaufen (p-System). Für die horizontale Aufteilung nutzt REMO ein sphärisches Arakawa-C-Gitter. Bis auf die Windvariablen werden alle Variablen im Zentrum eines Pixels definiert. Die u- und die v-Windkomponente hingegen sind an den Seiten des jeweiligen Pixels definiert (vgl. Abb. 4), da so die geostrophische Anpassung bis auf wenige Ausnahmen am besten nachgebildet werden kann (Arakawa & Lamb, 1977). Dieses spezielle Arakawa-Gitter besitzt zusätzlich gegenüber den anderen Gitter-Versionen weitere Vorteile und wird deshalb in vielen Zirkulationsmodellen eingesetzt (vgl. Mesinger & Arakawa, 1976).

**Räumliche Modellauflösung und hydrostatische Approximation:** In der Standardversion verwendet REMO die hydrostatische Approximation, bei der Druckgradientkraft und Schwerkraft

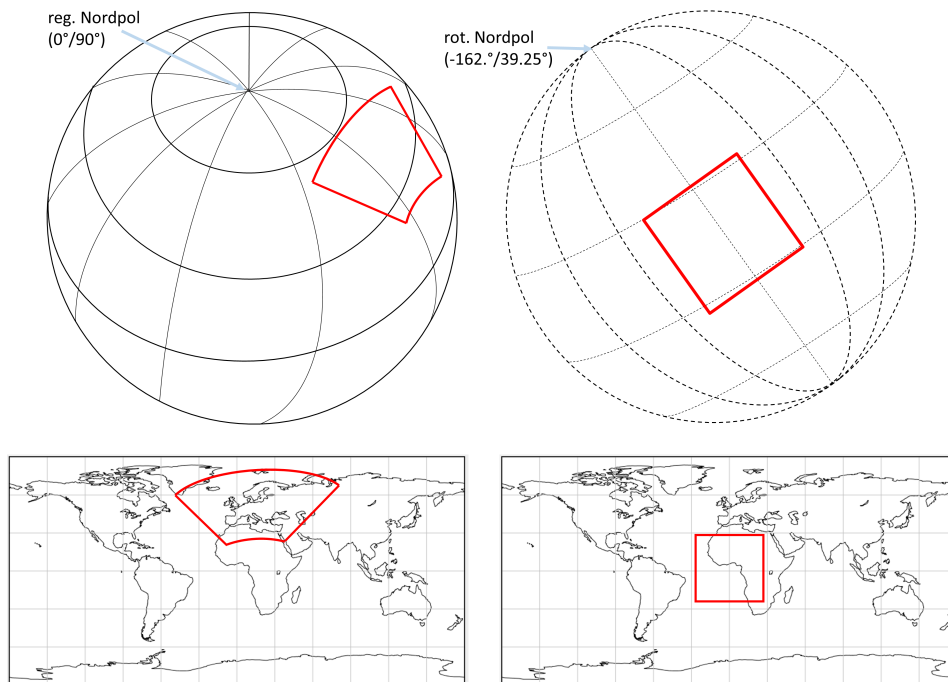


Abb. 3: Darstellung eines regulären und rotierten Modelluntersuchungsgebiets anhand des EUR-44-Testgebiets  
 links: reguläre Koordinaten  
 rechts: rotierte Koordinaten  
 oben: Darstellung als Globus  
 unten: Darstellung als Karte

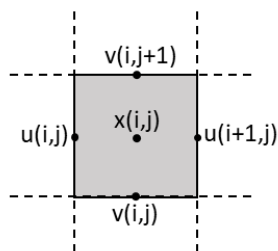


Abb. 4: Schematische Darstellung des Arakawa-C-Gitters, bei dem alle Werte eines Pixels im Mittelpunkt des Pixels definiert sind, die u- und v-Komponenten des Windes jedoch an den Seiten.

gleichgesetzt werden, sodass keine Vertikalbeschleunigung stattfindet. Dies ist bei den bisherigen Modellauflösungen, in denen das Modell benutzt wurde ( $0.11^\circ$ ,  $0.22^\circ$  oder  $0.44^\circ$ ), gängige Praxis, weil auf diesen Modellskalen der Einfluss der Vertikalbewegung im Vergleich zu dem der Horizontalbewegungen klein ist. Bei einer kleineren Maschenweite des Gitters muss, wie bereits erwähnt, ohne hydrostatische Approximation gerechnet werden, was einige Anpassungen im Modell erfordert. Eine entsprechende REMO-Version befindet sich in der Testphase, wird aber für diese Arbeit nicht verwendet.

**Downscaling, Nesting und Relaxationszone:** Zum Nesten wird in REMO der Einfluss der Antriebsdaten exponentiell zum Modellgebietsmittelpunkt abnehmend über acht Randgitterboxen (Relaxationszone) nach Davies (1976) simuliert. Zudem handelt es sich bei der in REMO angewendeten Methode um das „One-Way Nesting“, da nur Informationen vom gröber aufgelösten Modell ins regionale Modell fließen, nicht aber zurück. Die Erweiterung hin zu einem zwei-Wege Nesting, bei der auch das antreibende Modell von dem höher aufgelösten Modell Werte zurückbekommt, wurde für REMO bereits getestet (vgl. Lorenz, 2011). Die in dieser Arbeit verwendete Version nutzt weiterhin das standardmäßige One-Way Nesting.

**Antriebsdaten:** Da es sich bei REMO um ein regionales Klimamodell handelt, werden sogenannte Antriebs- bzw. Initialisierungsdaten benötigt, um das Modell mit Informationen von außerhalb des Modellgebiets zu versorgen. Diese Daten, auch Randbedingungen (engl. Lateral Boundary Conditions) (LBC) genannt, stammen entweder von Reanalysedaten, von einem GCM oder von einem in einem größeren Modellgebiet aber mit niedriger Auflösung betriebenen regionalen Klimamodell. Bei der Verwendung von Reanalysedaten spricht man von „Perfect Boundary Conditions“ (Jacob, 2001), da es sich dabei um aufbereitete Beobachtungsdaten handelt. Diese sind jedoch nur für das Simulieren des Klimas in der Vergangenheit (Hintcasting) einsetzbar. Für einen Reanalyse-Datensatz werden unregelmäßige Beobachtungsdaten mit Modelldaten kombiniert, die viele physikalischen und dynamische Prozesse erfassen, um einen räumlich homogenen und zeitlich kontinuierlichen Datensatz zu erhalten (Q. Sun et al., 2018). Es gibt mehrere Reanalyse-Datensätze, die in unterschiedlichen räumlichen und zeitlichen Auflösungen und für unterschiedliche Zeiträume verfügbar sind, z. B.

- NCEP/NCAR Reanalysis (Reanalysis-1) des National Centers for Environmental Prediction (NCEP) und des National Center for Atmospheric Research (NCAR) (Kalnay et al., 1996),
- NCEP/DOE Reanalysis (Reanalysis-2) des NCEP und des Department of Energy (DOE) (Kanamitsu et al., 2002),
- Climate Forecast System Reanalysis (CFSR) des NCEP (S. Saha et al., 2010),
- Europäische Reanalyse Datensätze ERA-40, ERA-Interim bzw. ERA-5 des European Center for Medium-Range Weather Forecasts (ECMWF) (Uppala et al., 2005; Simmons et al., 2007; Hersbach et al., 2020),
- Modern-Era Retrospective Analysis for Research and Applications (MERRA) Reanalysedaten des Goddard Space Flight Center (GSFC) der National Aeronautics and Space Administration (NASA) (Rienecker et al., 2011) oder
- Japanese 25- bzw. 55-year Reanalysis (JRA-25, JRA-55) der Japan Meteorological Agency (JMA) (Onogi et al., 2007; Ebita et al., 2011).

Die einzelnen Datensätze unterscheiden sich zudem in den jeweils verwendeten physikalischen Parametrisierungen, numerischen Schemata und den Assimilationsschemata. Einen Vergleich und nähere Informationen zu diesen Reanalysedaten liefern X. Zhang et al. (2016). In dieser Arbeit werden die ERA-Interim Reanalysedaten des ECMWF zum Antreiben verwendet, welche in Kapitel 3.4.1 genauer vorgestellt werden. Ferner ist zu beachten, dass je nach Klimamodell unterschiedliche Variablen als Anfangsbedingungen benötigt werden. Die prognostischen Variablen

in REMO sind: Luftdruck auf Bodenniveau, Lufttemperatur, spezifische Feuchte, Wolkenwasser und die beiden horizontalen Windkomponenten für unterschiedliche Atmosphärenschichten. Diese Variablen werden für jeden Pixel im Modellgebiet und für jeden Zeitschritt neu berechnet.

Alle weiteren Variablen, die in diesen sogenannten a-files (Antriebsdaten) enthalten sind, wie beispielsweise die Oberflächentemperatur oder die Feuchte, werden zu Beginn eines Modellaufes an jedem Pixel aus den Antriebsdaten initialisiert und anschließend von REMO ohne weiteres externes Forcing berechnet.

**Zeitliche Diskretisierung:** REMO arbeitet bei der zeitlichen Diskretisierung mit dem sogenannten „leap-frog Schema“, das auch „Bocksprungverfahren“ genannt wird und eine explizite Methode ist. Dabei wird die Prognose nicht vom aktuellen Rechenzeitpunkt  $t$ , sondern vom vorhergehenden Zeitpunkt  $t - \Delta t$  ausgeführt. Diese numerische Integration hilft, genauere Berechnungen durchzuführen, wenngleich die Methode das Speichern der zu prognostizierenden Variablen von drei statt zwei Zeitpunkten nötig macht. Durch das Anwenden des Zeitfilters nach Asselin (1972) wird ein semi-implizites Korrekturverfahren benutzt, um zum einen das Bilden von zwei Zeitfamilien zu vermeiden und zum anderen größere Zeitschrittintervalle verwenden zu können (Majewski, 1991, Göttel, 2009). Je nach horizontaler Auflösung muss der zeitliche Abstand zwischen zwei Rechenschritten dennoch angepasst werden. Beispielsweise werden in der hydrostatischen Version und bei einer Modellauflösung von  $0.44^\circ$  üblicherweise 240 Sekunden und bei einer Modellauflösung von  $0.11^\circ$  60 Sekunden in REMO angenommen.

### 3.2.2 Ablauf eines Modellaufs

Ein regionaler Klimamodellauf besteht aus mehreren Teilen. In Abbildung 5 ist der klassische Workflow von REMO schematisch dargestellt. Zunächst muss ein Untersuchungsgebiet mit der gewünschten Auflösung und der Anzahl der Pixel definiert werden. Anschließend werden im Präprozessor in der sogenannten Bodenbibliothek (BodLibKit) die Landoberflächenparameterkarten erstellt. Diese werden mit den klimatologischen Antriebsdaten (g-files) zu a-files zusammengebracht. Danach findet der eigentliche Modellauf statt. In der Postprozessierung wird dann der Modelloutput entrotiert und für die gewünschten Analysen aufbereitet. In den folgenden Unterkapiteln werden die einzelnen Schritte genauer erläutert.

#### 3.2.2.1 Präprozessor

Um einen REMO-Modellauf zu starten, werden zunächst die Präprozessierungen benötigt. Diese gliedern sich in zwei Schritte:

1. Erstellen der Landoberflächenparameter in der Bodenbibliothek
2. Interpolieren und Rotieren der Antriebsdaten auf das Modellgebiet und Zusammenführen mit den Parametern der Bodenbibliothek

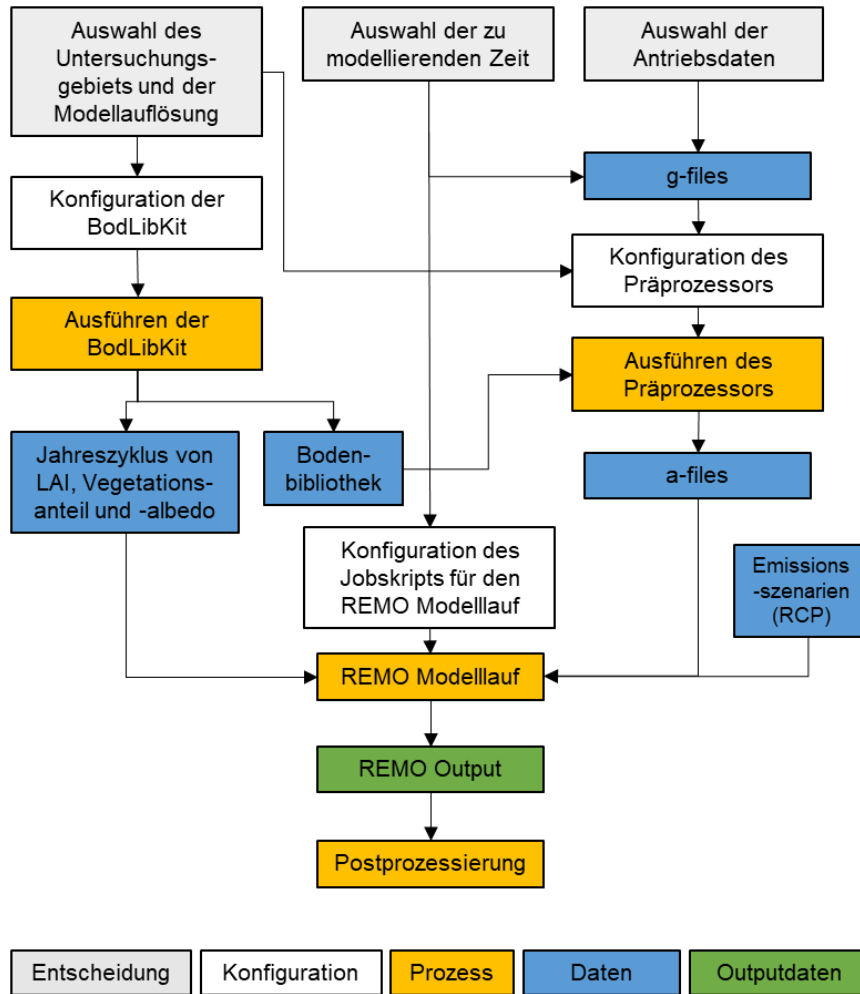


Abb. 5: Klassischer Workflow beim Ausführen eines REMO-Modelllaufs

Zum Erstellen der Bodenbibliothek sind wiederum mehrere Schritte notwendig. Der Prozess gliedert sich nach den verwendeten Eingangsdaten, deren Aufbereitung und Kombination in neun Schritte. Eine Übersicht der Daten der BodLibKit ist in Kapitel 3.4.2 und eine genaue Beschreibung der Skripte ist in Anhang D zu finden. Nach dem Fertigstellen der BodLibKit werden diese konstanten Werte, die dann auf das jeweilige Untersuchungsgebiet und die gewünschte Modellauflösung angepasst sind, zusammen mit den zeitlich variablen Antriebsdaten im Präprozessor in die dort erstellten Dateien geschrieben.

Der Präprozessor wird im Laufe dieser Arbeit so abgewandelt, dass die Landoberflächendaten nicht mehr in die a-files geschrieben werden müssen, sondern direkt vom Modelllauf aus der in der Bodenbibliothek erstellten Datei eingelesen werden (vgl. Abb. 6). Dadurch kann sowohl die Rechenleistung beim Erzeugen der a-files als auch die Speicherauslastung durch die a-files minimiert werden. Des Weiteren wird der Austausch der Landoberflächenparameter bei sonst gleichbleibenden Rahmenbedingungen erleichtert. Somit können neue konstante Variablen direkt an das Modell übergeben werden, da die a-files nach dieser Anpassung nicht mehr verändert werden müssen.



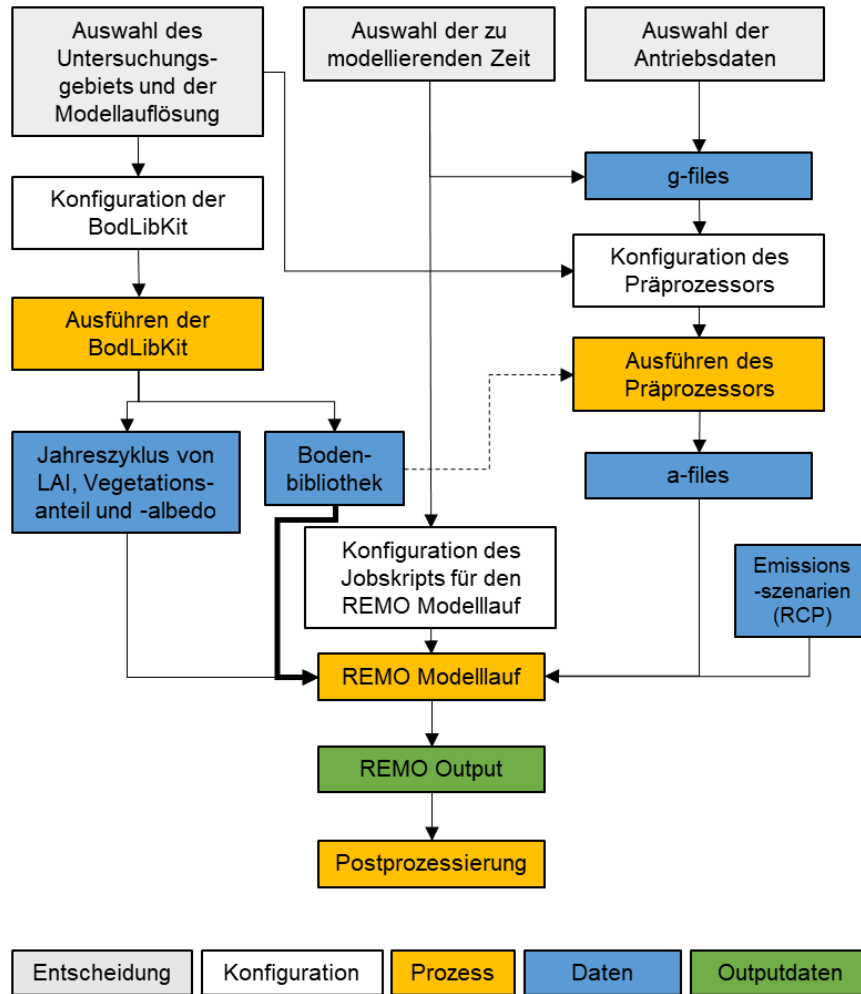


Abb. 6: Neuer Workflow beim Ausführen eines REMO-Modelllaufs

### 3.2.2.2 Modelllauf

Nach dem Vorbereiten der a-files muss das sogenannte Jobskript konfiguriert werden. Mit diesem wird beispielsweise die Namelist angelegt, in der festgelegt wird für welchen Zeitraum, mit welchen Einstellungen für verschiedene Parametrisierungen und mit welchem Zeitschritt REMO rechnen soll. Gleichzeitig wird dort definiert, welche Variablen mit welcher Häufigkeit herausgeschrieben werden. Dann kann das Modell mit dem Ausführen des Jobskripts gestartet werden. Das Modell ist in der Programmiersprache FORTRAN geschrieben und setzt sich aus einem Hauptprogramm und vielen Unterprogrammen zusammen. In Anhang E ist beispielsweise ein Teil des komplexen Zusammenhangs zwischen den Programmen als sogenannter Calltree dargestellt. Neben dem Einlesen und Herausschreiben (engl. Input/Output) der Daten (IO) gliedert sich REMO in zwei Hauptteile auf: die Dynamik und die Physik. Während der dynamische Teil bei beiden anfangs erwähnten REMO Physik-Parametrisierungen gleich ist, unterscheiden sich die DWD-Physik und die ECHAM4-Physik. Ein Vergleich der beiden Parametrisierungen liefert Clemens (1998). In dieser Arbeit wird, wie bereits erwähnt, mit der ECHAM4-Physik gearbeitet.

Der Calltree in Anhang E wurde durch einen sogenannten Profiling eines REMO-Modelllaufs (Modellzeit 1h) mittels des Kompilierbefehls *gprof* und mit dem frei verfügbaren Pythonprogramm *gprof2dot.py* (Fonseca, 2008-2017), das zur anschließenden Darstellung der Profiling-Ergebnisse genutzt wird, erzeugt. Der vereinfachte Baum stellt nur die Programme dar, die einen Rechenzeitanteil von über 0.5% haben. Die Prozentzahlen in der Abbildung stellen diesen Anteil der Rechenzeit an der Gesamtrechenzeit des Modells für jedes Programm mit seinen jeweiligen Unterprogrammen (=„Kinder“) und für jedes einzelne Programm in Klammern dar. Die Einfärbung der Boxen markiert das Verhältnis der jeweils kumulierten Rechenzeit der Programme.

Das Hauptprogramm *remorg* steuert *control* an, welches wiederum die Initialisierung der Variablen und das Einlesen der Anfangsbedingungen mit *init\_remo*, den eigentlichen Lauf und das Abspeichern der Ergebnisse über *timeloop* und das durch das Abspeichern und Schließen aller erstellten Dateien definierte Modellende mit *finish\_remo* (nicht im Teilbaum dargestellt) aufruft. Anhand der dritten Zahl in den Boxen der Abbildung kann man erkennen, wie oft ein Programm über den jeweils gerechneten Modelllauf aufgerufen wird. Weil der dargestellte Lauf aus der Analyse eines einstündigen Laufes des GER-11-Gebiets entstanden ist, welches mit einem Modellzeitschritt von 60 Sekunden angetrieben wurde, werden die meisten Kinder von *timeloop* 60x aufgerufen. Die längste Zeit wird mit 96.6% in den von *progec4* aufgerufenen Programmen verbracht. Die Kinder von *phyec* stellen zusammen den physikalischen Teil von REMO dar und benötigen 67.7% der Rechenzeit. Das Unterprogramm *progecp*, das als einzelnes Programm mit knapp 8% an der Gesamtrechenzeit relativ lange dauert, gehört jedoch beispielsweise zum dynamischen REMO-Teil. Bei den mit „MOD“ beginnenden Unterprogrammen handelt es sich um Module, die hauptsächlich zur Deklaration von Variablen und Konstanten sowie dem Definieren von einfachen Funktionen, wie beispielsweise dem Herausschreiben der Variablen zu einem bestimmten Zeitschritt, benutzt werden.

Alle in dieser Arbeit thematisierten Parametrisierungsänderungen sind durch Änderungen in dem in Anhang E gelb markierten Programm *surf* und in dem nicht im Teilbaum dargestellten Programm *init*, welches nur zu Beginn eines Laufs unter anderem zur Definition der Bodenparameter aufgerufen wird und damit einen zu geringen Rechenzeitanteil hat, vorgenommen worden. Für das Hinzufügen von neuen Schaltern und Variablen sowie für die Veränderungen der Anzahl der Dimensionen von speziellen Variablen mussten zudem die Variablendeklarationen in vielen Unterprogrammen geändert und ergänzt werden, die jedoch ebenfalls nicht alle im unvollständigen Calltree im Anhang zu sehen sind.

#### 3.2.2.3 Postprozessierung

Alle REMO-Daten liegen in dem in Kapitel 3.2.1 erläuterten rotierten Koordinatensystem vor, sodass vor der Analyse und z. B. bei einem Vergleich mit Beobachtungsdaten die Modelldaten zurück in ein reguläres Koordinatensystem transformiert werden müssen. Der verwendete Algorithmus ist in Anhang A dargestellt. REMO schreibt in der Standard-Version sowohl stündliche Daten für ausgewählte Variablen (xe-Daten) als auch 6-stündige hauptsächlich atmosphärische

Daten (xt-Daten) heraus, die bei Bedarf als Nesting-Daten für den Antrieb von noch höher aufgelösten Modellsimulationen benutzt werden können. Diese Zeitintervalle können jedoch im Jobskript angepasst werden. Eine Übersicht, welche Variable in welcher Datei und in welcher zeitlichen Auflösung abgespeichert ist, liefert die Tabelle in Anhang B. Darin ist außerdem festgehalten, wie REMO die stündlichen Werte berechnet. Es sind fünf verschiedene Arten möglich. Für die meisten Variablen wird der Mittelwert über diese eine Stunde gebildet. Als zweite und dritte Möglichkeit der Aggregation werden bei ausgewählten Variablen, wie z. B. der Minimum- und Maximumtemperatur, der minimale bzw. maximale Wert, der in der Stunde erreicht wird, abgespeichert. Die vierte Art, die vor allem bei hydrologischen Variablen verwendet wird (z. B. Menge aller Niederschlagsarten), gibt die Summe über den Zeitraum aus. Bei der letzten Möglichkeit wird der letzte in dieser Stunde berechnete Wert in die Ausgabedatei geschrieben. Diese Unterscheidung ist für die Analyse der Daten und das zeitliche Aggregieren auf Tages-, Monats- oder Saisonbasis wichtig.

#### 3.2.2.4 Kalt- und Warmstart

Wird das Klimamodell zum ersten Mal mit den originalen Anfangswerten gestartet, nennt man diesen Vorgang Kaltstart. Dieser Lauf sollte so lange dauern, bis ein Gleichgewicht zwischen den Antriebsdaten und der Modellphysik (Equilibrium) erreicht ist. Der sogenannte Warmstart hingegen ist ein Neustart des Modells mit den im Gleichgewicht befindlichen Werten eines Kaltstarts. Die Zeit, die ein Kaltstart benötigt, bis er das Equilibrium erreicht, wird Spin-Up Zeit (Einschwingzeit) genannt. Es gibt zwei Möglichkeiten an Warmstarts:

- kontinuierlicher Modelllauf: Die Daten der Spin-Up Zeit werden nicht verwendet. Folglich muss der Startzeitpunkt des Modelllaufs ausreichend vor dem gewünschten Analysezeitraum liegen.
- Neustart des Modelllaufs: Wird das Equilibrium erreicht, werden die Werte in den Variablen zu diesem Zeitpunkt benutzt, um das Modell erneut zum Anfangszeitpunkt zu reinitialisieren. (Diese Methode wird in dieser Arbeit verwendet.)

Welche Art für welche Modelle am besten geeignet ist, wie lange der Kaltstart dauert und welche Variablen eingeschwungen werden müssen, ist Gegenstand von zahlreichen Forschungen. Neben den sehr langsam reagierenden Ozeanmodellen, die mehr als 50 Jahre Spin-Up Zeit benötigen, sind es vor allem Bodentemperatur und -wassergehalt, die am längsten eingeschwungen werden müssen - auch in atmosphärischen Modellen (vgl. Giorgi & Mearns, 1999; J. Wang & Kotamarthi, 2014). Wird ein Modell zu einem ungünstigen Zeitpunkt initialisiert, verlängert sich die Einschwingzeit signifikant (Ridder, 2009). Dagegen können günstige Bedingungen oder gewisse Eigenschaften eines Modells die Zeit extrem verkürzen (z. B. verwenden Quante & Colijn (2016) nur 1-3 Jahre Spin-Up für ein regionales Ozeanmodell). Zudem zeigen Bonekamp et al. (2018) anhand des Weather Research and Forecasting (WRF) Modells (Skamarock & Klemp, 2008), dass die Wahl der Spin-Up Zeit, wegen des großen Einflusses auf den Bodenwasserhaushalt und damit indirekt auf die Wärmeflüsse, Auswirkungen beispielsweise auf den Modellniederschlag hat. Zu beachten

ist, dass sich diese Studie auf ein Wettermodell bezieht und die typische Spin-Up Zeit für diese Art von Modellen bei wenigen Stunden liegt. Für regionale Klimamodelle werden allerdings meist einige Jahre zum Einschwingen des Bodenwassergehalts hergenommen (Giorgi, 2019). Montávez et al. (2017) hingegen beziffern die optimale Einschwingzeit für RCMs für Bodentemperatur und -wassergehalt auf nur zwei bis sechs Monate.

Wegen der Temperatur in der tiefsten Modellbodenschicht ( $-10$  m) wird in der REMO Modell-dokumentation eine Spin-Up Zeit von 5-10 Jahren angegeben (Buntemeyer et al., 2015). In der Literatur findet man speziell für REMO unterschiedliche Angaben von einem Jahr (Rechid, 2009), über sechs (Davin et al., 2019) und zehn (Rechid, 2001) Jahre bis hin zu 20 (Kumar et al., 2014) und 35 Jahren (Pietikäinen et al., 2017). Für den weiteren Verlauf der Arbeit wird daher eine eigene Analyse des Einschwingvorgangs für das Untersuchungsgebiet durchgeführt. Da die Zeit von verschiedenen Faktoren abhängt und vor allem, weil das Modell mit veränderten Bedingungen und Parametrisierungen (speziell in den Bodenschichten) sehr häufig ausgeführt wird, sollte die Einschwingzeit so klein wie möglich gehalten werden, um unnötigen Rechen- und Speicheraufwand zu vermeiden.

Für REMO werden standardmäßig folgende 15 Variablen eingeschwungen. (Die Zahlen in den Klammern entsprechen den Variablen IDs von REMO, vgl. Anhang B)

- Oberflächentemperaturen für Land-, Wasser- und Eisflächen (54, 55, 56)
- spezifische Feuchte an der Landoberfläche (84)
- Bodenwassergehalt und Wassergehalt auf Blatt- und Bodenoberfläche (140, 194)
- Schneetiefe und -temperatur (141, 206)
- Bodentemperaturen (207, 208, 209, 170, 183)
- Meereis- und Gletschermaske (210, 232)

Bildet man das jeweilige Flächenmittel des Kalt- und Warmstarts für das Untersuchungsgebiets und stellt beide als Zeitreihe dar, erkennt man bei den Variablen, die in Abbildung 7 exemplarisch dargestellt sind, nur im ersten Jahr bzw. bei der Schneetiefe (Abb. 7B) in den ersten beiden Jahren größere Differenzen zwischen beiden Läufen. Der Literatur folgend, weisen nur die Variablen der tiefen Bodenschichten klar ersichtliche und sich über mehrere Jahre hinweg erstreckende Einschwingvorgänge auf. Dabei benötigt die Bodentemperatur umso länger zum Einschwingen, je tiefer sie liegt (vgl. Abb. 8).

Durch das systematische Testen von verschieden langen Spin-Up Zeiten mit Start im Januar wird untersucht, was die minimale Einschwingzeit für das gewählte Untersuchungsgebiet ist. Es ist bekannt, dass der Startzeitpunkt des Kaltstartes die Spin-Up Zeit erheblich beeinflusst. Kalte und trockene Anfangsbedingungen brauchen besonders lange, um auf ein neues Niveau zu kommen. Die Qualität des jeweiligen Spin-Ups wird im Vergleich zu einem über 15 Jahre eingeschwungenen Boden beurteilt. Bei einem Vergleich der Ergebnisse des Kalt- und Warmstartes lässt sich feststellen, dass sowohl die Oberflächentemperatur als auch Schneetemperatur sowie die Temperaturen der obersten zwei Bodenschichten innerhalb eines Jahres fast identisch sind. Die Variablen Meereis- und Gletschermaske, die im Untersuchungsgebiet nur marginal

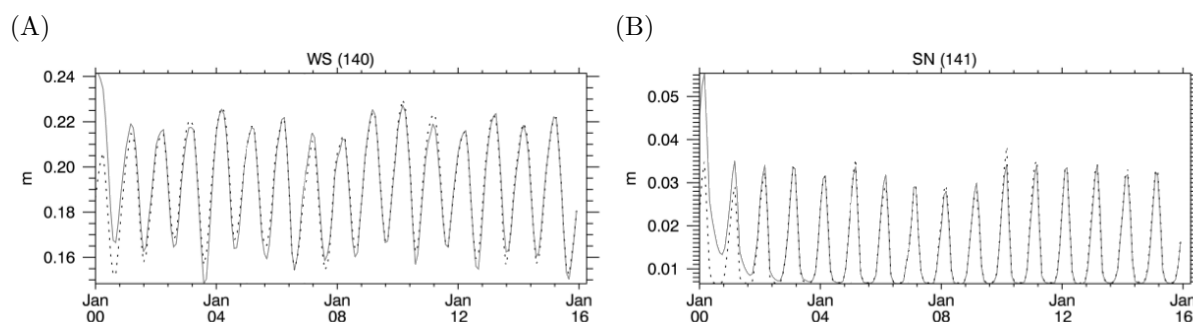


Abb. 7: Kalt- und Warmstart im Vergleich - Flächenmittel des EUR-44 Testgebiets für zwei unterschiedliche Variablen

Kaltstart: graue durchgezogene Linie

Warmstart: schwarz gepunktete Linie

(A) Bodenwassergehalt (WS)

(B) Schneetiefe (SN)

auftreten, benötigen nur wenige Zeitschritte. Daher kann die optimale Einschwingzeit anhand der Bodentemperaturen festgelegt werden. In Abbildung 8 sind sowohl der Kaltstart als auch ein 15 Jahre langer Warmstart (15a) aufgetragen. Zusätzlich sind für die Jahre 2005 und 2006 die Differenz zwischen Kalt- und 15-jährigen Warmstart sowie die Differenzen zwischen den drei Testwarmstarts und dem 15-jährigen Warmstart dargestellt. Während in der vierten Bodenschicht die Temperatur aller Testwarmstarts (Abb. 8A) schon 2005 kaum von der des langen Warmstarts abweicht und kein signifikanter Erwärmungstrend über den Zeitraum zu erkennen ist, weisen die Differenzkurven der untersten Bodenschicht (Abb. 8B) mit  $0.4\text{ °C}$  beim 6-jährigen Warmstart und  $0.1\text{ °C}$  sowohl beim 8- als auch beim 10-jährigen Warmstart den charakteristischen Erwärmungstrend auf. Es lässt sich festhalten, dass der Unterschied zwischen acht und zehn Jahren sehr klein ist. Zudem sind die Differenzen beider Läufe im Vergleich zum 15-jährigen Warmstart nicht mit dem Aufwand und der Rechenzeit für weitere sieben bzw. fünf Modelljahre rechtfertigbar. Daher wären für das gewählte Modellgebiet schon acht Jahre Spin-Up ausreichend gewesen.

Alle anderen Standardausgabeveriablen (vgl. Anhang B) wurden ebenfalls auf die für den Einschwingvorgang charakteristischen Merkmale für das EUR-44 Flächenmittel geprüft. Es stellte sich heraus, dass durch die Auswahl der 15 Variablen auch andere potentiell für einen Spin-Up in Frage kommende Variablen, wie beispielsweise der Eisanteil pro Bodenschicht, bereits ausreichend angeglichen werden. Daher ist das Ersetzen weiterer Variablen für diese Arbeit nicht notwendig. Wird in Zukunft das 5Layer-Schema benutzt, sollten die fünf Wassergehaltsvariablen aber in der Liste der Spin-Up Variablen ergänzt werden.

In dieser Arbeit werden nachfolgend die Bodenvariablen des längsten getesteten Laufs (15a) für die Warmstarts verwendet, da dieser Lauf für die Analyse der Spin-Up Zeit bereits vorliegt und nicht neu gerechnet werden muss.

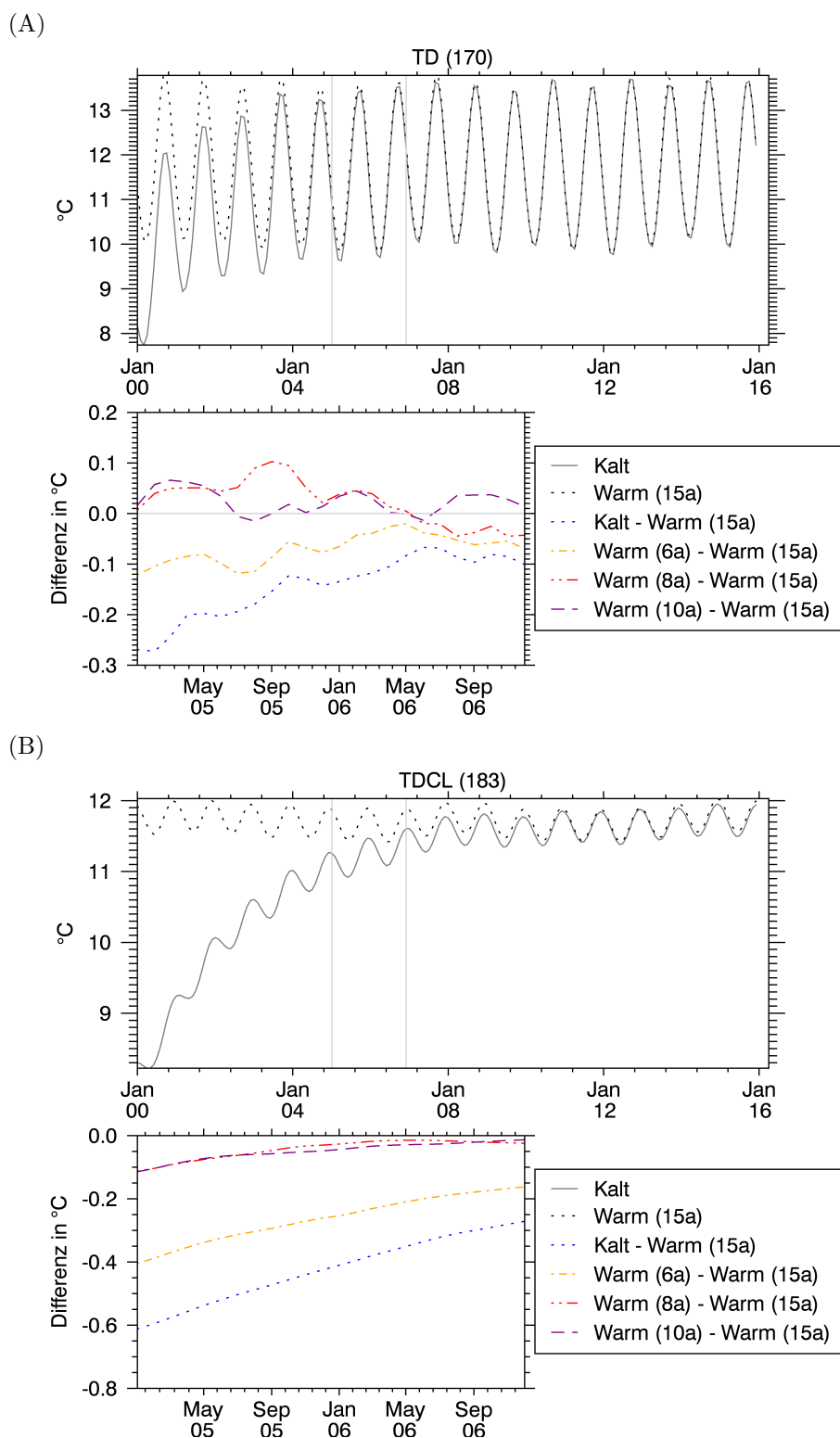


Abb. 8: Vergleich des Flächenmittels des EUR-44 Testgebiets für unterschiedliche lange Spin-Up Zeiten

Diagramm oben: Kalt- und 15-jähriger Warmstart für 2000-2015 (Die zwei grauen horizontalen Linien kennzeichnen den Zeitausschnitt für das zweite Diagramm.)

Diagramm unten: jeweilige Differenz zum 15-jährigen Warmstart für 2005-2006

(A) Bodentemperatur der vierten Schicht (TD)

(B) Bodentemperatur der fünften Schicht (TDCL)

### 3.3 Beschreibung verschiedener Parametrisierungen für Prozesse in REMO

In REMO wird die Interaktion zwischen der Erdoberfläche, dem System „Boden“ und der Atmosphäre in unterschiedlichen Schritten behandelt. Zunächst wird im Präprozessor eine Definition der Erdoberfläche mittels unterschiedlicher Datensätze generiert. Dazu wird die Oberfläche des Modellgebiets (anders als in der originalen ECHAM4-Parametrisierung) zunächst anteilmäßig anhand der korrespondierenden Fläche in einem Modellpixel in eine Land-, Meer- und Eisklasse eingeteilt (Semmler, 2002) (vgl. Abb. 9). Dieser Ansatz („tile approach“) hat den Vorteil, dass ein kontinuierlicher Übergang zwischen z. B. räumlich benachbarten Meeres- und Landpixeln möglich ist. Außerdem werden für jeden Bedeckungsanteil eine eigene Rauigkeitslänge, Albedo und Oberflächentemperatur sowie verschiedene Energieflüsse berechnet (vgl. Variablenliste in Anhang B). Des Weiteren können sich im Gegensatz zu einer starren Land-Meer-Maske die Anteile in REMO über einen Modelllauf hinweg ändern. Seit 2007 gibt es eine neuere Modellversion, bei der die bisherige Eisklasse (Seeeis) von Kotlarski (2007) um das Gletschereis auf dem Land erweitert wurde. Diese Einstellung ist aber standardmäßig deaktiviert und wird in dieser Arbeit nicht verwendet.

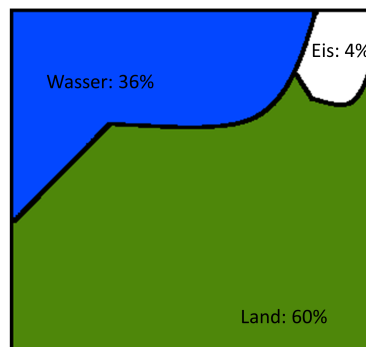


Abb. 9: Schematisches Beispiel einer Landoberflächeneinteilung eines Pixels in REMO in die drei Oberflächenklassen Land, Wasser und Eis

Die turbulenten Flüsse und der Oberflächenstrahlungsfluss werden pro Oberflächenklasse und für jeden Zeitschritt berechnet und anschließend anteilmäßig gewichtet gemittelt. Für die Vertikaldiffusion und die turbulenten Oberflächenflüsse wird in REMO die Monin-Obukhov-Ähnlichkeitstheorie (Louis, 1979) verwendet. Die Eddy-Diffusionskoeffizienten werden abhängig von der turbulenten kinetischen Energie berechnet. Anhand des Ansatzes von Morcrette et al. (1986) mit zusätzlichen Treibhausgasen und Aerosolarten werden die Strahlungsprozesse abgebildet. Der Wolkenwassergehalt wird nach Sundqvist (1978) aus einer Bilanzgleichung abgeleitet. Die Kumuluskonvektion wird nach dem Massenflussschema von Tiedtke (1989) mit Modifikationen von Nordeng (1994) realisiert. Für die Parametrisierung der nicht-konvektiven Wolken und die Abbildung der Wolkenmikrophysik wird das Schema von Lohmann & Roeckner (1996) verwendet (vgl. Roeckner et al., 1996; Kotlarski, 2007; Remedio, 2013; Bunttemeyer et al., 2015).

Der Boden wird in REMO in fünf Temperaturschichten eingeteilt (vgl. Abb. 10), die vom ECHAM Modell übernommen wurden (vgl. Deutsches Klimarechenzentrum, 1993). Die Schichtdicke nimmt mit zunehmender Tiefe zu und reicht bis in ca. 10 m Tiefe. Dabei wird der Wärmetransfer mittels einer Diffusionsgleichung bestimmt und für die unterste Bodenschicht auf Null gesetzt, was für Klimamodelle üblich ist, da die unterste Schicht sonst zu einer unendlichen Senke von Wärme werden würde (Warrilow et al., 1986).

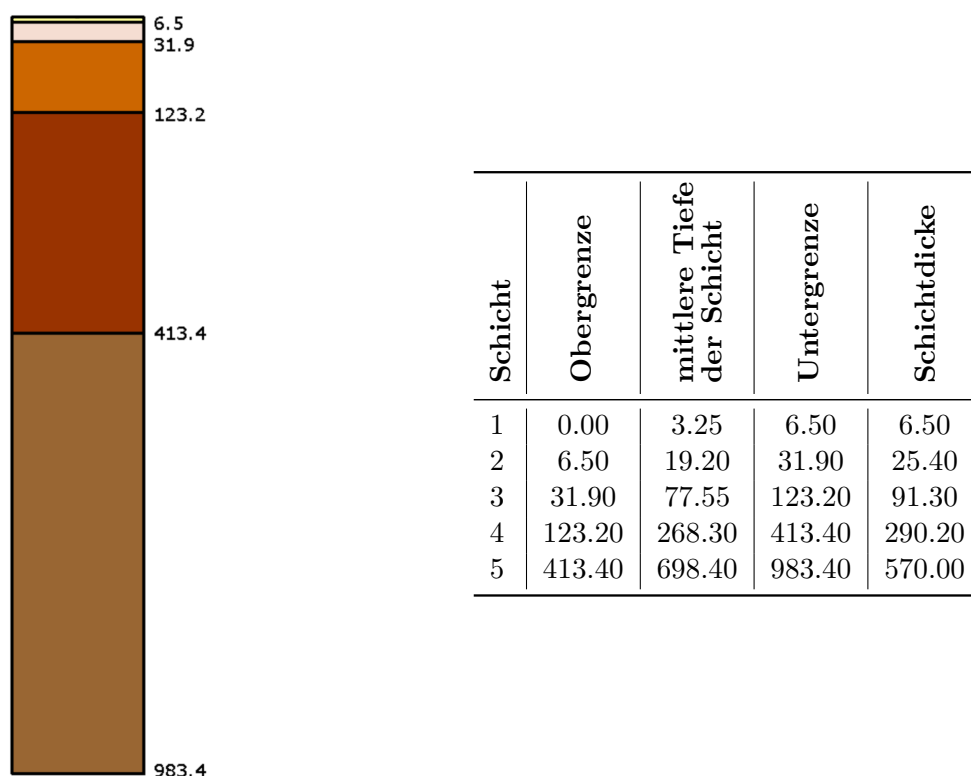


Abb. 10: Bodenschichten in REMO

links: proportionale Abbildung der Schichten mit der jeweiligen Schichtuntergrenze in cm

rechts: Tabelle mit den verschiedenen Tiefenangaben in cm

Mit dieser unteren Randbedingung und durch das Formulieren der Diffusionsgleichung für jeden Modellzeitschritt für jede der fünf Bodenschichten ergibt sich ein Gleichungssystem, welches in REMO bisher folgendermaßen gelöst wird (vgl. Formeln 3.1). Zunächst wird die rechte Seite der Gleichung als  $\frac{2}{3}$  der Bodentemperatur des vorherigen Zeitschritts für die jeweilige Schicht definiert. Anschließend werden für die oberen vier Schichten die Gleichungen für die Wärmeflüsse definiert und dann sukzessive von oben nach unten eingesetzt. Wegen der unteren Randbedingung kann die Gleichung der untersten Schicht gelöst werden und die restlichen Gleichungen von unten nach oben zurück substituiert werden. Zudem ist zu beachten, dass nur für die oberste Schicht der Bodenwärmestrom ( $TEM1$ ), der durch die Umformulierung von Gleichung 2.3 aus den anderen Komponenten berechnet wird, zusätzlich in die Gleichung einfließt (vgl. Formel 3.1e). Zum Schluss wird auf diesen veränderten Wert noch das übrige  $\frac{1}{3}$  der vorherigen Bodentemperatur pro Schicht addiert, um die neue Temperatur zu erhalten. Durch dieses Vorgehen wird die Änderung der



Bodentemperatur von einem auf den nächsten Zeitschritt im Modell numerisch eingeschränkt.

Initialisierung:

$$ZSDIF_k = \frac{2}{3} \cdot T_{k,t-1}, \text{ für } k \in [1, 5] \quad (3.1a)$$

Aufstellen der Diffusivitätsgleichung für die verschiedenen Bodenschichten:

$$ZCFH_k = \frac{1.5 \cdot 2 \cdot \Delta t}{\Delta z_k + \Delta z_{k+1}} \cdot \frac{ZDIFI_{k+1} \cdot \Delta z_{k+1} + ZDIFI_k \cdot \Delta z_k}{\Delta z_{k+1} + \Delta z_k}, \text{ für } k \in [1, 4] \quad (3.1b)$$

Einsetzen für die oberste Bodenschicht:

$$ZSA = \frac{\Delta t}{ZCGN_1 \cdot \Delta z_1} \quad (3.1c)$$

$$ZSB = \frac{\Delta t \cdot \frac{1}{0.5 \cdot (\Delta z_1 + \Delta z_2)} \cdot ZDIFI_1}{\Delta z_1} \quad (3.1d)$$

$$ZSDIF_1 = \frac{1}{1 + \frac{ZCFH_1}{\Delta z_1}} \cdot \left( ZSDIF_1 + TEM1 \cdot \frac{ZSA}{1 - 0.5 \cdot \left( ZSA \cdot \min \left( \frac{ZSB}{ZSA}, TEM2 \right) - ZSB \right)} \right) \quad (3.1e)$$

$$ZEB_1 = \frac{1}{1 + \frac{ZCFH_1}{\Delta z_1}} \quad (3.1f)$$

Einsetzen für die mittleren Bodenschichten:

$$ZSDIF_k = \frac{1}{1 + \frac{ZCFH_{k-1}}{\Delta z_k} \cdot (1 - ZEB_{k-1}) + \frac{ZCFH_k}{\Delta z_k}} \cdot \left( ZSDIF_k + \frac{ZCFH_{k-1}}{\Delta z_k} \cdot ZSDIF_{k-1} \right), \text{ für } k \in [2, 4] \quad (3.1g)$$

$$ZEB_k = \left( \frac{1}{1 + \frac{ZCFH_{k-1}}{\Delta z_k}} \right) \cdot \frac{ZCFH_k}{\Delta z_k}, \text{ für } k \in [2, 4] \quad (3.1h)$$

Einsetzen für die unterste Bodenschicht:

$$ZSDIF_5 = \frac{1}{1 + \frac{ZCFH_4}{\Delta z_5} \cdot (1 - ZEB_4)} \cdot \left( ZSDIF_5 + \frac{ZCFH_4}{\Delta z_5} \cdot ZSDIF_4 \right) \quad (3.1i)$$

Rücksubstitution:

$$ZSDIF_k = ZSDIF_k + ZEB_k \cdot ZSDIF_{k+1}, \text{ für } k \in [4, 1] \quad (3.1j)$$

finale Temperaturermittlung:

$$T_{k,t} = ZSDIF_k + \frac{1}{3} \cdot T_{k,t-1}, \text{ für } k \in [1, 5] \quad (3.1k)$$

mit:

$\Delta t$  = Zeitschrittintervall in s

$k$  = Index der Bodenschicht

$t$  = Index des Zeitschritts

$z_k$  = Schichtuntergrenze der Bodenschicht  $k$  in m mit  $z_0 = 0$

$\Delta z_k = z_k - z_{k-1}$  = Schichtdicke der Bodenschicht  $k$  in m

$T_{k,t}$  = Bodentemperatur in der Bodenschicht  $k$  zum Zeitpunkt  $t$  in K

$ZDIFI_k$  = thermale Diffusivität der Bodenschicht  $k$  in  $\frac{m^2}{s}$

$TEM1 = -ZAIR = ZSOFL + ZTRFL + THFL$  = Wärmefluss in den Boden in  $\frac{W}{m^2}$

$ZSOFL = \frac{(1-ALSOL) \cdot SRFL}{1-ALBEDO}$  = einfallende Solarstrahlung an der Oberfläche in  $\frac{W}{m^2}$

$ALSOL$  = Albedo des Landteils eines Pixels

$ALBEDO$  = Gesamtalbedo eines Pixels

$SRFL$  = Nettosolarstrahlung an der Oberfläche in  $\frac{W}{m^2}$

$ZTRFL = EMTER \cdot STBO \cdot TSLM1M^4$  = langwellige Strahlung in  $\frac{W}{m^2}$ , vgl. Formel 2.2

$EMTER$  = Emissivität

$STBO = 5.670475 \cdot 10^{-8} \frac{W}{m^2 K^4}$  = Stefan-Boltzmann Konstante

$TSLM1M$  = Oberflächentemperatur des Landanteils im vorherigen Zeitschritt in K

$TEM2 = -ZDAIR = 4 \cdot ZEMISS \cdot STBO \cdot TSLM1M^3 - DHFT$

= Änderung des Bodenwärmestroms

$ZEMISS = 0.996$  = konstante Emissivität

$DHFT$  = Ableitung des thermalen Wärmeflusses

$THFL$  = thermaler Wärmefluss = Summe aus sensiblen und latenten Wärmefluss in  $\frac{W}{m^2}$

$ZCGN_k$  = volumetrische Wärmekapazität des Bodens in Schicht  $k$  in  $\frac{J}{m^3 K}$

$ZCFH, ZSA, ZSB, ZSDIF_k, ZEB_k$  sind Zwischenergebnisse

Die für diese Berechnung benötigten Bodenparameter für die volumetrische Wärmekapazität ( $c_v$ ) und die Wärmeleitfähigkeit ( $\lambda$ ) werden wie folgt ermittelt:

$$c_v = c_{sol,dry} + \rho_w \cdot \theta \cdot (c_{p,w} \cdot (1 - f_i) + c_{p,i} \cdot f_i) \quad (3.2)$$

$$\lambda = \lambda_{dry} + \left(0.25 + \frac{0.3 \cdot \Delta\lambda}{1 + 0.75 \cdot \Delta\lambda}\right) \cdot \Delta\lambda \cdot \min\left(\frac{4 \cdot \theta}{\theta_s}; 1 + \left(\frac{4 \cdot \theta}{\theta_s} - 1\right) \cdot \frac{1 + 0.35 \cdot \Delta\lambda}{1 + 1.95 \cdot \Delta\lambda}\right) \quad (3.3)$$

mit:

$c_v$  = volumetrische Wärmekapazität des Bodens ( $ZCGN$ )

$c_{sol,dry}$  = volumetrische Wärmekapazität des trockenen Festanteils des Bodens ( $RGCGN$ )

$f_i$  = Anteil des gefrorenen Wassers an der Wassermenge im Boden ( $ZWI$ )

$\theta$  = aktueller Wassergehalt im Boden ( $ZWQ$ )

$\rho_w = 1000 \frac{kg}{m^3}$  = Dichte von Wasser (*RHOH2O*)

$c_{p,w} = 4186.84 \frac{J}{kgK}$  = spezifische Wärmekapazität von Wasser (*CLW*)

$c_{p,i} = 2093 \frac{J}{kgK}$  = spezifische Wärmekapazität von Eis (*ZCPICE*)

$\lambda$  = Wärmeleitfähigkeit des Bodens

$\lambda_{dry}$  = Wärmeleitfähigkeit von trockenem Boden (*TLAMBDA*)

$\Delta\lambda$  = Amplitude der Wärmeleitfähigkeit von gesättigtem zu trockenem Boden (*DLAMBDA*)

$\theta_s$  = Wassergehalt des gesättigten Bodens = Porenvolumen (*PORVOL*)

Die volumetrische Wärmekapazität für den gesamten Boden in REMO (Formel 3.2) entspricht der anteilmäßigen Berechnung der Wärmekapazität, die im Theoriekapitel 2 vorgestellt wurde. Die Formel 3.3 für die Wärmeleitfähigkeit des Gesamtbodens stammt aus der Physik des EMs und DMs (vgl. Majewski, 1991) und stellt eine empirische Modellierung der Wärmeleitfähigkeit auf Grundlage der Daten von Benoit (1976) und van Wijk & de Vries (1963) dar, die beispielsweise auch im Consortium for SmallScale Modeling (COSMO) Modell (Doms et al., 2011) aufgrund des darin integrierten Bodenmodells TERRA eingesetzt wird (vgl. Block, 2007; Di Giuseppe et al., 2011). Die resultierende Kurve für die Wärmeleitfähigkeit abhängig vom Wassergehalt zeigt zwei zusammengesetzte linear steigende Geraden. Für die meist sehr trockenen Böden Afrikas testete Eggert (2011) eine Anpassung dieser Formel im unteren Wertebereich des Wassergehaltes durch eine weitere flachere lineare Gerade mit guten Ergebnissen. Die Anpassung ist jedoch in der verwendeten Standardversion von REMO nicht enthalten und wird daher nicht in dieser Arbeit genutzt.

Die für die Formeln 3.2 und 3.3 benötigten Parameter sind ebenfalls im EM/DM-Manual festgelegt und in Tabelle 3 dargestellt. Es gilt zu beachten, dass der als volumetrische Wärmekapazität von trockenem Boden bezeichnete Wert den zum PWP-Anteil mit Wasser gefüllten Boden definiert und nicht den von ofentrockenem Boden. In einer älteren Version von REMO wurde wie beim EM/DM-Modell zur Berechnung der thermalen Bodenparameter ein mittlerer Wassergehalt verwendet (Schalter *LWDIF=FALSE* im Jobskript). In der in dieser Arbeit verwendeten aktuellen Version (REMO2015) wird der tatsächlich modellierte Wassergehalt im Boden genutzt (*LWDIF=TRUE*). Bei dieser Version werden die Werte für den trockenen Boden benutzt und mit einem Aufschlag an Wärmeleitfähigkeit bzw. Wärmekapazität für den jeweiligen Wassergehalt ( $\theta$  in Formel 3.3) wird der Wert für den feuchten Boden berechnet.

Im Gegensatz zur Bodentemperatur werden die hydrologischen Prozesse im Boden in dem sogenannten „Bucket-Schema“ (Manabe, 1969) modelliert, was bei LSMs der 1. Modellgeneration entspricht (vgl. Kapitel 2). Das im Boden enthaltene Wasser wird als konstant über die gesamte Bodensäule verteilt angenommen, sodass keine vertikale Bewegung innerhalb dieser Säule stattfindet. Ein Wassereintrag findet durch Niederschlag oder Schneeschmelze statt. Diese Infiltration wird in der aktuellen REMO-Standardversion durch das verbesserte Arno-Schema (Hagemann & Gates, 2003) (vorher: Arno-Schema von Dümenil & Todini (1992)) modelliert. Ist der Boden gesättigt, wird die überschüssige Menge Wasser als Oberflächenabfluss gespeichert. Die Verteilung der Bodenwasserkapazität wird neben den verschiedenen Wasserkapazitäten über den Kurvenparameter  $b$  bestimmt, der als Funktion der topographischen Variabilität auch von

Tab. 3: Bodenparameter für die jeweilige Bodentextur (verändert nach: Majewski, 1991)

		Eigenschaft	volumetrische Wärmekapazität von trockenem Boden	thermische Konduktivität von trockenem Boden	Änderung der thermischen Konduktivität	Porenvolumen	Feldkapazität	permanenter Welkepunkt
		Einheit	$10^6 J/Km^3$	$W/Km$	$W/Km$	$m^3/m^3$	$m^3/m^3$	$m^3/m^3$
Bodentexturklasse	1	Sand	1.28	0.30	2.40	0.364	0.196	0.042
	2	sandiger Lehm	1.35	0.28	2.40	0.445	0.260	0.100
	3	Lehm	1.42	0.25	1.58	0.455	0.340	0.110
	4	lehmiger Ton	1.50	0.21	1.55	0.475	0.370	0.185
	5	Ton	1.63	0.18	1.50	0.507	0.463	0.257
	6	Torf	0.58	0.06	0.50	0.863	0.763	0.265

der Modellauflösung abhängt. Bei besonders steilen Gelände wird der Kurvenparameter  $b$  so verändert, dass mehr Oberflächenabfluss generiert wird als bei flachem Gelände (Hagemann & Gates, 2003):

$$b = \begin{cases} b_w + b_{oro}, & \text{für } b_{oro} > 0.01 \\ b_w, & \text{sonst} \end{cases} \quad (3.4)$$

mit:

$b_w$  = Kurvenparameter abhängig von den Subgrid-Feldkapazitäten ( $BETA$ , vgl. Tab. 6)

$$b_{oro} = \max\left(\min\left(\frac{\max(\sigma_{oro} - \sigma_0; 0)}{\sigma_{oro} + \sigma_{max}}; 0.5\right); 0.01\right)$$

$\sigma_{oro} = \sqrt{VAROR}$  = Standardabweichung der Topographie

$$\sigma_0 = 10m$$

$$\sigma_{max} = 820m$$

Durch die Verdunstung des Wassers aus den obersten 10 cm, durch Transpiration aus der gesamten Bodensäule (Hagemann & Stacke, 2015) sowie durch Drainage am unteren Rand des Modellbodens kann der Wassergehalt im Boden abnehmen. Die Erweiterung und die Validation von hydrologischen Prozessen auf die fünf definierten Bodenschichten wird in einer parallel laufenden Dissertation bearbeitet (vgl. Abel et al., 2020).

## 3.4 Benötigte Eingangsdaten für REMO, deren Auflösung und Aufgaben

### 3.4.1 Antriebsdaten

Die in dieser Arbeit verwendeten Antriebsdaten werden dem atmosphärische Reanalyse-Datensatz ERA-Interim des ECMWF entnommen. Dieser ist global in einer Auflösung von  $0.75^\circ$  (ca. 80 km) für den Zeitraum Januar 1979 bis August 2019 verfügbar und ein Folgeprodukt von ERA-40. Durch die Verwendung einer vierdimensionalen Variationsanalyse und durch verbesserte Feuchtigkeitsanalysen sowie durch die Verwendung von neuen Beobachtungssystemen und Biaskorrekturmethode konnte in ERA-Interim im Vergleich zu ERA-40 eine verbesserte Darstellung des hydrologischen Kreislaufs und der stratosphärischen Zirkulation erreicht werden. Der Datensatz liefert verschiedenste Variablen für 00:00, 06:00, 12:00 und 18:00 UTC und die atmosphärischen Variablen liegen für 60 Vertikalschichten vor (vgl. Berrisford et al., 2009; Dee et al., 2011). Anhang B zeigt die von REMO zum Antreiben benutzten Variablen in der Spalte „a-files“. Vor der Verwendung in REMO werden die ERA-Interim-Daten mittels des Präprozessors (vgl. Kapitel 3.2.2.1) auf das Untersuchungsgebiet zugeschnitten und auf das rotierte Modellgitter interpoliert.

### 3.4.2 Bisherige externe Oberflächendaten

Der Einfluss der Landbedeckung, die laut Di Gregorio (2005) die beobachtete (bio-)physikalische Schicht auf der Erdoberfläche darstellt, auf das Klima in unterschiedlichen räumlichen und zeitlichen Skalen wurde in vielen Studien bereits nachgewiesen (z. B. Rechid & Jacob, 2006; Paeth et al., 2009; Wu et al., 2017; Santanello et al., 2018; Gogoi et al., 2019; Jia et al., 2019). Die Art der Bedeckung wirkt sich auf unterschiedliche Prozesse zwischen Erdoberfläche und der darüber liegenden Atmosphäre aus. Sie bestimmt den Austausch von Wasser und Energie zwischen den beiden Systemen. Daher ist es wichtig, dass die Landbedeckung und die daraus abgeleiteten Oberflächenparameter in einem Klimamodell gut abgebildet sind. Je höher die Auflösung des Modells, desto genauer müssen auch diese Daten sein. Um im Klimamodell diese benötigten Oberflächenrahmenbedingungen (engl. boundary conditions) für die Parametrisierung verwenden zu können, werden in REMO standardmäßig die in Tabelle 4 aufgelisteten externen Datensätze und einige von diesen Daten abgeleitete Datensätze zur Erstellung der Bodenbibliothek genutzt. Diese Daten stammen aus unterschiedlichen Quellen und besitzen unterschiedliche Originalauflösungen.

Tab. 4: Übersicht der bisher verwendeten Datensätze in der BodLibKit

Nr. (vgl. Abbildung in Anhang D)	Datensatzbezeichnung	Zweck	Auflösung der REMO-Ausgangsdaten	Referenzen
(1), (2)	GLCCD, LSPII	Landklassen + dazugehörige Oberflächen- parameter	1 km	Eidenshink & Faundeen (1994), Earth Resources Observation And Science (EROS) Center (1994), Olson (1994), Hagemann (2002)
(3), (4), (5)	Meteosat	Albedokorrektur	1 km	Pinty et al. (2000), Knorr et al. (2001), Hagemann (2002)
(6)	GTOPO30	Topographie	1 km	Earth Resources Observation And Science (EROS) Center (1999), Gesch et al. (1999)
(7), (8), (9)	FAO	Bodentexturklassen	50 km	FAO-UNESCO (1971-1981), Zobler (1986)
(10), (11)	MODIS	Albedo für Jahresgang	50 km	Strahler et al. (1999), Lucht et al. (2000), Schaaf et al. (2002), F. Gao et al. (2005), Rechid et al. (2009)
(12)	f_lim	Wachstumsfaktor (klimatologischer Jahresgang für Vegetationsindizes)	50 km	Knorr (1998), Legates & Willmott (1990), Hagemann (2002)

Für die anteilmäßige Berechnung der Vegetationsparameter pro Pixel wird in REMO der Landbedeckungsdatensatz Global Land Cover Characterization Dataset (GLCCD) verwendet, der beispielsweise auch bei der Erstellung der Reanalyse-Daten ERA-Interim benutzt wurde (Dee et al., 2011). Dieser vom United States Geological Survey (USGS) erstellte Datensatz enthält mehrere Landbedeckungsklassifikationen, welche jeweils in einer Auflösung von 1 km für die ganze Welt zur Verfügung stehen. Als Hauptquelle der unüberwachten Klassifikation der 10-Tages Normalized Difference Vegetation Index (NDVI)-Komposite dienten Advanced Very High Resolution Radiometer (AVHRR)-Bilder, die zwischen April 1992 und März 1993 aufgenommen wurden (Eidenshink & Faundeen, 1994). Die Global Ecosystem Klassifikation nach Olson (1994) mit 99 Klassen ist in REMO als Standardklassifikation definiert. Speziell für diese Klassifikation und diesen Datensatz wurde von Hagemann eine Tabelle konzipiert (Hagemann et al., 1999) und später nochmals überarbeitet (Hagemann, 2002), die jeder Landbedeckungsklasse einen Wert für die von REMO benötigten neun Vegetationsparameter zuweist. Diese als LSPI bzw. LSPII bezeichnete Tabelle (vgl. Anhang C) liefert Werte für:

- Vegetationsalbedo  $\alpha_{veg}$ ,
- Rauigkeitslänge der Vegetation  $z_{0,veg}$ ,

- Vegetationsanteil in der Ruhe- und Wachstumsphase (engl. dormancy/ growing phase)  $c_d$  und  $c_g$ ,
- Blattflächenindex (engl. Leaf Area Index) (LAI) ebenfalls in Ruhe- und Wachstumsphase  $lai_d$  und  $lai_g$ , Waldanteil  $c_f$ ,
- Feldkapazität  $W_{cap}$  und
- permanenten Welkepunkt  $f_{pwp}$ .

Als topographischen Datensatz, der zur Berechnung der mittleren Höhe, der Varianz der Höhe und der Rauigkeitslänge eines Pixels benötigt wird, verwendet REMO bisher das globale Digital Elevation Model (DEM) GTOPO30, welches aus verschiedenen globalen und regionalen topographischen Datensätzen generiert und 1996 vom U.S. Geological Survey Center for Earth Resources Observation and Science (EROS) veröffentlicht wurde (Gesch et al., 1999). Der Datensatz besteht aus 33 Kacheln, die zusammen die komplette Erde abdecken, und besitzt eine horizontale Auflösung von 30 Bogensekunden, was ca. 1 km entspricht.

Basierend auf der Bodenkarte der FAO-UNESCO (1971-1981), die von Zabler (1986) überarbeitet wurde und die in der vorliegenden Version eine räumliche Auflösung von  $0.5^\circ$  (ca. 50 km) hat, werden in REMO aus den 15 Klassen des Ausgangsdatsatz fünf Bodentexturklassen und die Klasse „Torf“ reklassifiziert (vgl. Tab. 5). Im eigentlichen Modelllauf dienen diese Texturklassen der Zuweisung der jeweiligen Bodenparameter (vgl. Tab. 3).

Tab. 5: Reklassifikation der FAO-Bodentexturen in der Bodenbibliothek von REMO

Texturklassen			
im überarbeiteten FAO-Datensatz		in REMO	
ID	Textur	ID	Textur
1	Sand	1	Sand
2	lehmgiger Sand	2	lehmgiger Sand
3	sandiger Lehm	2	lehmgiger Sand
4	Lehm	3	Lehm
5	schluffiger Lehm	3	Lehm
6	sandig-toniger Lehm	3	Lehm
7	schluffig-toniger Lehm	4	toniger Lehm
8	toniger Lehm	4	toniger Lehm
9	sandig Lehm	4	toniger Lehm
10	schluffiger Lehm	4	toniger Lehm
11	Ton	5	Ton
12	grob	2	lehmgiger Sand
13	mittel	3	Lehm
14	fein	4	toniger Lehm
15	Torf	6	Torf

Seit 2009 wird in REMO standardmäßig zusätzlich ein Jahresgang von Vegetationsanteil, LAI und Albedo verwendet. Diese drei zusätzlichen Dateien mit jeweils 12 Monatswerten pro Modellpixel (vgl. Abbildung in Anhang D) ersetzen die konstanten Variablen VGRAT, VLT und ALB aus Tabelle 6, wenn im Jobskript, das zum Starten des Klimamodelllaufs benutzt wird, der Schalter „LVEG“ auf TRUE gesetzt wird, was standardmäßig der Fall ist. Die Grundlage

dieser drei Jahreszyklen sind die mittels des Moderate-resolution Imaging Spectroradiometer (MODIS) erhobenen Daten, aus denen Karten für eine schneefreie Oberflächenalbedo und für die Vegetationsalbedo sowie 12 Monatskarten eines sogenannten Wachstumsfaktors erstellt werden (vgl. Hagemann, 2002). Die auf 0.5° interpolierten Karten, die aus den ursprünglich 0.05° aufgelösten MODIS Albedo- und LAI-/FPAR-Produkten (MOD43C1, MOD15A2) abgeleitet sind, liegen bereits in der für den BodLibKit-Prozess benötigten Form vor. Die Aufbereitung und Auswertung der Daten kann bei Rechid et al. (2009) nachgelesen werden.

Aus diesen vorgestellten Datensätzen werden insgesamt 13 zeitlich konstante Variablen (vgl. Tab. 6) und je drei Monatszeitreihen für jeden Pixel des vorher definierten Untersuchungsgebiets in der gewünschten Modellauflösung in dem Präprozess, der vor dem eigentlichen Klimamodelllauf durchgeführt wird, abgeleitet. Die Abbildung in Anhang D zeigt den Prozess mit den jeweiligen Zwischenschritten und den dafür verwendeten Datensätzen (vgl. Tab. 4) als Übersicht. Der Vorgang wird in Anhang D anhand der einzelnen Skripte, die zur Erstellung verwendet werden, genauer erläutert.

Tab. 6: Variablen der bisherigen Bodenbibliothek

Name	Beschreibung	CodeID	Einheit
FIB	Orographie	129	m
BLA	Landseemaske	172	-
AZO	Rauhigkeitslänge	173	m
ALB	Hintergrundalbedo	174	-
VGRAT	Vegetationsanteil	198	-
VAROR	Varianz der Orographie	199	m <sup>2</sup>
VLT	Blattflächenindex	200	-
FOREST	Waldanteil	212	-
FAO	Bodentexturklasse	226	-
WSMX	Feldkapazität	229	m
BETA	Parameter für das Arno Schema	272	-
WMINLOK	minimale Feldkapazität auf dem Submodellgitter	273	m
WMAXLOK	maximale Feldkapazität auf dem Submodellgitter	274	m



# 4 Aufbereitung und Vergleich der neuen Oberflächendatensätze und Validationsdaten

Zur Verbesserung von REMO werden in dieser Arbeit verschiedene neue Datensätze für die Topographie und den Boden aus unterschiedlichen Quellen getestet. In den folgenden Unterkapiteln werden die Methoden zur Anpassung und Aufbereitung dieser Daten für den Einsatz in REMO näher beschrieben sowie die verschiedenen Datensätze miteinander verglichen. Die zusätzlich zur Validation verwendeten Datensätze und deren Aufbereitung werden ebenfalls vorgestellt.

## 4.1 Statistische und programmiertechnische Methoden zur Datenaufbereitung und -analyse

### 4.1.1 Verwendete Interpolationen

#### Horizontale Interpolation

In der Bodenbibliothek werden je nach Datensatz unterschiedliche räumliche Interpolationen vorgenommen, um die Originaldaten auf das rotierte Koordinatensystem mit der gewünschten Modellauflösung zu transformieren. Die bisher verwendeten Methoden wurden im vorherigen Kapitel 3.4.2 und in Anhang D aufgezählt. Aufgrund der erhöhten Auflösung und der damit verbundenen stark vergrößerten Anzahl an zu verarbeitenden Datenpunkten sowie aufgrund der neuen Datensätze und deren Anpassung an die Modelleigenheiten müssen neue Interpolationsmethoden gewählt werden. Die Modifikationen der bisherigen Methoden, der neuen Methoden und ihre jeweilige Verwendung werden in den folgenden Abschnitten kurz vorgestellt.

Eine Anpassung der bereits enthaltenen räumlichen Interpolation des Bodendatensatzes wird wegen der um den Faktor 50 höheren räumlichen Auflösung notwendig. Abbildung 11 zeigt in der oberen Reihe schematisch die bisher zur Ermittlung des Bodentexturwertes eines Modellpixels verwendete Interpolationsmethode in Skript 4 (vgl. Anhang D) für die ersten vier Iterationsschritte. Ausgehend von dem mit „x“ markierten Startpixel wird in jedem Schritt ein um je zwei Pixel in x- und y-Richtung größeres Fenster von oben links nach unten rechts zeilenweise nach einem Landpunkt durchsucht. Der Prozess bricht ab, sobald dieser gefunden wird. Der Texturwert in diesem Pixel des Ausgangsdatsatzes wird dem Modellpixel zugewiesen.

Beim neuen Algorithmus (vgl. untere Reihe in Abb. 11) wird das Prinzip des immer größer werdenden Suchfensters beibehalten. Allerdings werden in jedem Schritt nur die neu ins Fenster gekommenen Pixel geprüft. Dieser Vorteil der kleineren Anzahl an zu testenden Pixeln und einer damit geringeren Rechenzeit des Programms beim neuen Algorithmus macht sich schon bei kleinen Fenstergrößen trotz gleichbleibender Eingangsauflösung bemerkbar. Er potenziert sich um ein Vielfaches, wenn die Auflösung der Eingangsdaten erhöht wird und macht das Interpolieren

bei einer um (50 x 50)fach verbesserten Auflösung in dieser Arbeit auf diese Weise überhaupt erst möglich. Bei der bisherigen Interpolation bricht der Suchalgorithmus außerdem ab, sobald ein geeigneter Landpixel gefunden wurde, ungeachtet dessen, ob nicht noch ein näherer Punkt die Bedingung erfüllt. Dieser Nachteil wird beim neuen Algorithmus ausgeglichen, da alle neu in einem Fenstergrößenschritt dazugekommenen Pixel getestet werden und anschließend der Pixel mit der geringsten Distanz für den Wert des Zielmodellpixels ausgewählt wird.

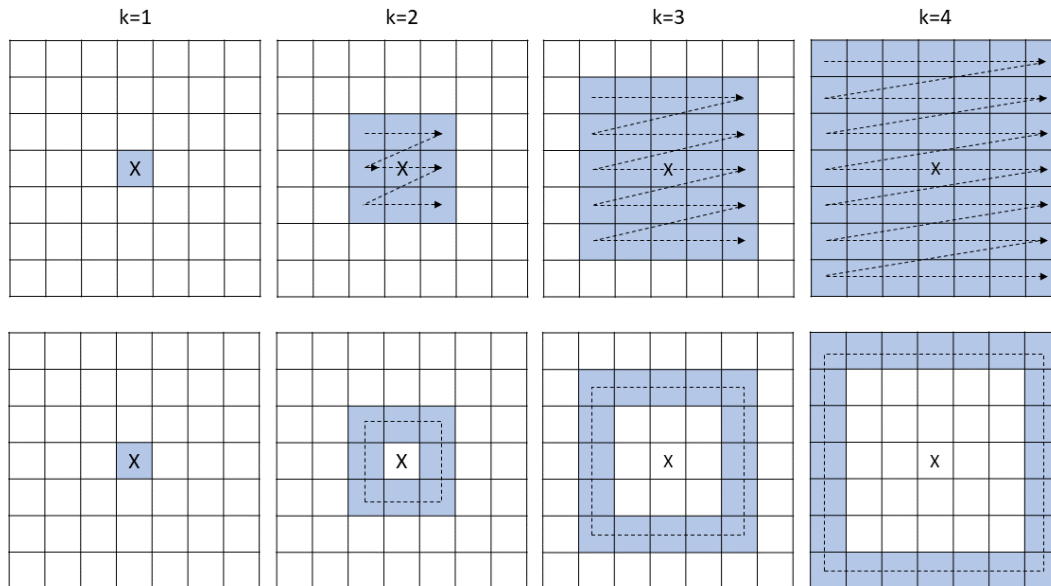


Abb. 11: Schemata des bisherigen und des neuen Suchalgorithmus zur Interpolation des Bodendatensatzes in der BodLibKit  
oben: bisheriger Algorithmus  
unten: neuer Algorithmus  
blau: im aktuellen Iterationsschritt getestete Pixel  
x ist der entsprechende Startpixel für dessen Koordinaten der Texturwert gesucht wird  
... kennzeichnet die Reihenfolge des Tests

### Vertikale Interpolation

Neben der horizontalen Interpolation werden zur Aufbereitung der für unterschiedliche Bodentiefen vorliegenden neuen Bodendaten vertikale Interpolationen benötigt, um diese auf die REMO-Schichten (vgl. Abb. 10) zu bringen. Abhängig von der Variable und der Anzahl der vorliegenden Schichten in den neuen Datensätzen werden verschiedene Interpolationsarten gewählt.

Für die Variablen Sand- und Tonanteil wird eine lineare Zu- bzw. Abnahme der Werte in die Tiefe angenommen. Daher werden für die jeweiligen Schichtanteile lineare Gleichungen auf Basis der Ausgangsdatsätze FAOn und HWSD (nähere Beschreibung siehe Kapitel 4.3.2) gebildet. Die in Formel 4.1 dargestellte Vorgehensweise, einen Anteil  $y$  für eine beliebige Tiefe  $x$  zu berechnen, entspricht einer einfachen linearen Spline-Interpolation. Für tiefer liegende Schichten, für die in den Ausgangsdatsätzen keine Werte vorhanden sind, wird die aufgestellte Formel der darüber

liegenden Schicht verwendet.

$$y_i = m \cdot x_i + t \quad (4.1a)$$

$$y_{i+1} = m \cdot x_{i+1} + t \quad (4.1b)$$

woraus folgt:

$$m = \frac{y_{i+1} - y_i}{x_{i+1} - x_i} \quad (4.1c)$$

$$t = y_i - m \cdot x_i$$

sodass für  $x_i \leq x < x_{i+1}$  gilt:

$$y(x) = m \cdot x + t \quad (4.1d)$$

mit:

$x_i$  = mittlere Tiefe von Schicht i

$y_i$  = Wert von Schicht i

$t$  = y-Achsenabschnitt der linearen Funktion

$m$  = Steigung der linearen Funktion

$y(x)$  = lineare Funktion zur Ermittlung des Wertes für eine beliebige Tiefe x

Für den SoilGrids-Datensatz (nähere Beschreibung siehe Kapitel 4.3.2), der sieben Schichten bis zu einer Tiefe von 2 m differenziert, wird jeweils ein mit der jeweiligen Schichtdicke gewichteter Mittelwert für die fünf REMO-Bodenschichten berechnet:

$$p_1 = \frac{100 \cdot x_1 + 15 \cdot x_2}{100 + 15} \quad (4.2a)$$

$$p_2 = \frac{85 \cdot x_2 + 100 \cdot x_3 + 6.3 \cdot x_4}{85 + 100 + 6.3} \quad (4.2b)$$

$$p_3 = \frac{93.7 \cdot x_4 + 100 \cdot x_5 + 23.2 \cdot x_6}{93.7 + 100 + 23.2} \quad (4.2c)$$

$$p_4 = \frac{76.8 \cdot x_6 + 100 \cdot x_7}{76.8 + 100} \quad (4.2d)$$

$$p_5 = x_7 \quad (4.2e)$$

mit:

$x_i$  = Wert der Soilgrid-Schicht i

$p_i$  = Wert der REMO-Schicht i

Für den organischen Anteil wird wie bei der Erstellung des originalen SoilGrids-Datensatzes (vgl. Hengl et al., 2014) ein exponentieller logarithmischer Verlauf (log-log relationship) angenommen.

Mit den Schichttiefen und Werten aus dem jeweiligen Bodendatensatz kann eine tiefenabhängige Formel abgeleitet werden:

aus:

$$y_i = \exp^{(a+b \cdot \ln(x_i))} \quad (4.3a)$$

$$y_{i+1} = \exp^{(a+b \cdot \ln(x_{i+1}))} \quad (4.3b)$$

folgt:

$$b = \frac{\ln y_{i+1} - \ln y_i}{\ln x_{i+1} - \ln x_i}$$

$$a = \ln y_i - b \cdot \ln x_i$$

für:

$$y(x) = \exp^{(a+b \cdot \ln(x))} \quad (4.3c)$$

mit:

$x_i$  = mittlere Tiefe der Schicht  $i$

$y_i$  = Wert der Schicht  $i$

$a, b$  = Kurvenparameter der log-log-Funktion

$y(x)$  = log-log-Funktion zur Ermittlung der Parameter für eine beliebige Tiefe  $x$

Die mit dieser Methode auf die fünf REMO-Schichten interpolierten Bodenprofile werden in Kapitel 5.1.2 miteinander verglichen. Neben diesen Anpassungen auf horizontaler und vertikaler räumlichen Ebene werden auch „inhaltliche“ Anpassungen im Code durch die neuen Daten notwendig, weil diese eine feinere Gliederung der Texturklassen enthalten, für die bisher jedoch keine Bodenparameter (vgl. Tab. 3) ausgewiesen sind.

### **Interpolation für Texturmischklassen**

Bei der Approximation der Bodenparameter aus vorhandenen Stützstellen (vgl. eigene PTF in Kapitel 5.1.1) wird mit einer gewichteten bilinearen Extrapolation gearbeitet. Der durch Stützstellen für einen Parameter  $p$  definierte lineare Zusammenhang hängt dabei von zwei Komponenten ab. Für beide Komponenten wird eine Regressionsgerade mit Steigung  $m$  und Achsenabschnitt  $t$  so definiert, dass der Abstand der Punkte zur Geraden minimiert ist (nähere Beschreibung dieser „Kleinste-Quadrate-Methode“ in Kapitel 6.1.4). Anschließend werden die Werte der beiden Geradengleichungen mit dem jeweiligen Korrelationskoeffizienten gewichtet gemittelt, um den Einfluss beider Korngrößenfraktionen auf die Bodenparameterwerte zu berücksichtigen (vgl. Formeln 4.4).

$$m_x = \frac{\sum_{i=1}^n (x_i - \bar{x}) \cdot (p_i - \bar{p})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (4.4a)$$

$$m_y = \frac{\sum_{i=1}^n (y_i - \bar{y}) \cdot (p_i - \bar{p})}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

$$t_x = \bar{p} - m_x \cdot \bar{x}$$

$$t_y = \bar{p} - m_y \cdot \bar{y}$$

$$r_x = \frac{\sum_{i=1}^n (x_i - \bar{x}) \cdot (p_i - \bar{p})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^n (p_i - \bar{p})^2}} \quad (4.4b)$$

$$r_y = \frac{\sum_{i=1}^n (y_i - \bar{y}) \cdot (p_i - \bar{p})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2} \cdot \sqrt{\sum_{i=1}^n (p_i - \bar{p})^2}}$$

$$p(x, y) = \frac{r_x \cdot (m_x \cdot x + t_x) + r_y \cdot (m_y \cdot y + t_y)}{r_x + r_y} \quad (4.4c)$$

mit:

$x_i$  = zugewiesener Sandanteil der Texturklasse i

$y_i$  = zugewiesener Tonanteil der Texturklasse i

$p_i$  = Parameterwert für Texturklasse i

$\bar{p}$  = Mittelwert des Parameters über alle Texturklassen

$p(x, y)$  = Funktion zur Ermittlung des gesuchten Parameterwert für einen beliebigen Sandanteil x und Tonanteil y

$\bar{x}$  = Mittelwert des Sandanteils aller Texturklassen

$\bar{y}$  = Mittelwert des Tonanteils aller Texturklassen

Durch diese Interpolationsmethoden kann der Korngrößenanteil in jeder beliebigen Tiefe geschätzt und darauf aufbauend für jeden Parameter  $p$  ein Wert für alle beliebigen Korngrößenanteilkombinationen ermittelt werden. Die Ergebnisse der Anwendung dieser Methoden auf die Daten und ein Vergleich mit den bisherigen Stützstellen wird in Kapitel 5.1.1 präsentiert.

## 4.1.2 Programmiertechnische Änderungen

### Automatisiertes Erstellen und Einlesen der Eigenschaften der topographischen Kacheln

Durch die höher aufgelösten Datensätze müssen größere Datenmengen verarbeitet werden. Daher ist es notwendig, dass nicht nur das Einlesen der Daten angepasst, sondern auch der technische Prozess optimiert wird. Bisher werden beispielsweise die geographischen Koordinaten

aller 33 Einzelgebiete des REMO Topographie-Datensatzes GTOPO einzeln als Hardcode im Programmskript (*second.sh*), das zur Interpolation auf die gewünschte Modellauflösung benutzt wird, definiert. Für die 22.912 Kacheln des ASTER-Datensatzes, der als Alternative eingebaut werden soll, wäre das mit einem großen Aufwand verbunden und würde die Übersichtlichkeit des Codes enorm verschlechtern. Deshalb wurde ein Hilfsprogramm konzipiert, das einmalig bei der Integration eines neuen Datensatzes kalibriert und ausgeführt werden muss. Es berechnet die einem regelmäßigen Muster folgenden Eckpunktkoordinaten jeder Kachel und schreibt diese in eine Datei, die anschließend vom Hauptprogramm zu Beginn eingelesen wird und somit für jedes beliebige Modellgebiet verfügbar ist.

### **Automatisierte Auswahl der für ein Untersuchungsgebiet benötigten topographischen Kacheln**

Wie in Kapitel 3.4.2 und Anhang D ausführlich beschrieben, werden beim Vorbereiten einer Bodenbibliothek für ein Untersuchungsgebiet bisher automatisch alle 33 Kacheln in das richtige Verzeichnis des Modellgebiets kopiert und eingelesen. Der Benutzer konnte vor dem eigentlichen Prozessieren manuell eine Selektion vornehmen, um überflüssige Daten zu entfernen, sodass diese nicht eingelesen werden mussten. Durch die enorme Steigerung der Anzahl der Kacheln wird das manuelle Auswählen unmöglich und das Einlesen des gesamten hochaufgelösten globalen Datensatzes würde einen großen Arbeitsspeicher benötigen und sehr lange dauern. Daher stellt der neu eingebaute Point-in-Polygon (PiP)-Algorithmus eine weitere Optimierung in der Bodenbibliothek dar.

Es gibt verschiedene Ansätze, um herauszufinden, ob ein Punkt mit festgelegten Koordinaten innerhalb, auf den Kanten oder außerhalb eines Polygons mit beliebig vielen Ecken, welche allerdings durch Koordinaten ebenfalls fest definiert sind, liegt. Die Grundlage für den häufig genutzten „verbesserten Nordbeck and Rystedt-Algorithmus“ von Sloan (1985), der für alle Polygonarten gilt, ist die richtige Reihenfolge der Ecken und die Berechnung der Determinanten. Sind die Ecken eines konvexen Polygons gegen den Uhrzeigersinn sortiert, dann liegt ein Punkt, der innerhalb dieses Polygons liegt, immer links von allen Kanten. Um mathematisch herauszufinden, ob ein Punkt links oder rechts der Verbindungslinie liegt, die durch zwei Eckpunkte definiert wird, erzeugt man ein Dreieck mit allen drei Punkten. Liegt der Punkt links, ist die Fläche dieses Dreiecks positiv. Diese Fläche wird über die Determinante berechnet (vgl. Abb. 12). Man nutzt aus, dass jedes beliebige Dreieck in einem Koordinatensystem zu einem Parallelogramm (ABA'D) ergänzt werden kann, welches den doppelten Flächeninhalt des ursprüngliche Dreieck ( $\Delta ABC$ ) besitzt. Der Flächeninhalt dieses Parallelogramms lässt sich leicht durch das Abziehen des Flächeninhaltes von zwei kleineren Rechtecken (rosa und gelb) vom Flächeninhalt des größeren Rechtecks, das das Parallelogramm umschließt, berechnen. Die mathematische Formulierung der Flächenberechnung zeigt Formel 4.5.

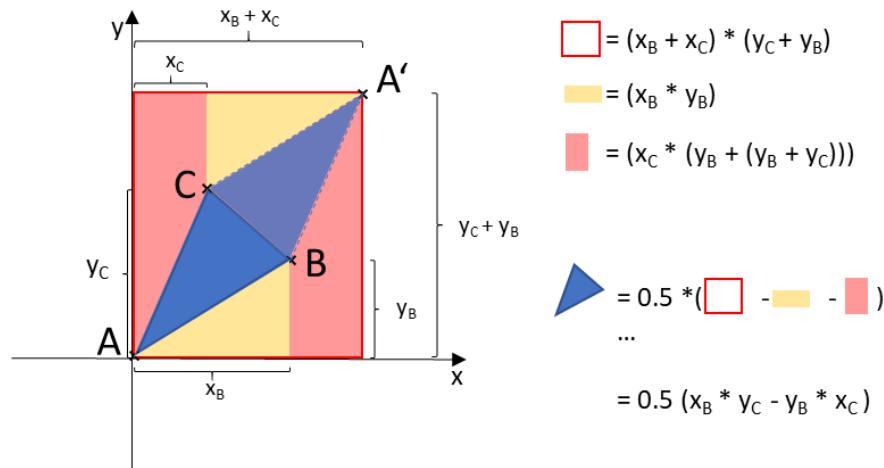


Abb. 12: Graphische Darstellung und Herleitung der Berechnung des Flächeninhaltes eines Dreiecks über die Determinante

$$Fläche_{\Delta ABC} = \frac{1}{2} \cdot \begin{vmatrix} x_B & x_C \\ y_B & y_C \end{vmatrix} = \frac{1}{2} \cdot (x_B \cdot y_C - y_B \cdot x_C) \quad (4.5)$$

mit:

$$A = (x_A/y_A)$$

$$B = (x_B/y_B)$$

$$C = (x_C/y_C)$$

$x_i$  = x-Koordinate (hier: Longitude) des Punktes i

$y_i$  = y-Koordinate (hier: Latitude) des Punktes i

Ein weiterer sehr effektiver PiP-Algorithmus, der für alle Arten von zweidimensionalen Polygonen gilt, basiert auf dem Jordanschen Kurvensatz. Bei diesem wird gezählt, wie oft eine von dem in Frage kommenden Punkt ausgehende Linie die Seiten des Polygons schneidet. Liegt kein Schnittpunkt oder eine gerade Anzahl an Schnittpunkten vor, befindet sich der Punkt außerhalb, ansonsten innerhalb des getesteten Polygons (Franklin, 2018).

Für die neue Bodenbibliothek wird deshalb beim Setup des Modellgebiets mittels des von Douglas (1974) veröffentlichten Codes, der dem vorgestellten und ebenfalls getesteten Code von Sloan (1985) bezüglich der Geschwindigkeit überlegen war, für jeden Pixel im Modellgebiet geprüft, ob dieser in einer topographischen Kacheldatei enthalten ist. Ist dies der Fall, wird der Name dieser Kachel in eine Liste aufgenommen. Nur die sich in dieser Liste befindenden Kacheln werden anschließend in das Arbeitsverzeichnis kopiert und eingelesen. Zum einen wird dadurch der Speicherplatz für ein jeweils ausgewähltes Untersuchungsgebiet minimiert und zum anderen wird die Rechenzeit enorm verkürzt.

Zu beachten ist, dass die Auswahl der Kacheln notwendig ist, bevor die regulären Koordinaten des Untersuchungsgebiets berechnet sind, da das Modellgebiet vom Benutzer durch die Angabe der Koordinaten der linken unteren Ecke im rotierten System, der Auflösung und der Anzahl der Pixel definiert wird. Die Koordinaten für jeden Pixel im Gebiet liegen daher nur im rotierten

Zustand als regelmäßige Koordinaten vor (vgl. Abb. 3 in Kapitel 3.2.1). Überführt man die Eckpunkte der in Frage kommenden Kacheln ebenfalls ins rotierte Koordinatensystem, kann sich durch die geometrische Verzerrung bei der Rotation die Lage der Eckpunkte zueinander ändern. Daher wird die für den PiP-Algorithmus benötigte Reihenfolge kontrolliert und bei Bedarf werden die Eckpunkte neu sortiert. Dieses Sortieren der mit Formel 4.6 berechneten Winkel von klein nach groß bringt die benötigte Reihenfolge entgegen dem Uhrzeigersinn.

$$\text{angle}_i = \text{atan2}(y_i - \bar{y}, x_i - \bar{x}) \quad (4.6)$$

mit:

$\text{angle}_i$  = Winkel des Referenzpunkts (hier: Mittelpunkt des Polygons) zum Punkt  $i$

$x_i$  = Longitude des Punktes  $i$

$y_i$  = Latitude des Punktes  $i$

$\bar{x}$  = Mittelwert aller Longituden der Punkte des Polygons = Longitude des Mittelpunktes

$\bar{y}$  = Mittelwert aller Latituden der Punkte des Polygons = Latitude des Mittelpunktes

### **Beschleunigung der Varianzberechnung**

Des Weiteren können durch das Setzen von If-Bedingungen und durch geschicktes Umformulieren von Gleichungen, wie beispielsweise die Formel zur Berechnung der Varianz, Rechenschritte gespart werden, was die Rechenzeit massiv minimiert.

$$\text{VAR}(x) = E[(x - \mu)^2] = E[x^2] - E[x]^2 \quad (4.7a)$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{n}{n-1} \left( \frac{\sum_{i=1}^n x_i^2}{n} - \left( \frac{\sum_{i=1}^n x_i}{n} \right)^2 \right) \quad (4.7b)$$

Nutzt man die Umformulierung der Varianzformel (vgl. Formel 4.7b) von einem Two-Pass zu einem One-Pass Algorithmus (in dieser Form auch naiver Algorithmus oder Textbook One Pass genannt), wird ein erneutes Durchgehen aller Pixel überflüssig, da der Mittelwert nicht benötigt wird, um diesen von den einzelnen Werten abzuziehen. Dadurch kann mit einem Schleifendurchlauf sowohl der Mittelwert als auch die Varianz berechnet werden. Allerdings muss hierbei beachtet werden, dass bei großen Stichprobenumfängen, die zu großen Werten bei den beiden Erwartungswerten führen, größere Ungenauigkeiten aufgrund von Rundungsfehlern auftreten können (vgl. Kamat & Nandi, 2017, Schubert & Gertz, 2018). Um dem entgegen zu wirken, werden alle beteiligten Variablen konvertiert und die Rechnungen als Double Precision statt Single Precision durchgeführt.



## 4.2 Analysemethoden zum Vergleich verschiedener Datensätze

Neben den typischen Kennzahlen wie arithmetischem Mittelwert, Varianz und Standardabweichung, eignen sich auch andere Lagemaße zur Beschreibung und zum Vergleich verschiedener Daten. Vor allem bei Datensätzen mit extremen Ausreißern helfen die folgenden Werte, um die Charakteristik einer Stichprobe erfassen und analysieren zu können.

Als Minimum bzw. Maximum werden der erste bzw. letzte Wert einer sortierten Stichprobe bezeichnet. Der Median, der auch zweites Quartil oder 0.5 bzw. 50% Quantil genannt wird, ist hingegen der Wert an der mittleren Stelle. Mathematisch ausgedrückt:

$$q_{0.5} = \begin{cases} x(\frac{n+1}{2}), & \text{für einen ungeraden Stichprobenumfang } n \\ \frac{x(\frac{n}{2}) + x(\frac{n}{2} + 1)}{2}, & \text{für einen geraden Stichprobenumfang } n \end{cases} \quad (4.8)$$

für einen sortierten Datensatz bei dem gilt:  $x(i-1) \leq x(i) \leq x(i+1)$

Als erstes und drittes Quartil (bzw. 25% und 75% Quantil) werden die Werte an den mittleren Stellen zwischen Minimum und Median bzw. Median und Maximum bezeichnet. Die Differenz des dritten und ersten Quartils wird als Interquartilsabstand (IQR) bezeichnet. Der prozentuale Wert der Quantile gibt an, wie viele Werte der Stichprobe unterhalb des jeweiligen Wertes liegen. Ganzzahlige Anteile werden auch Perzentile genannt. Die Perzentile und der IQR sind anders als Mittelwert und Standardabweichung nicht so stark von Einzelwerten im Datensatz abhängig.

Eine geeignete Darstellung dieser Kennziffern gibt der sogenannte Boxplot oder Box- und Whiskerplot. Diese Art der Darstellung wurde von Tukey (1977) eingeführt. Die Box wird dabei immer von 1. und 3. Quartil definiert und vom Median durchschnitten. Die beiden Whisker, die von der Ober- und Untergrenze der Box zu den jeweiligen Extremen zeigen, können je nach Methode unterschiedliche Längen annehmen. Bei der einfachsten Methode enden die Whisker beim tatsächlichen Minimal- und Maximalwert. Bei anderen gängigen Methoden werden die Längen durch den 1.5-fachen Interquartilsabstand definiert oder sie reichen bis zum 2.5% bzw. 97.5% Quantil. In dieser Arbeit wird immer die Methode mit dem 1.5fachen IQR zur Darstellung der Werteverteilung einer Stichprobe verwendet. Über- oder unterschreitet diese jeweilige Whiskerlänge den Maximum- oder Minimumwert des Datensatzes wird der berechnete Whiskerwert auf den jeweiligen Extremwert gesetzt (Wilks, 2006).

Möchte man noch genauere Einblicke in die Verteilung der Werte im Wertebereich erhalten, helfen Histogramme. Diese stellen die absolute oder relative Häufigkeit der in Intervallklassen eingeteilten Werte als Balken oder Säulen dar. An einem Histogramm kann der Wertebereich, die Verteilung und die Symmetrie einer Stichprobe abgelesen werden. Die mit dem Histogramm zusammenhängende Darstellung der (kumulierten) Häufigkeitsverteilung bietet im Vergleich zum Säulendiagramm aufgrund der linienhaften Darstellung eine noch bessere Vergleichsmöglichkeit verschiedener Verteilungen. Eng mit dieser Darstellungsart für die Häufigkeiten einer Stichprobe verwandt ist die Kerndichteschätzung (engl. Kernel Density Estimation (KDE) oder Kernel

Density Smoothing). Sie ist eine statistische Methode zum Schätzen der Wahrscheinlichkeitsverteilung. Aus einer Stichprobe wird eine Funktion der Verteilung der Grundgesamtheit geschätzt. Durch die Platzierung der Klassenmittelpunkte bei der Darstellung eines Histogramms können gleiche Stichprobendaten trotz gleicher Klassenbreite unterschiedliche Verteilungen anzeigen. Dieser Effekt wird durch die Verwendung des Kernel Density Smoothings egalisiert.

$$d_t = \frac{1}{n \cdot h} \sum_{j=1}^n K\left(\frac{(x_j - t)}{h}\right) \quad (4.9)$$

mit:

$n$  = Stichprobenumfang

$x_j$  = j-ter x-Wert einer Stichprobe

$t$  = x-Wert, für den die Wahrscheinlichkeitsdichte geschätzt werden soll

$h$  = Klassenbreite bzw. Bandbreite bzw. Smoothing Parameter

$K()$  = ausgewählte Kernfunktion

Eine Kernfunktion erfüllt folgende Eigenschaften:

1. symmetrisch, d. h.  $K(t) = K(-t)$
2.  $\int_{-\infty}^{\infty} K(t) dt = 1$
3.  $K(t) \geq 0$ , für  $-\infty < t < \infty$  (gilt nicht für higher order kernels, vgl. Marron (1994))
4.  $\lim_{t \rightarrow \infty} K(t) = \lim_{t \rightarrow -\infty} K(t) = 0$

Es existieren verschiedene Kernfunktionen (Kernels), die diese Eigenschaften erfüllen, mit unterschiedlichen Formeln und daraus resultierenden Eigenschaften. Der in dieser Arbeit verwendete Epanechnikov-Kern minimiert beispielsweise die mittlere quadratische Abweichung zum Kerndichteschätzer und hat folgende Formel:

$$K(t) = \begin{cases} \frac{3}{4} \cdot (1 - t^2), & \text{wenn } t \in [-1;1] \\ 0 & \text{sonst} \end{cases} \quad (4.10)$$

Die Kerndichteschätzung hängt neben der Wahl der Kernfunktion auch von der gewählten Bandbreite  $h$  ab. Es existieren verschiedene Formeln für die optimale Wahl für  $h$ , z. B. abhängig vom Stichprobenumfang oder vom IQR. Wird eine zu kleine Bandbreite gewählt, ergibt die resultierende Funktion zu viele Peaks. Analog dazu wird die Funktion bei einer zu großen Bandbreite zu stark geglättet. Außerdem ist zu beachten, dass die Eignung der Formel zur Ermittlung der geeigneten Bandbreite teilweise stark von den zugrunde liegenden Verteilungen beeinflusst wird. Als Richtwert und besonders geeignet für normalverteilte Daten gelten die sogenannten „rules of thumb“ (RoT) von Scott (1992) und Silverman (1986) und deren Anpassungen (vgl. Sheather, 2004; Wilks, 2006; Węglarczyk, 2018). In dieser Arbeit wird die Silverman-RoT (Formel 4.11) als Bandbreite verwendet.

$$h = 0.9 \cdot \min\left(s, \frac{IQR}{1.34}\right) \cdot n^{\frac{-1}{5}} \quad (4.11)$$

mit:

$n$  = Stichprobenumfang

$s$  = Standardabweichung

$IQR$  = Interquartilsabstand

## 4.3 Beschreibung, Aufbereitung und Einbau der neuen Oberflächendaten

Das Ziel bei der Auswahl der neuen Eingangsdatensätze war eine verbesserte räumliche Auflösung, damit diese neue realistischere Landoberflächenparameter für REMO für eine sehr hohe Modelauflösung liefern. Gleichzeitig war ein weiteres Auswahlkriterium die globale Abdeckung der neuen Datensätze, sodass wie bisher auch ein Datensatz die Informationen für jedes ausgewählte Untersuchungsgebiet zur Verfügung stellen kann, damit die Landoberflächenparameter für das Klimamodell in sich konsistent bleiben. Außerdem wurde darauf geachtet, dass der Informationsgehalt (z. B. Anzahl an Texturklassen), der in den Datensätzen enthalten ist, gleichwertig oder sogar höher ist.

### 4.3.1 Neue Topographie-Datensätze

Einen großen Anteil bei der Verbesserung des Modelloutputs von GCMs auf RCMs liefert die verbesserte Abbildung der Orographie, da durch diese genauere Darstellung vor allem bei komplexer Topographie Parameter wie Temperaturgradienten und Niederschlagsschatteneffekte auf Leeseiten von Gebirgen besser abgebildet werden können (Di Luca et al., 2015; Cabos et al., 2019). Deshalb ist eine gute Repräsentation der Orographie (FIB) und der davon abgeleiteten Parameter Varianz der Orographie (VAROR) und Rauigkeitslänge der Orographie (AZO) sehr wichtig, wenn die Modellauflösung weiter erhöht werden soll.

Die topographischen Daten stammen aus Satellitendaten. Es gibt sehr viele unterschiedliche Produkte. In Tabelle 7 sind die in dieser Arbeit verwendeten Datensätze mit ihren jeweiligen Aufnahmedaten und Informationen aufgelistet. Wie bereits in Kapitel 3.4.2 erwähnt, hat der bisherige Datensatz GTOPO30 eine Auflösung von ca. 1 km und wird aufgrund der globalen Abdeckung und der nachgewiesenen Präzision oft verwendet, z. B. im High Resolution Atmospheric Model (HIRAM) (Undén et al., 2002), für den Beobachtungsdatensatz E-OBS oder für den Reanalyse Datensatz ERA-Interim (vgl. Kapitel 4.5.1). Er dient in allen folgenden topographischen Analysen als Referenzwert.

Die Shuttle Radar Topography Mission (SRTM) der NASA, der National Geospatial-Intelligence Agency (NGA), des Deutschen Luft- und Raumfahrtzentrum (DLR) und der Italien Space Agency

#### 4 Aufbereitung und Vergleich der neuen Oberflächendatensätze und Validationsdaten

Tab. 7: Allgemeine Informationen zu den verwendeten topographische Datensätzen

Name	Datensatz-Version	Abdeckung	Auflösung	Aufnahmejahre	Quelle	Literatur
GTOPO	GTOPO30	90° N/S	30" / 1 km	1993- 1996	USGS EROS	Gesch et al. (1999)
SRTM	SRTM3 (V2.1)	56° N/ 60° S	3" / 90 m	2000	NASA, USGS	Farr et al. (2007), Rodriguez et al. (2006)
ALOS	AW3D30	83° N/S	1" / 30 m	2006- 2011	EORC, JAXA	Takaku et al. (2014), Tadono et al. (2014)
TANDEM	TANDEM- X-DEM	90° N/S	3" x 3-30"/ 90-900 m	2018	DLR	Wessel et al. (2018), Rizzoli et al. (2017)
ASTER	ASTER GDEM2	83° N/S	1" / 30 m	2009/ 2011	METI, NASA, USGS	Tachikawa, Kaku et al. (2011)

(ASI), die im Jahr 2000 startete, kennzeichnete einen 11-tägigen sinusoidalen Orbitflug des Spaceshuttles Endeavour (11.-22. Februar 2000) in einer Flughöhe von 233 km mit dem Ziel ein Höhenmodell der Erde zu erzeugen. Dabei wurde ein Radarsignal aktiv ausgesendet. Dieses zurückgestrahlte Signal wurde von den Radarsensoren im C- (NASA) und X-Band (DLR) gemessen, die sich an einer Antenne am Shuttle und an der Antenne, die an der Spitze eines 60 m langen Mast des Spaceshuttles angebracht war, befanden. Durch den Vergleich der beiden gleichzeitig aber mit zwei unterschiedlichen Antennen und dadurch mit leicht unterschiedlichen Aufnahmewinkel aufgenommenen Radarbildern kann mittels eines Interferogramms ein digitales Geländemodell (DGM) erstellt werden (Single Pass Verfahren), das zudem unabhängig von der jeweiligen Wolkenbedeckung und Sonneneinstrahlung ist (Farr et al., 2007). Der in der Literatur (vgl. Vergleich der Topographiedaten in Kapitel 4.4.1) als sehr genau beschriebene und sehr oft verwendete SRTM-Datensatz liegt allerdings nur für den Bereich zwischen 56°N und 60°S vor. Damit erfüllt SRTM nicht die Ansprüche einer globalen Abdeckung. Trotzdem wurde der SRTM-Datensatz aufbereitet und eingebaut, um ihn mit allen anderen ausgewählten Datensätzen vergleichen zu können. Bei der eingebauten Version handelt es sich um die Void Filled Version V2.1, der in der Produktdokumentation eine vertikale Genauigkeit von unter 16 m bescheinigt wird (Earth Resources Observation And Science (EROS) Center, 2017).

Die beiden nächsten Datensätze ASTER GDEM (Tachikawa, Kaku et al., 2011) und ALOS AW3D DEM (Tadono et al., 2014, Takaku et al., 2014) wurden auf Grundlage von optischen Satellitenbildern generiert. Sie können damit den passiven Systemen zugeordnet werden, weshalb Fehlwerte z. B. aufgrund von Wolkenbedeckung in den Originaldaten auftreten können. Im Gegensatz zu den aktiven Systemen senden die passiven Systeme keine eigene Strahlung aus, sondern empfangen nur reflektiertes Sonnenlicht oder Infrarotstrahlung (Fu et al., 2020).

Das Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) wurde im Dezember 1999 an Bord des Terra Spacecraft der NASA ins Weltall geschossen. Mit zwei Teleskopen werden auf 14 verschiedenen Bändern Werte von verschiedenen Wellenlängenbereichen aufgenommen. Durch die Aufnahme von Stereobildern kann ein DEM erzeugt werden. Dabei wird eine maximale horizontale Auflösung von 15 m erreicht. Jede Szene deckt einen Bereich von ca. 60 km × 60 km ab. Der ASTER-Datensatz wird von mehreren Quellen mit einer vertikalen Genauigkeit von ca. 10–25 m angegeben (z. B. Tachikawa, Kaku et al., 2011; Abrams et al., 2020). Die in dieser Arbeit verwendete Version 2 wurde 2011 veröffentlicht und stellt durch mehr aufgenommene Szenen und eine verbesserte Land-Meer-Maske eine Verbesserung hinsichtlich räumlicher Artefakte und Fehlwerte im Vergleich zu vorherigen Versionen dar (Meyer et al., 2012).

Der von dem Panchromatic Remote-sensing Instrument for Stereo Mapping (PRISM) des Advanced Land Observing Satellite (ALOS) aufgenommene ALOS World 3D 30 m (AW3D30) Datensatz, dessen Stereobilder aus Aufnahmen von 2006-2011 stammen (Szypuła, 2019), ist dafür bekannt, dass er größere Fehlwertbereiche aufgrund von nicht existenten Daten in den höheren Breiten besitzt (Abrams et al., 2020). Weil die Mission wegen eines Energieproblems 2011 für beendet erklärt wurde, können diese Fehlwerte nicht mehr direkt aufgefüllt werden (Tadono et al., 2014), weswegen versucht wird mittels anderer hochaufgelöster Datensätze die Datenlücken zu schließen. Die kommerzielle Version des ALOS-Datensatzes, aus dem die vorliegenden Daten (30 m Auflösung) abgeleitet wurden, besitzt sogar eine Auflösung von 5 m (Takaku et al., 2016).

Der TANDEM-X DEM-Datensatz ist ein Produkt der TanDEM-X Mission (TerraSAR-X add-on for Digital Elevation Measurements) des DLR. Die Messung beruht auf Radarinterferometrie (InSAR) von zwei Satelliten. Die horizontale Originalauflösung beträgt 12 m, wobei die angegebene absolute Genauigkeit im Bereich von unter 10 m liegt. Der in dieser Arbeit verwendete 30 m-Datensatz wurde aus diesem Datensatz aggregiert (Rizzoli et al., 2017).

Folgende topographische Datensätze haben ein oder mehrere Auswahlkriterien nicht erfüllt oder sind abgeleitete Produkte der oben genannten Datensätze und wurden daher nicht in dieser Arbeit verwendet: ETOPO1/2 (Amante, 2009; NOAA National Geophysical Data Center, 2006), WorldDEM (Riegler et al., 2015) (12 m Auflösung, kommerzielles TanDEM-X), Multi-Error-Removed Improved-Terrain DEM (MERIT DEM) (Yamazaki et al., 2017) (Kombination aus SRTM3 v2.1 und AW3D30 v1) oder NASADEM (NASA JPL, 2020) (basiert auf einer Neuprozessierung von SRTM Daten).

In Tabelle 8 sind die Eigenschaften der fünf ausgewählten Datensätze zusammengefasst. Es fällt auf, dass die Daten in unterschiedlichen Dateiformaten vorliegen. Um eine möglichst gute Automatisierung der Prozessierung in der Bodenbibliothek zu gewährleisten, wurden alle neuen Datensätze in das netCDF Dateiformat umgewandelt. Außerdem spielt auch der Endian (Byte-Reihenfolge), in dem die Daten vorliegen, eine wichtige Rolle. Abhängig vom System, auf dem die Bodenbibliothek erzeugt werden soll, und abhängig vom durch den Benutzer ausgewählten Topographiedatensatz werden die Daten automatisch in das richtige Endianformat konvertiert.

Tab. 8: Eigenschaften der verwendeten topographische Datensätzen im Vergleich

Name	GTOPO	SRTM	ALOS	TANDEM	ASTER
originales Dateiformat	.DEM	.hgt	.tif	.tif	.tif
Anzahl der Kacheln	33	14280	22125	19389	22912
Anzahl der Pixel pro Kachel	3600 x 7200 6000 x 4800	1201 x 1201	3601 x 3601	1201 x 1201 801 x 1201 961 x 1201 481 x 1201	3601 x 3601
EGM / vertikales Datum	orthometrische Höhe (EGM96 Geoid auf WGS84 Ellipsoid)	orthometrische Höhe (EGM96 Geoid auf WGS84 Ellipsoid)	orthometrische Höhe (EGM96 Geoid auf WGS84 Ellipsoid)	Ellipsoidhöhe auf WGS84 Ellipsoid	orthometrische Höhe (EGM96 Geoid auf WGS84 Ellipsoid)
Fehlwert	-9999	-32768	-9999	-32767	-9999
Endian	big endian	big endian	big endian	little endian	little endian
Datenformat	16-bit signed integer	16-bit signed integer	16-bit signed integer	32-bit float single precision	16-bit signed integer

Für den TANDEM-Datensatz ist in Tabelle 8 zu erkennen, dass dieser in einem anderen Earth Gravitational Model (EGM) vorliegt. Das EGM96-Geoid ist ein Geopotentialmodell, auch vertikales Datum genannt, das Höhen und Tiefen der Erdoberfläche simuliert. Das WGS84-Ellipsoid hingegen ist eine leicht zusammengestauchte Kugel mit glatter Oberfläche. Die in den Datensätzen enthaltenen Höhen sind als die relativen Höhen über dem jeweiligen Referenzellipsoid oder -geoid als orthometrische ( $H$ ) bzw. ellipsoidische ( $h$ ) Höhe definiert (vgl. Abb. 14). Um auch den TANDEM-Datensatz in das gleiche Referenzsystem zu übertragen, wurde der WGS84-EGM96-Geoid-Höhen-Datensatz der NASA (Lemoine et al., 1998) heruntergeladen. Dieser Datensatz, der in einem 15 min-Gitternetz vorliegt (vgl. Abb. 13), wurde auf die jeweilige Modellauflösung jeder TANDEM-Kachel interpoliert und anschließend von der jeweiligen Kachel subtrahiert (vgl. Formel 4.12).

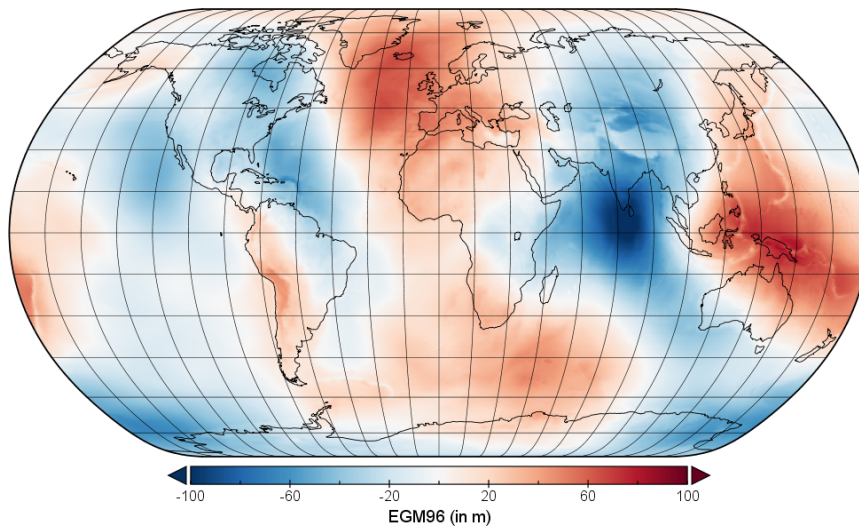


Abb. 13: Weltkarte der EGM96 Geoid Höhe (Datenquelle: Lemoine et al., 1998)

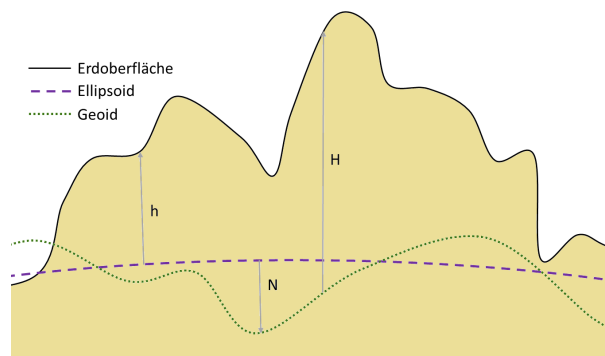


Abb. 14: Schaubild der Gravitationsmodelle (in Anlehnung an: Heck & Seitz, 2016)

$$H = h - N \quad (4.12)$$

mit:

$H$  = orthometrische Höhe

$h$  = Ellipsoidhöhe

$N$  = Geoidhöhe (bezogen auf die Ellipsoidnormale)

Zusätzlich zur allgemeinen Datenaufbereitung, die hier beschrieben wurde, wird für jeden Datensatz eine „Kachelliste“ erstellt. Jeder Datensatz hat unterschiedlich große Kacheln mit teils unterschiedlicher Anzahl an Pixeln und im Fall von TANDEM sogar unterschiedlicher Auflösung (vgl. Tab. 8). In dieser Kachelliste sind für jede in dem jeweiligen Datensatz enthaltene Kachel der Name, die Anzahl der Pixel in Longitude und Latitude sowie die geographischen Koordinaten des Mittelpunkts des linken unteren Punktes dieser Kachel abgespeichert. Dadurch kann das im Kapitel 4.1.2 beschriebene Programm, das mittels des PiP-Algorithmus die für das ausgewählte Modellgebiet benötigten Kacheln für jeden Datensatz selbstständig auswählt, ausgeführt werden, ohne dass der Benutzer zusätzliche Informationen eingeben muss. Zuvor muss die Liste mit den Angaben für einen neuen Topographie-Datensatz entsprechend nur ein einziges Mal erstellt werden.

### 4.3.2 Neue Boden-Datensätze

Eine weitere Verbesserung der Landoberflächenparameter erfolgt durch neue Bodendaten. Bisherige Bodenkarten wie der globale Bodenklassendatensatz von Wilson & Henderson-Sellers (1985) (1° Auflösung; verwendet im Biosphere-Atmosphere Transfer Scheme (BATS) (Dickinson et al., 1986) und im Joint UK Land Environment Simulator (JULES) (Cox et al., 1999)) oder die von Zobler (1986) überarbeitete FAO Soil Map of the World (mit 0.5° aktuell in REMO und mit 1° im Simple Biosphere Model (SiB) eingebaut) (vgl. Dai et al., 2013) werden wegen ihrer globalen Abdeckung, der damit einhergehenden Konsistenz der Daten und wegen ihrer Einfachheit oft in LSM benutzt. Aufgrund dieser genannten Eigenschaften sind die Daten jedoch nicht für den Einsatz in hochaufgelösten regionalen Klimamodellen geeignet. Generell ist die flächendeckende Erforschung der Bodenarten und -beschaffenheit schwierig. Anhand von Bodenstichproben, deren Eigenschaften und Zusammensetzungen im Labor analysiert werden, und mittels verschiedener Satellitendaten wie Topographie oder Landbedeckung werden die Punktdaten in die Fläche interpoliert. Zudem müssen die punktuell für unterschiedliche Tiefen erstellten Bodenprofile meist erst standardisiert und in die Tiefe interpoliert werden, um nicht nur die Bodenbeschaffenheit an der Oberfläche, sondern auch im Untergrund zu erfassen. Die zunehmenden Forschungen und Modellierungen in Themenfeldern wie Erosion durch Wind und Wasser, Grundwasserverfügbarkeit, Vegetation und Landwirtschaft, Bodenfeuchtigkeit sowie allgemein hydrologische Vorgänge benötigen jedoch diese hochaufgelösten und auch global verfügbaren Produkte (vgl. Ballabio et al., 2016). Eine Übersicht der existierenden Bodendatenbanken sowie Informationen zum Projekt „GlobalSoilMap“, das eine harmonisierte globale hochaufgelöste digitale Bodendatenbank zum Ziel hat, bieten Omuto et al. (2013) und Arrouays et al. (2017). Auch Nemes (2011) listet verschiedene internationale Datenbanken mit ihren jeweiligen Stärken und Schwächen sowie daraus ableitbaren physikalischen und hydrologischen Parametern auf. Einen Vergleich von ausgewählten globalen Bodendatenbanken geben Dai, Shangguan, Wei et al.



(2019). Sie halten beispielsweise fest, dass der SoilGrids-Datensatz (Hengl et al., 2014) und der Global Soil Dataset for Earth System Models (GSDE) (Shangguan et al., 2014) in ihrer Analyse am besten abgeschnitten haben.

Der bisherige Bodendatensatz, der in REMO verwendet wird, ist die von Zobler (1986) überarbeitete Bodenkarte der Food and Agricultural Organization of the United Nations (FAO). Sie liefert eine  $0.5^\circ$  aufgelöste Weltkarte mit 15 Texturklassen, welche im 6. Schritt der Bodenbibliothek (vgl. Beschreibung in Anhang D) auf fünf Texturklassen und Torf reklassifiziert wird (vgl. Tab. 5 in Kapitel 3.4.2). Zum Vergleich mit diesem bisherigen FAO-Datensatz wird die neuere Version vom FAO (FAO, 1988), die eine höhere räumliche Auflösung von  $0.05^\circ$  hat, in REMO eingebaut. Danach werden in dieser Arbeit der SoilGrids-Datensatz (Soilgrid) und die Harmonized World Soil Database (HWSD) als Alternativen für den bisherigen Bodendatensatz ausgewählt. Die allgemeinen Informationen dieser vier Datensätze sind in Tabelle 9 zusammengefasst. Die Einschränkung der Abdeckung des Soilgrid auf  $90^\circ\text{N}$  und  $56^\circ\text{S}$  ist nur eine technische Bemerkung, denn südlich von  $56^\circ\text{S}$  sind keine Landmassen vorhanden, sodass auch Soilgrid als globaler Bodendatensatz gilt.

Bei dem neuen FAO-Datensatz (FAOn) werden die  $0.05^\circ$  aufgelösten Daten für Sand- und Tonanteil mit den globalen Bodendaten des International Geosphere-Biosphere Programme (IGBP) der Global Soil Data Task Group (2000) in der Version von IGBP (2000) ergänzt, die ebenfalls auf FAO-Bodendaten (FAO, 1995) basieren und  $0.05^\circ$  Auflösung besitzen, jedoch weitere Variablen enthalten und auch Daten in verschiedenen Tiefen zur Verfügung stellen. Die Sand- und Tonanteil-Daten des FAOn für die Oberfläche werden beispielsweise auch im ECOCLIMAP-Datensatz (globale Datenbank von Landoberflächenparametern, Masson et al., 2003 und Champeaux et al., 2005) verarbeitet, den z. B. das LSM SURFEX (le Moigne, 2009) verwendet.

Die SoilGrids-Daten werden anhand von Daten aus ca. 110.000 Bodenprofilen verschiedener nationaler und internationaler Datenbanken (z. B. auch HWSD) mittels verschiedener Regressions- und Kriging-Techniken erstellt. Dadurch erreicht Soilgrid eine horizontale Auflösung von 30 Bogensekunden (ca. 1 km) und stellt Daten für sieben verschiedenen Tiefen von 0–2 m zur Verfügung (Hengl et al., 2014). Eine neue Version des SoilGrids-Datensatz, die sogar eine horizontale Auflösung von 250 m besitzt, wurde 2017 veröffentlicht (Hengl et al., 2017). Jedoch wurde aufgrund der hohen Speicherbelegung von mehreren globalen Karten (4 Variablen und 7 Schichten) mit 250 m Auflösung und der erschwerten Verarbeitung und Aufbereitung durch die Limitierung des Arbeitsspeichers auf den Einbau dieser Version verzichtet. Es soll zunächst überprüft werden, ob und wie sich der Einbau der 1 km Version auf das Modellergebnis auswirkt.

Der HWSD, der als Updateprodukt der FAO-UNESCO Soil Map of the World gilt, wird ebenfalls aus verschiedenen Datensätzen zusammengesetzt. Er enthält hauptsächlich die Daten der Europäischen Bodendatenbank (ESDB), der Soil Map of China (Shi et al., 2004), der Soil and Terrain (SOTER) Datenbank (van Engelen & Dijkshoorn, 2012) und der World Inventory of Soil Emission Potential (WISE) Datenbank (Batjes et al., 1997) sowie Informationen aus den vorgestellten FAO-Produkten, wo die anderen Quellen nicht ausreichend Daten zur Verfügung

#### 4 Aufbereitung und Vergleich der neuen Oberflächendatensätze und Validationsdaten

stellen. Der HWSD besitzt wie der Soilgrid eine globale Abdeckung und eine Auflösung von 30 Bogensekunden. Es werden 16 verschiedene physikalische und chemische Bodeneigenschaften bereitgestellt (Nachtergaele et al., 2009). Die vorliegende Version HWSD v1.2 wurde 2012 veröffentlicht (Nachtergaele et al., 2012).

Tab. 9: Allgemeine Informationen zu den verwendeten Boden-Datensätzen

Name	FAO	FAOn	Soilgrid	HWSD
Datensatz	FAO-Zobler	FAO + Global Soil Data	SoilGrids1km	Harmonized World Soil Database v1.2
Abdeckung	90° N/S	90° N/S	90°N/ 56°S	90° N/S
Auflösung in °	0.5	$\frac{1}{12}$	$\frac{1}{120}$	$\frac{1}{120}$
Anzahl an Schichten	1	1	7	2
verwendete Parameter	Bodentextur	Bodentextur, Sand-/ Ton-/ Schluffanteil, organischer Anteil	Bodentextur, Sand-/ Ton-/ Schluffanteil, organischer Anteil	Bodentextur, Sand-/ Ton-/ Schluffanteil, organischer Anteil
Quelle	FAO-UNESCO (1971-1981)	FAO (1988), IGBP (2000)	Hengl et al. (2014)	FAO (2012)

Für alle neuen Datensätze werden neben der bisher einzigen Bodenvariable Textur weitere Variablen für die Verwendung in REMO hinzugezogen, denn es liegen je nach Datensatz unterschiedliche Texturklassifikationen vor. Neue Klassifikationen können jedoch mittels der verfügbaren Sand-, Ton- und Schluffanteile berechnet werden. Zudem ist in den Texturklassifikationen der neuen Daten keine separate Ausweisung von Torf enthalten. Weil die „Texturklasse“ Torf jedoch sehr spezifische Bodeneigenschaften besitzt, die sich stark von den anderen Texturklassen unterscheiden, und weil laut Sato et al. (2015) das Miteinbeziehen von organischem Material bei der Modellierung von Bodenprozessen äußerst wichtig ist, werden auch die jeweiligen Datensätze für organisches Material (bei FAOn: Organic Carbon, bei Soilgrid: soil organic carbon content/ORCDRC und bei HWSD: organic carbon/OC) heruntergeladen und aufbereitet.

Dazu werden zunächst alle neuen Bodendatensätze in netCDF Format konvertiert, um den automatisierten Ablauf in der Bodenbibliothek aufrecht zu erhalten. Anschließend werden die jeweiligen Einheiten der Datensätze angeglichen. Im FAOn-, SoilGrids- und HWSD-Datensatz liegen Sand-, Ton- und Schluffanteil bereits jeweils in Gewichtsprozent vor. Für diese Variablen sind keine Anpassungen notwendig. Der organische Anteil wird jedoch beim Soilgrid in Promille bzw.  $\frac{g}{kg}$  und beim FAOn und HWSD in Gewichtsprozent angegeben, sodass die Werte von Soilgrid durch 1000 und die von FAOn und HWSD durch 100 dividiert werden müssen, um die gewünschte Einheit zu erhalten. Außerdem wird für die Pedotransferfunktion (PTF) der Soil Organic Matter (SOM)-Wert und nicht der Soil Organic Content (SOC)-Wert benötigt. Um den SOM-Wert aus einem SOC-Wert zu erhalten, werden die organischen Werte aller Datensätze mit dem Faktor 1.72 multipliziert. Dies ist ein gängiges, wenngleich aber auch umstrittenes Verfahren

(vgl. van Looy et al., 2017). Diese Umwandlung wird z. B. auch für die Aufbereitung des in dieser Arbeit verwendeten Bodenvalidierungsdatensatzes High Resolution Soil Map of Hydraulic Properties (HiHydroSoil) verwendet (de Boer, 2016).

In Tabelle 9 ist außerdem zu erkennen, dass die Variablen für unterschiedliche Anzahlen von Schichten vorliegen. Während die FAO-Datensätze die Parameter nur als Daten für die Oberfläche (Toplayer) zur Verfügung stellen, gibt es beim HWSO Daten für zwei Schichten (0–30 cm: Toplayer, 30–200 cm: Sublayer) und bei Soilgrid für sieben Schichten (0 cm, 5 cm, 15 cm, 30 cm, 45 cm, 100 cm, 200 cm). Für die realistischere Darstellung des Bodensystems und die neuen Parametrisierungen wurden alle Schichtdaten heruntergeladen. Anschließend wurden die für verschiedene Bodenschichten vorliegenden Originaldaten mit den in Kapitel 4.1.1 beschriebenen Methoden auf die fünf REMO-Schichttiefen umgewandelt. Bei den neuen FAO-Daten kann mit Hilfe des mitgelieferten Programms zusätzlich zur Auswahl der Variablen die Ober- und Untergrenze der Bodendaten festgelegt werden (Daten bis ca. 1.5 m), sodass die resultierende Karte den Mittelwert über die vom Benutzer festgelegte Tiefe ausweist. Die Skripte in der neuen BodLibKit wurden so umgeschrieben, dass jeweils fünf Schichtkarten statt einer Oberflächenkarte und vier statt einer Bodenvariable für den jeweils ausgewählten Datensatz herausgeschrieben werden. Wählt ein Benutzer weiterhin die bisherigen FAO-Daten für seine Bodenbibliothek, werden alle Schichtwerte auf den verfügbaren Wert der ersten Schicht gesetzt. Das gewährleistet, dass alle Benutzer unabhängig von der jeweiligen Wahl der Eingangsdaten den gleichen Präprozessierungscode benutzen können.

Die drei aus den neuen Bodendaten abgeleiteten Variablen SAND, CLAY und FORG für die verschiedenen Bodenschichten werden trotzdem unabhängig von der Einstellung des Benutzers als neue Variablen zusätzlich zu den bisherigen Variablen (vgl. Tab. 6) in die finale Bodenbibliotheksdatei geschrieben. Außerdem wurden weitere Variablen von anderen Bodendatensätzen und von den neu eingeführten Datensätzen abgeleitete Variablen in die neue Bodenbibliothek eingebaut. Diese Variablen werden für die anfangs erwähnte parallel laufende Dissertation zur Weiterentwicklung des 5Layer-Schemas benötigt und sind in der Tabelle in Anhang F aufgelistet. Für die Analysen dieser Arbeit werden alle Variablen außer SAND, CLAY und FORG nicht benutzt.

## 4.4 Vergleich der bisherigen und neuen Datensätze

In diesem Kapitel werden die vorgestellten Topographie- und Boden-Datensätze direkt miteinander verglichen. Die Analyse wird zum einen für alle CORDEX-Regionen für eine Auflösung von  $0.44^\circ$  durchgeführt, um eventuelle Schwächen und Probleme eines Datensatzes in gewissen Regionen aufzudecken und zu thematisieren. Zum anderen wird für die topographischen Datensätze eine Analyse auf der kleinsten gemeinsamen Originalauflösung von 1 km für das definierte Deutschlandgebiet (GER-11) vorgenommen. Zudem findet eine Zusammenfassung der Einschätzungen und dem Ranking der Datensätze aus der Literatur statt. Die Analyse der Datensätze, die tatsächlich für die Modellläufe der beiden Untersuchungsgebiete (EUR-44 und

GER-11) in dieser Arbeit verwendet wurden, wird im Detail in Kapitel 6.3.1 präsentiert.

#### **4.4.1 Topographie-Daten**

Zwar werden die Topographie-Daten (wie in Kapitel 3.4.2 beschrieben) auf die gewünschte bzw. vorgegebene Klimamodellauflösung interpoliert und dadurch das ganze Potential der bisher 1 km aufgelösten Daten im Moment noch nicht gänzlich genutzt, aber es ist dennoch sinnvoll einen Vergleich mit alternativen und noch besser aufgelösten Datensätzen zu ziehen. Neben den mittleren Höhen pro Pixel werden zudem auch die Varianz der Orographie (VAROR)  $s^2$  (vgl. Formel D.3) und die Rauigkeitslänge der Orographie (AZ0)  $z_{0,oro}$  im Modell verwendet (vgl. Formel D.4), welche beide stark von der Subgrid-Skala abhängen.

#### **Vergleich der vier neuen Topographiedatensätze mit dem bisherigen Datensatz für GER-1 km**

In Abbildung 15 sind die vier neuen Topographie-Datensätze (ASTER, ALOS, SRTM und TANDEM) und der bisherige Datensatz GTOPO in dessen Originalauflösung von  $0.0083^\circ$ , was ca. 1 km entspricht, für das definierte GER-Untersuchungsgebiet dargestellt. Auf den ersten Blick fällt auf, dass der ALOS-Datensatz in allen Teilabbildungen mehrere unnatürliche Streifen besonders im östlichen Bereich des ausgewählten Gebiets hat. An diesen Stellen existieren in den Originaldaten keine Werte (vgl. Kapitel 4.3.1), sodass in der Karte für den jeweiligen Pixel der Startwert Null ausgegeben wird. Das heißt innerhalb dieser  $1 \text{ km} \times 1 \text{ km}$  Pixel findet der Algorithmus keinen Wert in allen darin befindlichen  $30 \text{ m} \times 30 \text{ m}$  Pixeln von ALOS. Diese vermehrten Nullwerte verschieben auch die Quantile von ALOS in Abbildung 16A im Vergleich zu allen anderen Datensätzen nach unten. Zudem sieht man auch die größten Unterschiede in den Häufigkeitsverteilungen zwischen ALOS und den anderen Topographie-Daten in diesem Wertebereich (vgl. Abb. 16B). Aufgrund dieser im Untersuchungsgebiet auftretenden großflächigen Fehlwerte wird der ALOS-Datensatz in dieser Arbeit nicht weiter benutzt, obwohl beispielsweise Caglar et al. (2018) diesem Topographiedatensatz einen sehr geringen Root Mean Square Error (RMSE) von nur 1.78 m bescheinigen.

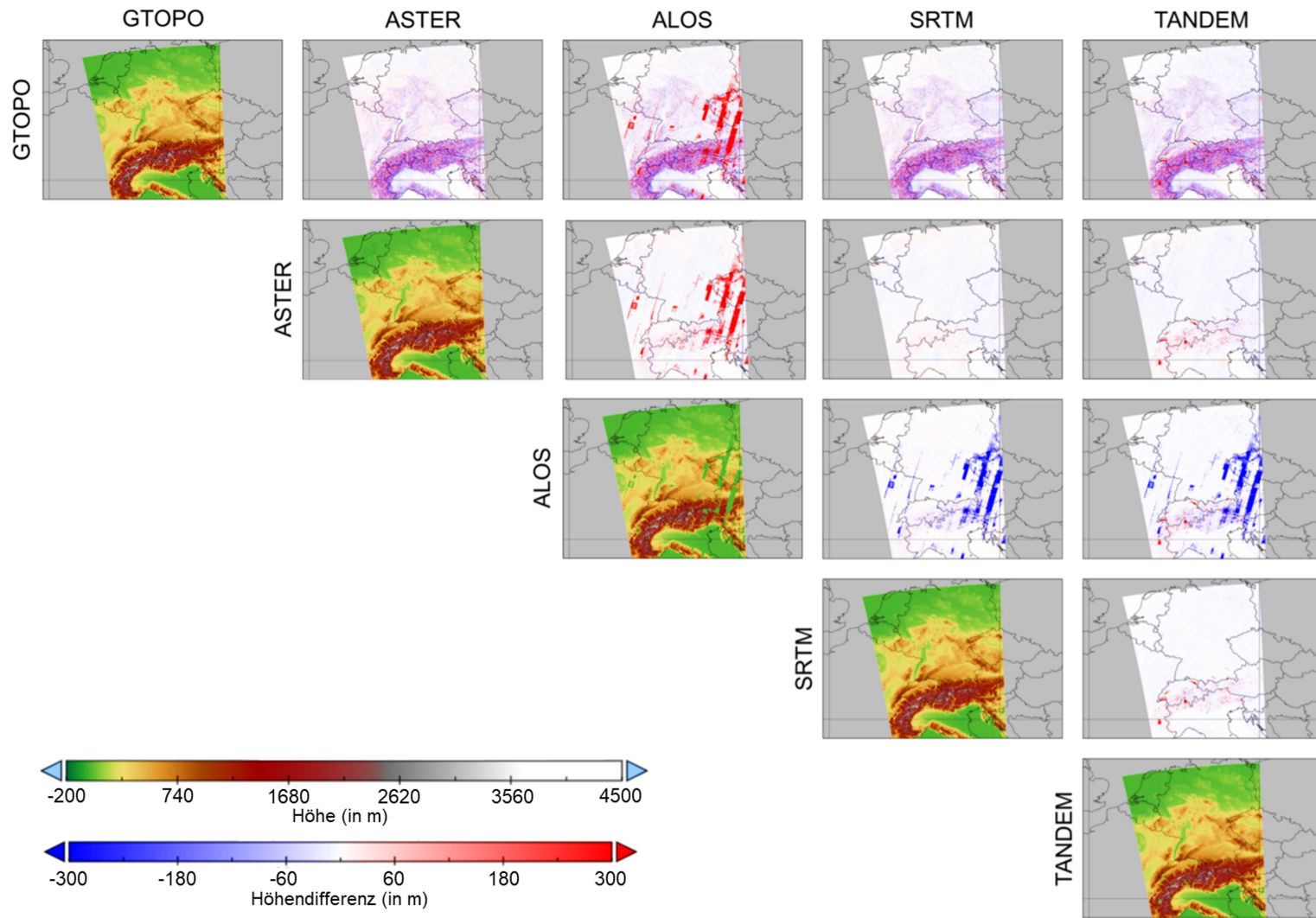
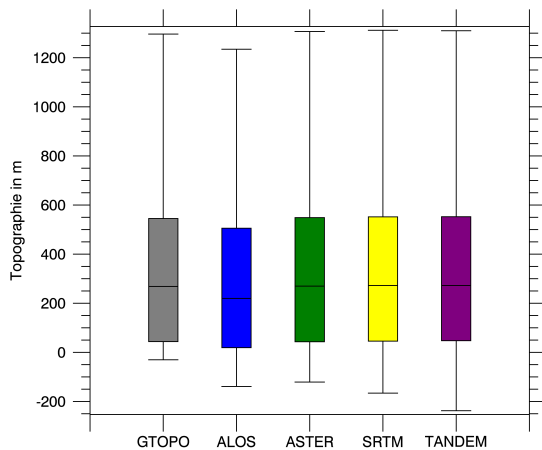


Abb. 15: Geographischer Vergleich der Topographie-Datensätze in GER-1 km

Tab. 10: Statistische Kennwerte der fünf verschiedenen Topographie-Datensätze für GER-1 km (Abb. 15) im Vergleich (Alle Höhenangaben sind in der Einheit m.)

	Minimum	Maximum	Mittelwert	Standardabweichung
GTOPO	-30.0	4469.0	436.01	560.1
ASTER	-121.05	4516.7	439.23	563.48
ALOS	-138.5	4450.0	397.78	548.15
SRTM	-166.02	4509.0	440.75	562.68
TANDEM	-191.51	4377.8	480.53	565.42

(A)



(B)

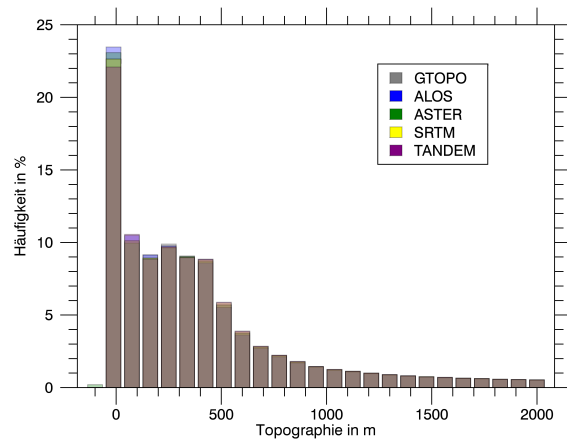


Abb. 16: Statistischer Vergleich der Topographie-Datensätze in GER-1 km

(A) Boxplot

(B) Häufigkeitsverteilung

Aus Abbildung 15 ist weiterhin abzulesen, dass die drei verbleibenden Höhendatensätze (ASTER, SRTM, TANDEM) ähnliche Differenzkarten im Vergleich zu GTOPO aufweisen. Dies zeigt sich auch daran, dass die jeweiligen Differenzkarten zwischen ASTER-SRTM, ASTER-TANDEM und SRTM-TANDEM nur geringe Differenzwerte von  $\pm 50$  m haben. Die größten Differenzen zwischen bisherigem und den neuen Daten liegen erwartungsgemäß in den Alpen und in den Mittelgebirgen. In gebirgigen Regionen, in denen sich die topographische Höhe horizontal gesehen schnell ändert, spielt die Auflösung eine größere Rolle als in flachen und sich über weite Strecken hinweg kaum ändernden Landschaften. Es ist zu erkennen, dass kein systematischer Differenzwert erreicht wird, sondern dass sich positive und negative Werte abwechseln. Die Beträge der Differenzen liegen in diesem Bereich meist über 100 m.

Die jeweiligen Quantile der drei neuen Datensätze sind nahezu identisch (vgl. Abb. 16). Dennoch können Unterschiede beim unteren Whisker festgestellt werden, die sich auch im Histogramm in den unterem Wertebereich erkennen lassen. Der sehr niedrige Wert beim unteren TANDEM-Whisker (Abb. 16A) ist durch die Geoid-Ellipsoid-Korrektur zu erklären. Da durch diese Korrektur alle Originalwerte im Bereich von Deutschland um ca. 70 m reduziert werden, sind die größeren Wasserflächen nicht mehr Null, wie in den anderen Datensätzen. Eigentlich wäre daher eine

Wassermasken-Korrektur nötig. Da jedoch der TANDEM-Datensatz aufgrund der in Kapitel 4.4.1 erläuterten Probleme für die weiteren Analysen ausgeschlossen wird, ist dies für die vorliegende Arbeit nicht erforderlich. Generell ist SRTM fast im gesamten Gebiet ein paar Meter niedriger als TANDEM. ASTER ist an der Nordseeküste und im alpinen Bereich etwas höher und sonst etwas tiefer als SRTM und TANDEM (vgl. Abb. 15), was vermutlich an der höheren Originalauflösung von ASTER liegt.

Die in Tabelle 10 dargestellten statistischen Kennwerte zeigen, dass GTOPO zwar den höchsten Minimumwert, aber den geringsten Mittelwert und die geringste Standardabweichung (abgesehen von ALOS mit vielen fehlerhaften Nullwerten) besitzt. Der kleinste Maximalwert wird in TANDEM erreicht, welcher gleichzeitig auch den kleinsten Minimalwert, aber den höchsten Mittelwert und die größte Standardabweichung hat. ASTER und SRTM liegen im Mittelfeld und weisen sehr ähnliche Kennziffern im Untersuchungsgebiet auf.

### **Topographiedatenanalyse und -vergleich in der Literatur**

In der Literatur spielt der Vergleich von verschiedenen Topographiedatensätzen eine große Rolle, da diese Daten für viele unterschiedliche Anwendungsgebiete wie z. B. Reliefanalysen, Risikobewertungen, gravimetrische Interpolationen oder als Eingangsdaten von hydrologischen und klimatologischen Modellen benötigt werden. Grohmann (2018) vergleicht beispielsweise die (nahezu) globalen aktuellen Topographiedaten SRTM, ASTER, ALOS und TANDEM-X, die alle ebenfalls in dieser Arbeit verwendet werden, für Brasilien miteinander und merkt an, dass die beiden aus optischen Systemen generierten Datensätze ASTER und ALOS nur in offenem Terrain ähnlich gute Ergebnisse wie TANDEM-X und SRTM erzielen. Jedoch weißt er auch darauf hin, dass er zwar die SRTM-Version mit 30 m Auflösung verwendet hat, die tatsächliche Auflösung des Datensatzes jedoch geringer scheint. Zudem ist zu beachten, dass in dieser Studie die 12 m-Version von TANDEM-X als Referenzdatensatz verwendet wird, sodass das gute Abschneiden der daraus abgeleiteten TANDEM-X 30 m-Version nicht überrascht.

Der Vergleich und die Einschätzung der Güte der Datensätze hängt neben der jeweiligen Version der Daten und den damit verbundenen Auflösungen, auch von deren Aufbereitung und den Untersuchungsgebieten ab, was an folgenden Beispielen gut gezeigt werden kann. Satgé et al. (2015) stellen in ihrer Studie fest, dass der ASTER GDEM V2 in Südamerika genauer ist als SRTM V4 (ca. 8.8 m), wohingegen P. Li et al. (2012) und Jing et al. (2013) für China angeben, dass eine höhere Genauigkeit beim SRTM V4.1 im Gegensatz zu ASTER GDEM V2 vorliegt. Auch in Australien (Rexer & Hirt, 2014) und in der Türkei (Bildirici & Abbak, 2017) schnitt generell jeweils der SRTM3-Datensatz im Vergleich zu ASTER GDEM V2 etwas besser ab. Bildirici & Abbak (2017) kommen jedoch zum Ergebnis, dass vor allem in gebirgigen Regionen der ASTER-Datensatz dem SRTM überlegen ist. Diesen Eindruck bestätigen Enßle et al. (2012) für Süddeutschland und Morais et al. (2017) für Brasilien. Trotzdem stellt Sertel (2010) für die Türkei fest, dass schon ASTER GDEM V1, welcher laut Tachikawa, Hato et al. (2011) und Abrams et al. (2020) schlechter performt als V2, ein geeigneter Datensatz zur Darstellung der türkischen Topographie ist. In einer Studie von Santillan & Makinano-Santillan (2016) erreichte

der ALOS-Datensatz auf den Philippinen die höchste Genauigkeit verglichen mit SRTM V3 (30 m) und ASTER GDEM V2. Zum gleichen Ergebnis kommen Courty et al. (2017) bei einer Flutmodellierung eines Flusseinzugsgebiets in Mexiko. Teile dieser und weiterer Studien zum Vergleich der verschiedenen Versionen von SRTM- und ASTER-Datensätzen sind bei Sahid et al. (2018) aufgelistet.

Des Weiteren wird in den oben genannten Studien festgehalten, dass die Genauigkeit von Topographieprodukten stark von den jeweiligen Oberflächenbedeckungen und der grundlegenden Beschaffenheit der Oberfläche abhängt. So nimmt die Genauigkeit beispielsweise für steile Hänge ab, weil unter anderem Bereiche während der Aufnahme im Radarschatten gelegen waren. Zudem werden nicht-bewaldete Gebiete besser abgebildet. Wasserkörper können ebenfalls aufgrund der sehr niedrigen Radarrückstrahlung von Wasser ungenauer sein oder werden wie z. B. bei SRTM gar nicht erfasst.

Die Auswahl des richtigen Datensatzes hängt auch von der Art der Nutzung ab. Forkuor & Maathuis (2012) empfehlen beispielsweise SRTM V4 für hydrologische Anwendungen, da dieser im Vergleich zu ASTER GDEM V1 die Höhe eher über- als unterschätzt. Suwandana et al. (2012) merken dagegen an, dass SRTM V4 zwar die Topographie genauer abbildet, ASTER GDEM V2 hingegen für die Analyse von Überschennungsgebieten herangezogen werden sollte. Zusätzlich ist die richtige geographische Verortung der jeweiligen Daten wichtig. Nikolakopoulos et al. (2007) konnten beispielsweise beim Vergleich von SRTM und ASTER für Kreta einen leichten Versatz der beiden Datensätze von 200–400 m feststellen. Einen kleineren Versatz bemängeln auch Hirt et al. (2010) für SRTM V4.1 beim Vergleich mit anderen Topographie-Datensätzen in Australien.

Generell lässt sich durch diese Studien erkennen, dass es keinen Datensatz gibt, der den anderen weltweit in Sachen Genauigkeit deutlich überlegen ist. Es ist sinnvoll, mehrere Datensätze zur Verfügung zu haben, um die jeweiligen Schwächen eines Datensatzes zu detektieren und je nach Untersuchungsgebiet einen geeigneten auszuwählen.

#### **Fehlwerte im TANDEM-X Datensatz**

In den neuen TANDEM-X-Daten, die im Jahr 2018 vom DLR zur Verfügung gestellt wurden, wurden beim Vergleich der Topographiedatensätze vor allem an den Küsten und in küstennahen Bereichen auffällige Werte entdeckt (z. B. eine nicht existente Insel mit Erhebungen bis zu 2200 m westlich vor Afrika, vgl. Abb. 17A). Eine Korrektur dieser Werte mittels der mitgelieferten Wassermaske konnte einige (vgl. links unten in Abb. 17A), aber nicht alle kritischen Werte eliminieren. Der Datensatz weist weiterhin einige Schwachstellen in diesen Bereichen auf, da die zur Korrektur eingesetzte Wassermaske aus den gleichen Ausgangsdaten prozessiert wird wie die zu korrigierende Höhenkarte. Abbildung 17B zeigt beispielsweise eines dieser trotz angewandeter Wassermaskenkorrektur übrig gebliebenen Relikte im Mittelmeer. Anhand der lückenhaften Struktur und der quadratischen Abgrenzung der Fehlwerte sind diese klar als Rauschen in einer Originalaufnahme erkennbar. Die Werte, die in dem gelb markierten Kreis erreicht werden, sind mit ca. 5000 m so hoch, dass diese sich selbst beim Aggregieren auf die 50 km Auflösung wegen



der Mittelwertbildung stark auswirken. Mit einem gewissen Mehraufwand könnte eine externe Wassermaske auf die TANDEM-Kacheln, die in unterschiedlichen Auflösungen für unterschiedliche Teilgebiete der Erde vorliegen, angepasst werden und die TANDEM-Daten dadurch vielleicht besser korrigiert werden. Da mit den ASTER-Daten allerdings eine gute Alternative vorliegt, wurde auf die Korrektur und das Benutzen von TANDEM in dieser Arbeit verzichtet.

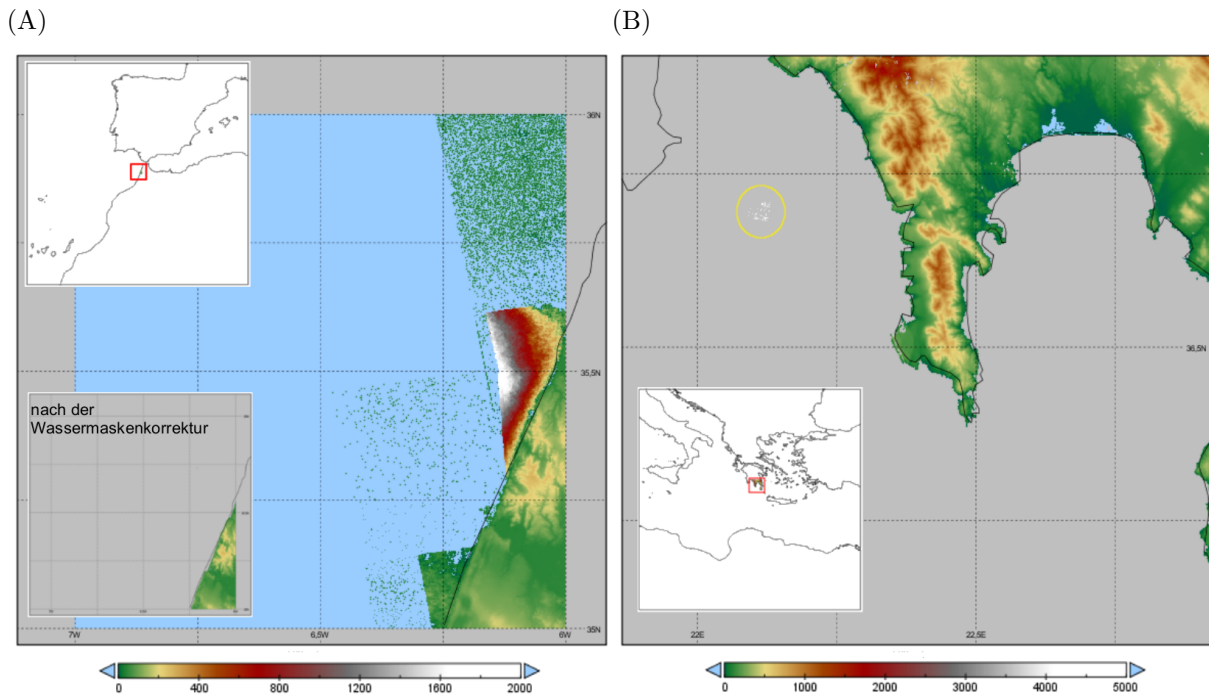
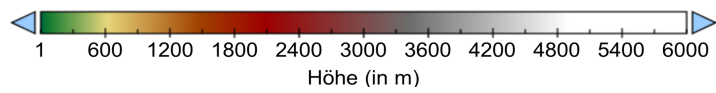
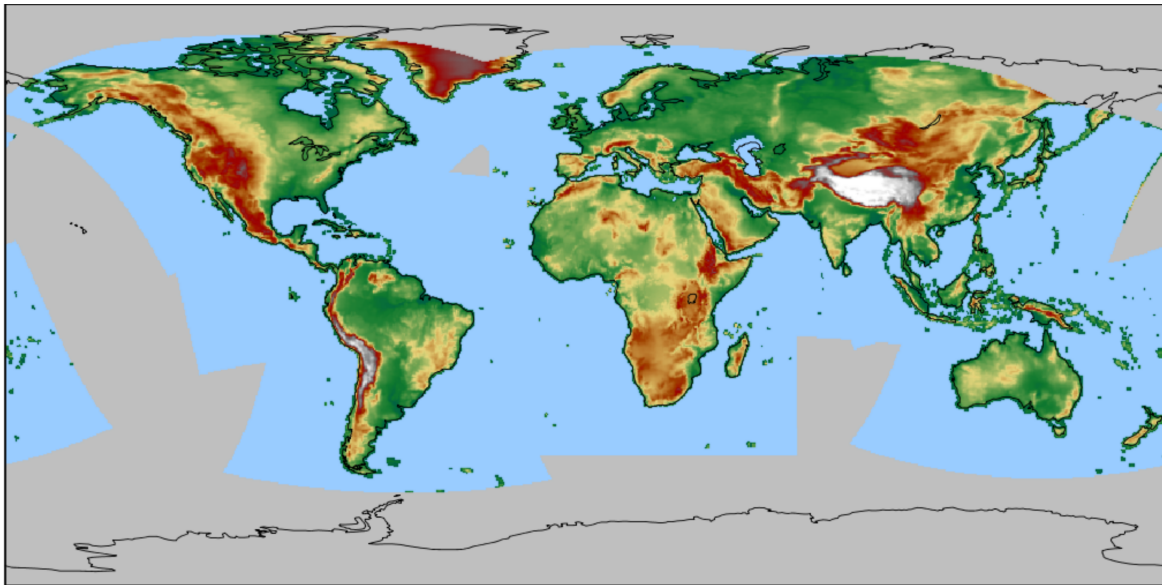


Abb. 17: Beispiele von Fehlwerten im TANDEM-X Datensatz (Höhenangaben in m)  
 (A) Atlantik: Fehlwerte sind erfolgreich korrigiert nach Wassermaskenkorrektur  
 (B) Mittelmeer: Fehlwerte bleiben bestehen trotz angewandter Wassermaskenkorrektur

### Vergleich der ASTER- und GTOPO-Daten für die CORDEX-Regionen auf 0.44° Auflösung

Nachfolgend wurde der ASTER-Datensatz, der sich in dieser Arbeit als Favorit aller getesteten Daten ausgezeichnet hat, für die neun CORDEX-Gebiete auf einer Auflösung von 0.44° erstellt und mit dem bisherigen GTOPO-Datensatz verglichen. Diese Analyse dient der globalen Überprüfung des neuen Topographie-Datensatzes.

(A)



(B)

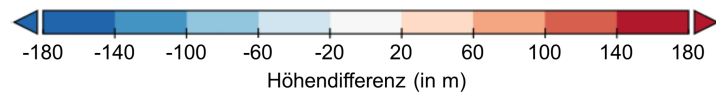
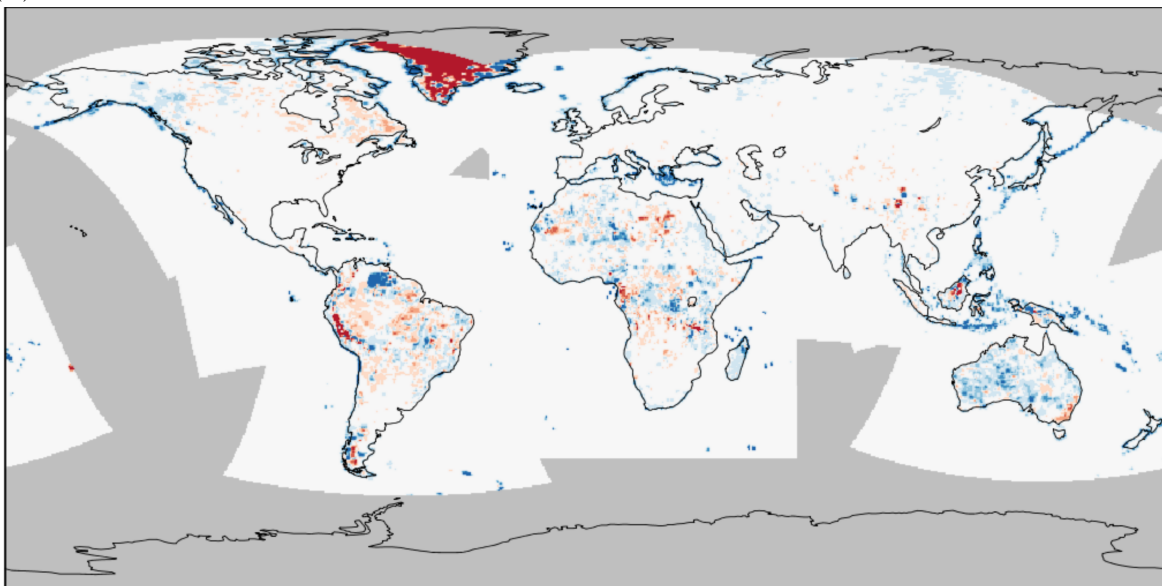


Abb. 18: Vergleich der topographischen Höhe (FIB) von GTOPO- und ASTER-Datensatz

(A) GTOPO

(B) ASTER-GTOPO

Bei dem in Abbildung 18 dargestellten Höhenvergleich fällt als erstes Grönland auf, das in den ASTER-Daten wesentlich höhere Werte erzeugt als die GTOPO-Daten. Auf dieses „Grönland-Problem“ wird in Kapitel 4.4.1 näher eingegangen. Weitere markante Stellen sind:

- eine größere Fläche in Venezuela in Südamerika (SAM), die im ASTER-Datensatz mehr als 200 m niedriger ist
- eine größere Fläche in Peru (SAM), die ASTER ca. 150 m höher darstellt
- einige flächige Regionen in Australien (AUS), die laut ASTER 60–140 m niedriger sind
- die in einem Halbbogen angeordnete und pixelweise höheren und niedrigeren Werte südlich des Himalajas im ostasiatischen Gebiet (EAS)
- sowohl positive als auch negative Differenzen von  $\pm 100$  m in Nord- und Zentralafrika (AFR)
- alle Inseln (besonders gut erkennbar in Indonesien) und Küstenlinien, die beim neuen Datensatz geringere Höhen erreichen als bisher

Im Hauptuntersuchungsgebiet Europa sind die Unterschiede auf dieser 50 km Auflösung nicht so stark ausgeprägt.

Die in Abbildung 19A dargestellten Boxplots zeigen, dass der neue ASTER-Datensatz in allen Kennwerten des Boxplots niedrigere Werte erreicht als GTOPO. Dass ASTER besonders an den Küsten niedrigere Werte hat als GTOPO, lässt sich auch in der niedrigsten Histogrammklasse ( $-62.5-0$  m) an dem deutlich höheren Anteil von über 20% (bei ASTER) im Vergleich zu 12% (bei GTOPO) ablesen. In den folgenden Klassen gleichen sich die beiden Datensätze im Bezug auf die kumulierte Häufigkeit wieder an, weil in diesem Wertebereich GTOPO höhere Häufigkeiten erreicht (vgl. Abb. 19B). Danach sind die Balken nahezu identisch, was sich auch in den kumulierten Häufigkeiten in Abbildung 19C, die ab ca. 1000 m übereinander liegen, widerspiegelt.

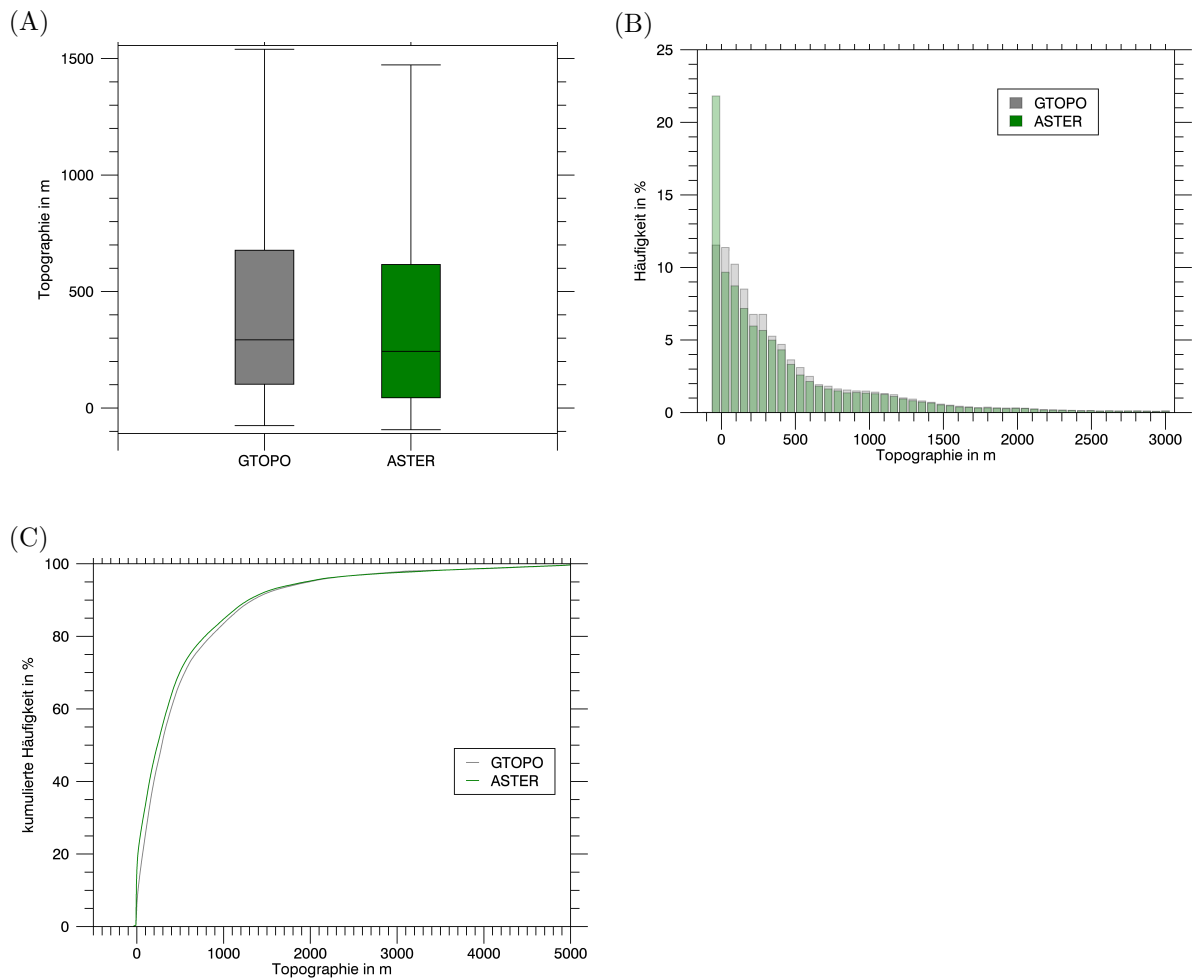


Abb. 19: Statistischer Vergleich von GTOPO und ASTER für FIB in 0.44° Auflösung

- (A) Boxplot
- (B) Häufigkeitsverteilung
- (C) kumulierte Häufigkeitsverteilung

Schaut man sich die Varianz der Höhe an (Abb. 20), unterscheiden sich die schon bei der Variable FIB aufgefallenen Gebiete der Erde hier ebenfalls. Aufgrund des hohen Wertebereichs, welcher aus dem Quadrieren in der Berechnungsformel (vgl. Formel 6.2) resultiert, kristallisieren sich diese Unterschiede noch stärker heraus.

Vergleicht man die beiden Datensätze anhand des Boxplots und der Häufigkeitsverteilung der Werte (Abb. 21), ist zu erkennen, dass wieder alle Kennziffern des Boxplots bei ASTER geringer ausfallen als bei GTOPO. ASTER besitzt eine wesentlich höhere Häufigkeit bei sehr geringen Varianzwerten, die, betrachtet man die dazugehörigen Karten in Abbildung 20B, besonders an den Küsten zu finden sind. In der kumulierten Häufigkeit (Abb. 21C) fällt auf, dass die beiden Kurven einen Schnittpunkt bei ca. 45 000 m<sup>2</sup> haben. Das heißt, im unteren Wertebereich weist ASTER mehr Pixel mit niedrigeren Varianzwerten aus, im mittleren Wertebereich der Varianz werden dagegen weniger Pixel gefunden und bei den höheren Varianzwerten wieder mehr Pixel als bei GTOPO.

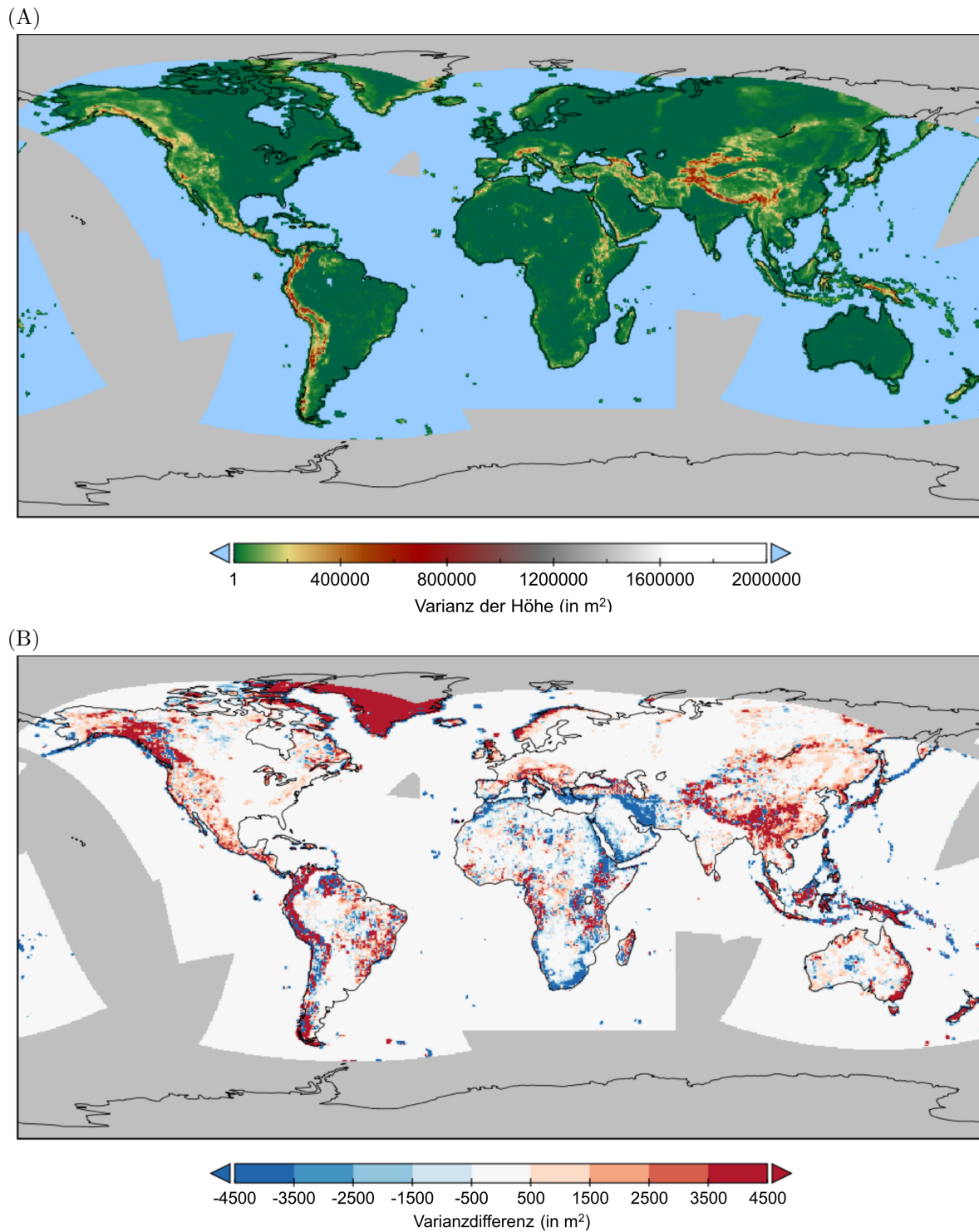


Abb. 20: Vergleich der Varianz der topographischen Höhe (VAROR) von GTOPO- und ASTER-Datensatz

(A) GTOPO

(B) ASTER-GTOPO

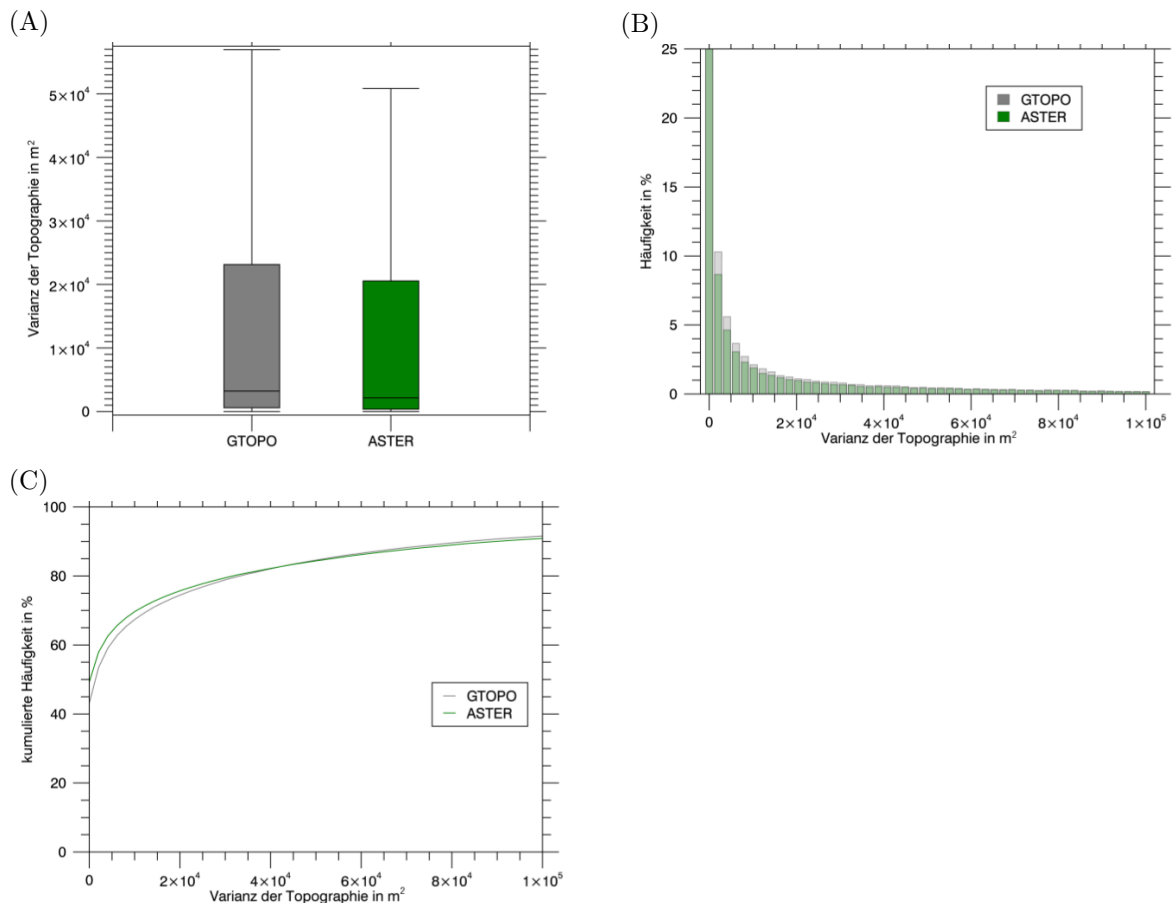
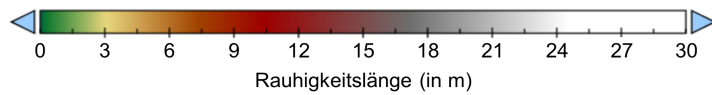
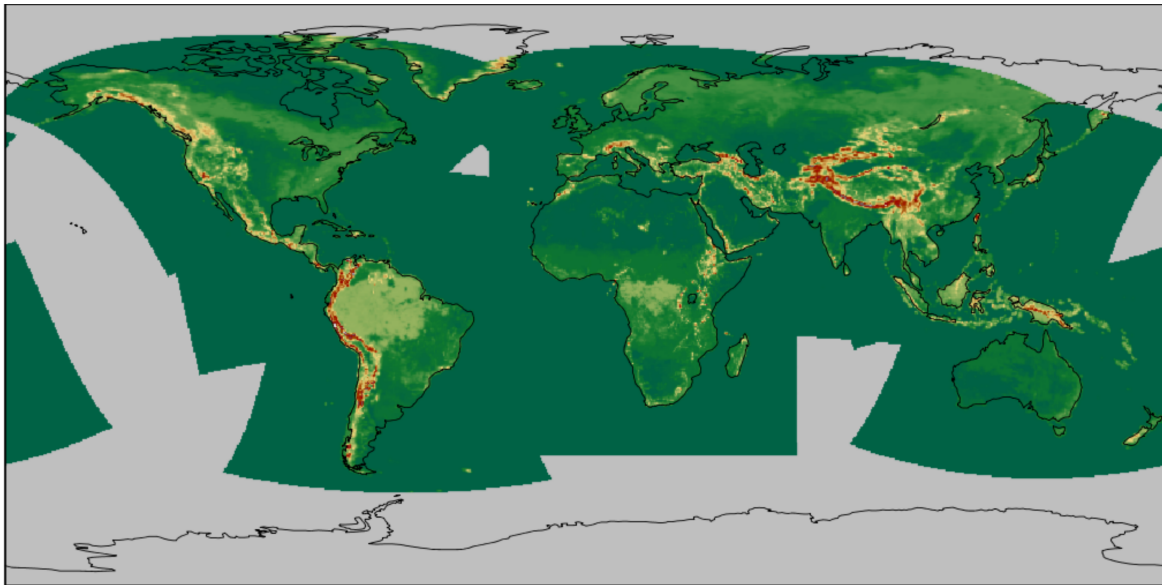


Abb. 21: Statistischer Vergleich von GTOPO und ASTER für VAROR in 0.44° Auflösung  
(A) Boxplot  
(B) Häufigkeitsverteilung  
(C) kumulierte Häufigkeitsverteilung

Bei der Rauigkeitslänge, die in Abbildung 22 dargestellt ist, werden die Unterschiede zwischen den Topographie-Datensätzen insgesamt kleiner. Vereinzelt sind jedoch sehr große Abweichungen festzustellen. Die Skala und Klassenbreite der Werte wurde dementsprechend gewählt, weshalb man so aber in der Differenzabbildung 22B in den meisten Gebieten kaum Unterschiede erkennt. Dennoch sind wie bei der Varianz beispielsweise vermehrt an der Westküste von Südamerika und an der Südseite des Himalaja sowohl positive als auch negative Differenzen zu erkennen. Die größten Differenzwerte mit über 1.75 m sind wieder in Grönland zu finden. Diese flächendeckenden höheren Werte erklären auch das höhere 3. Quartil und den höheren oberen Whiskerwert für ASTER in Abbildung 23A. Auch am Histogramm (Abb. 23B) und an der kumulierten Häufigkeitskurve (Abb. 23C) ist gut abzulesen, dass geringere Rauigkeitswerte öfter bei GTOPO gezählt werden, während im mittleren Bereich beide Topographiedatensätze relativ ähnliche Häufigkeiten erreichen und ASTER höhere Werte im Wertebereich oberhalb von 2.5 m hat.

(A)



(B)

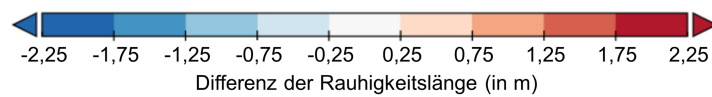
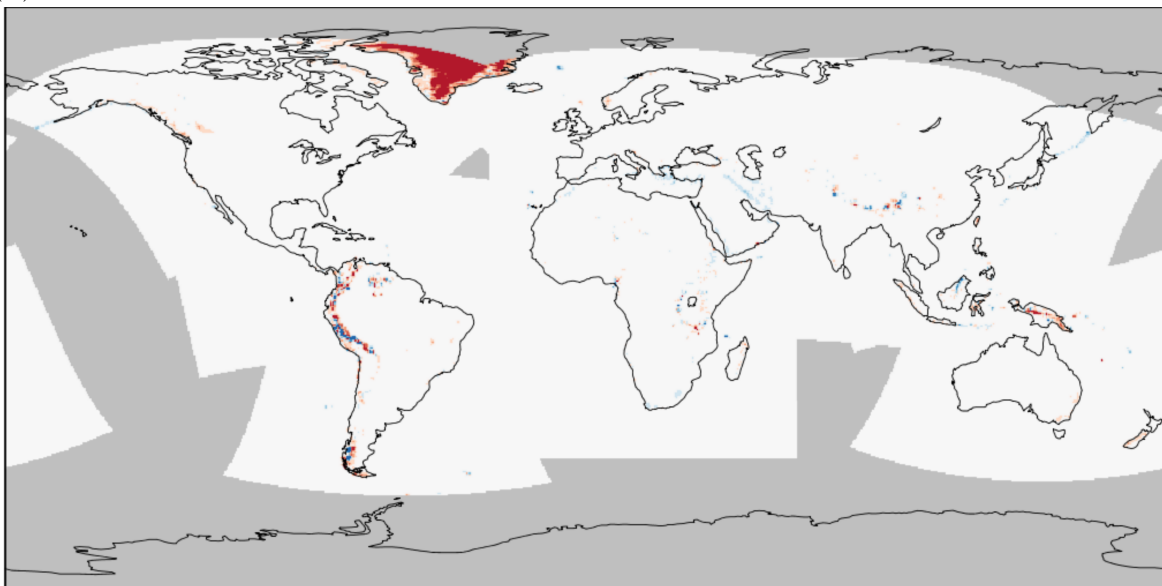


Abb. 22: Vergleich der Rauhigkeitslänge ( $AZ_0$ ) von GTOPO- und ASTER-Datensatz

- (A) GTOPO
- (B) ASTER-GTOPO

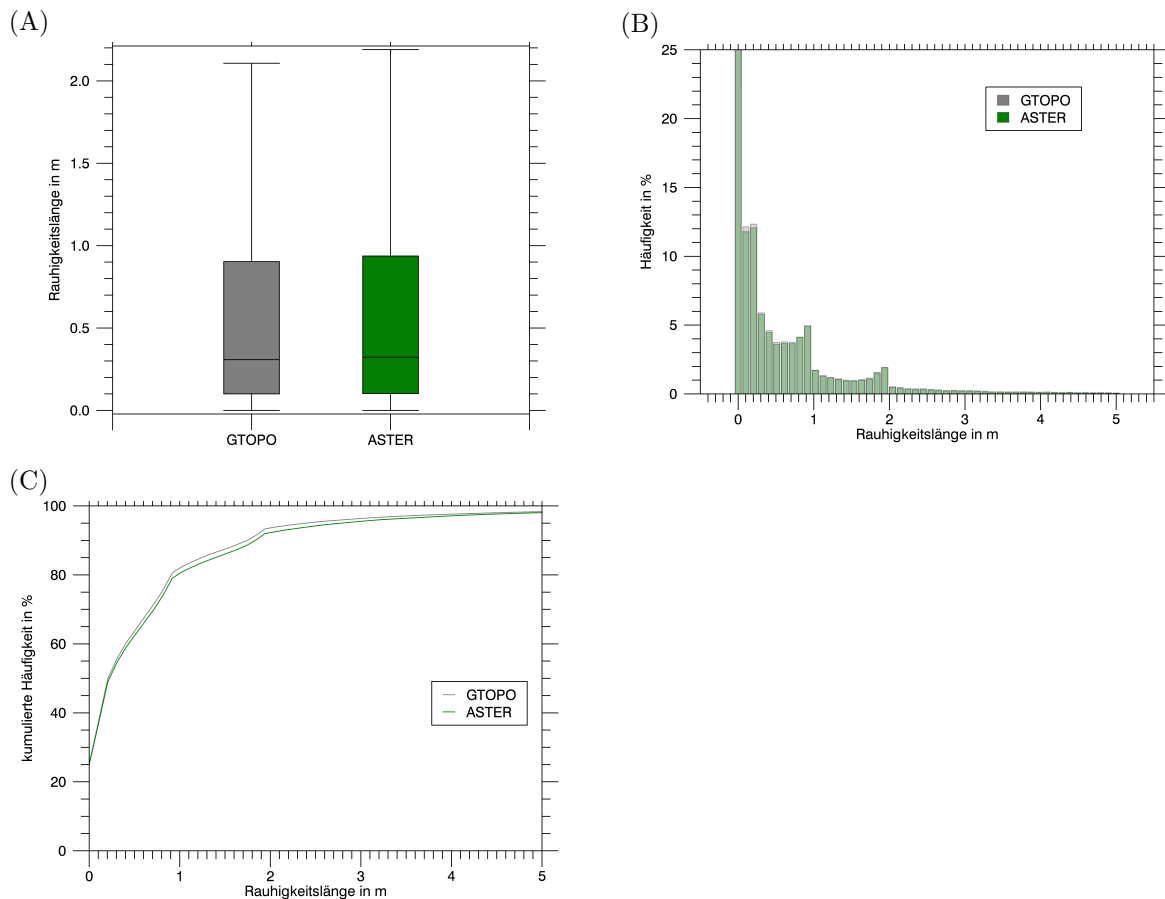


Abb. 23: Statistischer Vergleich von GTOPO und ASTER für AZ0 in 0.44° Auflösung  
(A) Boxplot  
(B) Häufigkeitsverteilung  
(C) kumulierte Häufigkeitsverteilung

#### Validation der Topographie von Grönland mit zusätzlichen Topographiedaten

Um genauere Informationen über das Ausmaß der Differenzen über Grönland zu bekommen und festzustellen, ob ASTER oder GTOPO Grönland schlecht abbildet, wurden zusätzlich die Daten vom Greenland Ice Mapping Project (GIMP) (Howat et al., 2014) und das ArcticDEM (Porter et al., 2018) heruntergeladen und aufbereitet. Abbildung 24 zeigt diese beiden Datensätze zusammen mit den ASTER-Daten und ihre jeweiligen Differenzen für einen Großteil Grönlands mit einer Auflösung von 90 m, da dies die größte Originalauflösung der drei Datensätze ist. In den Einzelabbildungen fällt auf, dass die ASTER-Höhendaten auf dem Eisschild der Insel viel höhere Werte erreichen, als die beiden Validationsdatensätze. Die Maximalwerte der jeweiligen Datensätze mit über 6500 m bei ASTER im Vergleich zu 3736 m (GIMP) bzw. 4478 m (ArcticDEM) zeigen, dass hier eine Schwäche des ASTER-Datensatzes vorliegt. Es ist bekannt, dass Rasteraufnahmen von großen Eisschilden aufgrund der kontrastarmen Oberfläche zu teilweise verfälschten Ergebnissen führen können. Zudem sind diese Gebiete oft in ihrer Befliegungshäufigkeit schlecht abgedeckt (Abrams et al., 2020), wodurch Artefakte in Form von Schatten oder Wolken mehr ins Gewicht fallen und die Daten verfälschen, wenn sie nicht korrekt ausmaskiert werden.



An den Küsten ähneln sich ASTER und GIMP mehr als ASTER und ArcticDEM, da der GIMP-Datensatz dort aus den ASTER-, den SPIRIT DEM- (Korona et al., 2009) und den Bamber DEM- (Bamber et al., 2001) Daten generiert wurde (vgl. Howat et al., 2014). Hvidegaard et al. (2012) bestätigen zudem das gute Abbilden der Topographie in den Küstenregionen von Grönland durch die ASTER-Daten mit aus Flugzeugen aufgenommenen Light Detection And Ranging (LiDAR)-Daten.

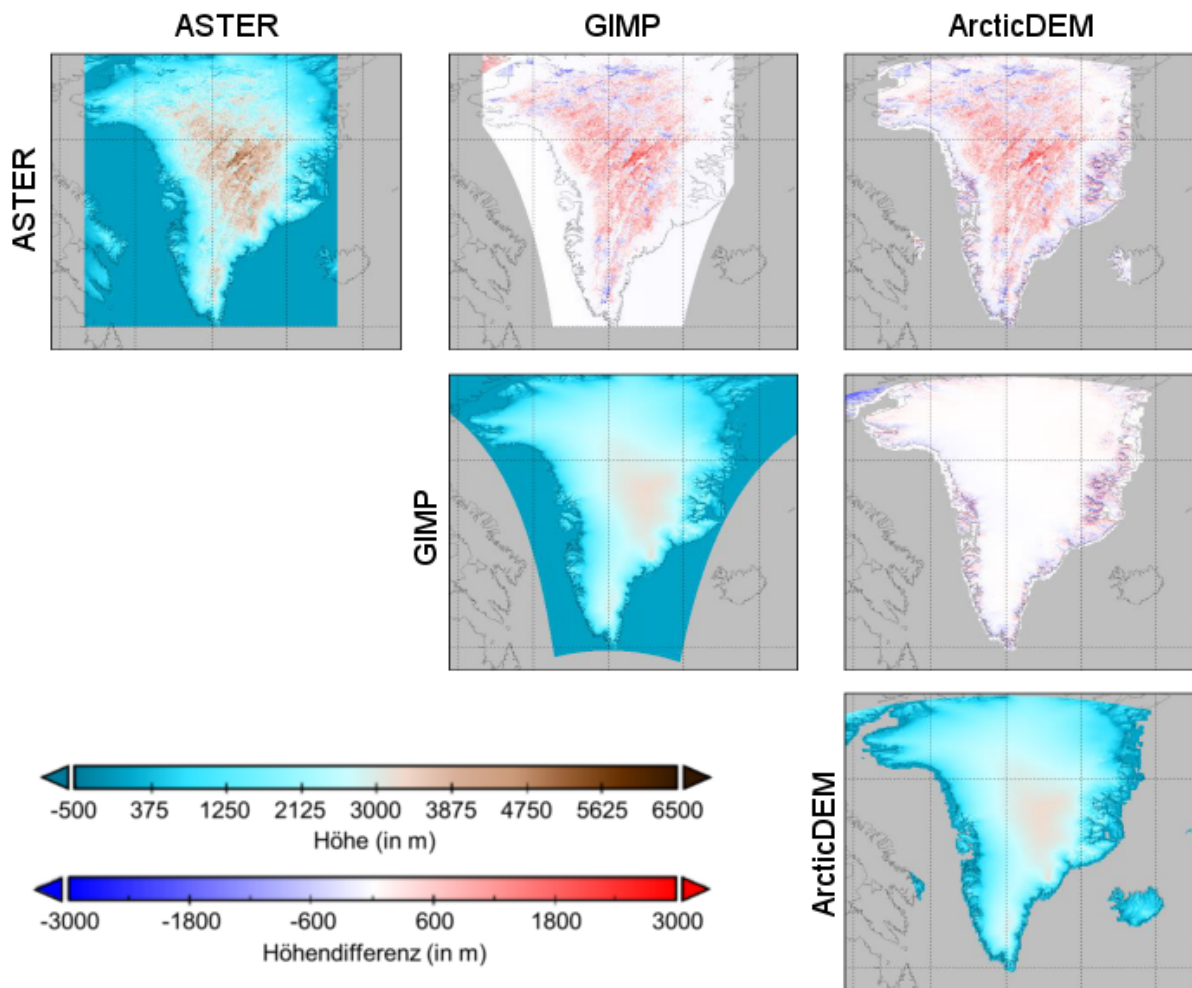


Abb. 24: Topographievergleich von ASTER, GIMP und ArcticDEM für Grönland

Wenn der ASTER-Datensatz für weitere Studien vor allem im Zentralbereich Grönlands verwendet werden soll, könnten die betroffenen „Grönland-Kacheln“ durch den GIMP-Datensatz ersetzt werden oder man verwendet die 2019 erschienene finale Version 3 des ASTER GDEM-Datensatzes. Diese Version schneidet laut Carabajal & Boy (2016), die mittels des Geoscience Laser Altimeter System (GLAS) auf dem Ice, Cloud and land Elevation Satellite (IceSat) der NASA die neuen ASTER DEM-Daten für Grönland validierten, besser ab als die ASTER GDEM V2. Da in dem in dieser Arbeit verwendeten Untersuchungsgebiet EUR-44 die fehlerhaften Topographiewerte am Rand des Gebiets in der Relaxationszone liegen, die nicht in die Analysen mit einfließt, wird auf diese Korrektur bzw. die Verwendung der ASTER GDEM V3 in dieser Arbeit verzichtet.

### **Zusammenfassung der Erkenntnisse des Topographie-Datenvergleiches**

Insgesamt lässt sich feststellen, dass der ASTER-Datensatz global keine systematischen Fehler im Vergleich zu GTOPO enthält. Es ist aber zu erkennen, dass sich selbst bei einer Aggregation von 1 km bzw. 30 m auf 50 km Auflösung vor allem in den gebirgrigeren Regionen und an den Küsten der Erde hohe Differenzen von mehreren 100 m zwischen den beiden Datensätzen ergeben. Eine detailliertere Untersuchung der Differenz beider Datensätze für das EUR-44- und das GER-11-Gebiet sowie die Auswirkungen dieser Differenzen auf die Modellperformanz wird in Kapitel 6.3.1 gegeben.

#### **4.4.2 Bodendaten**

Vergleicht man zunächst die anteilmäßigen globalen Texturklassenverteilungen der drei neuen Datensätze (FAOn, Soilgrid und HWSD) in ihrer jeweiligen Originalauflösung (Abb. 25), stellt sich heraus, dass zwar die generelle Verteilung bei allen übereinstimmt, einzelne Texturklassen aber größere prozentuale Unterschiede aufweisen. Beispielsweise hat die Klasse „Lo“ bei allen drei Datensätzen den höchsten Anteil. Jedoch fällt dieser Anteil beim FAOn-Datensatz mit 52.3% im Vergleich zum HWSD mit 32.4% viel höher aus. Die Anteile der Klassen „Sa“ und „SaLo“ weisen keine großen Unterschiede auf. Außerdem kommt die Texturklasse „Si“ in allen drei Datensätzen nicht vor. Auch die Mischklassen „SiCl“, „SaCl“ und „SiClLo“ treten mit Anteilswerten von unter einem Prozent kaum auf. Ein Unterschied zwischen SoilGrids und HWSD liegt vor allem in den Klassen „LoSa“ (2.34% und 10.53%), „SiLo“ (3.31% und 9.70%), „ClLo“ (10.06% und 4.70%) und „Cl“ (1.60% und 9.27%) vor. Bei den ersten drei genannten Klassen ist der FAOn-Datensatz jeweils dem SoilGrids-Werten ähnlich, bei „Cl“ jedoch dem HWSD-Wert. Nur bei der Klasse „SaClLo“ ähneln sich SoilGrids und HWSD stärker und FAOn bildet mit 0.77% im Vergleich zu 12.24% bzw. 8.31% den Ausreißer. Der Vergleich mit der bisherigen Texturkarte (FAO) fällt aufgrund der anderen Klassifizierung mit fünf statt der bei den neuen Datensätzen angewendeten USDA-Klassifikation mit 12 Texturklassen aus.

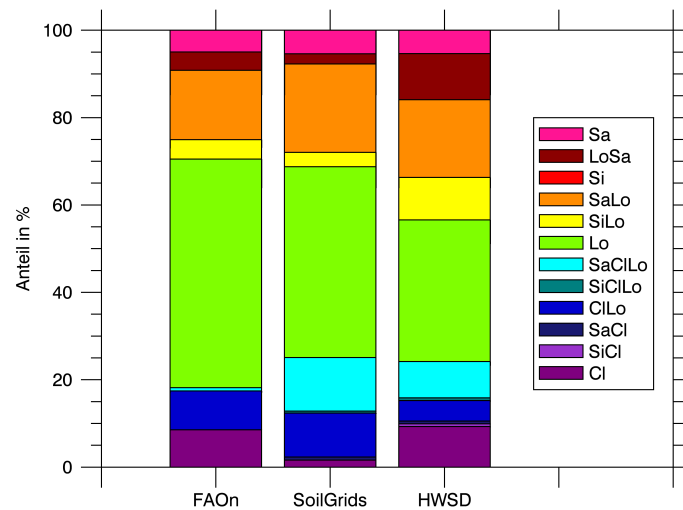
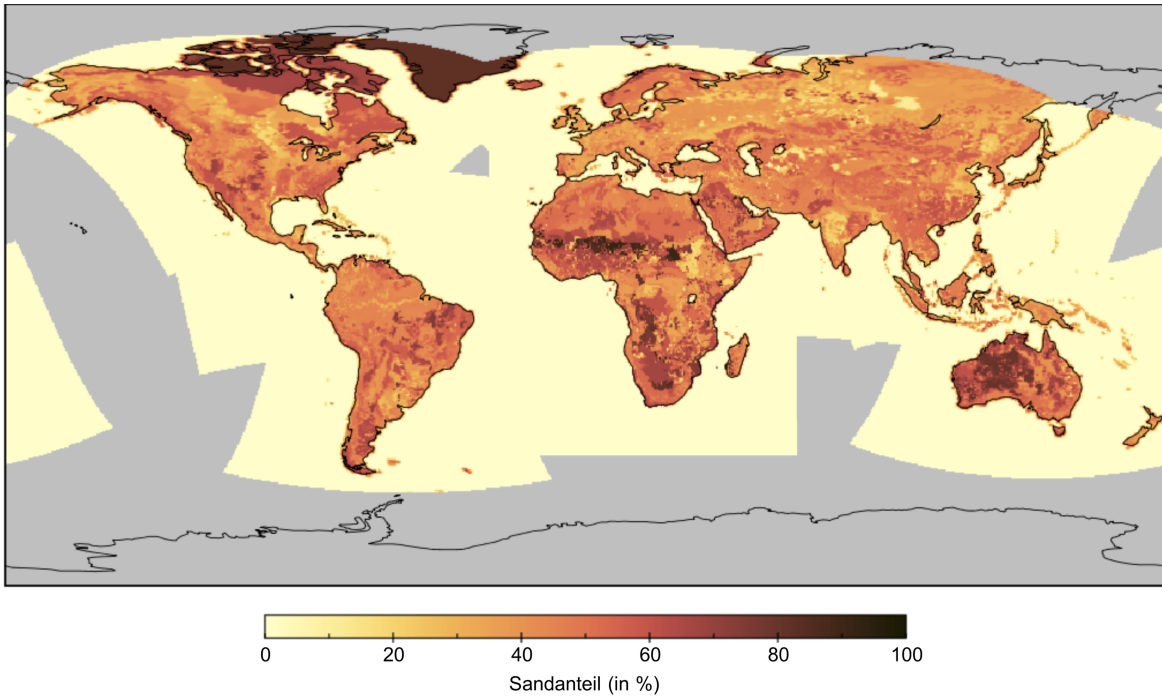


Abb. 25: Anteilsmäßige Texturklassenverteilung der neuen Bodendatensätze in deren Originalauflösung Sa: Sand ; Cl: Ton (Clay) ; Si: Schluff (Silt) ; Lo: Lehm (Loam)

In den Abbildungen 26, 28 und 31 werden die Variablen Sand-, Ton- und organischer Anteil vom neuen FAOn- und SoilGrids-Datensatz für die ganze Welt in einer Auflösung von  $0.44^\circ$  verglichen. Es ist zu beachten, dass dabei der FAOn-Datensatz bisher ebenfalls nicht in der Bodenbibliothek von REMO eingebaut war und lediglich als Referenz und zur Darstellung der Unterschiede von den Bodenvariablen dient. Der räumliche Vergleich mit dem HWSD-Datensatz wird nur für die beiden Untersuchungsgebiete EUR-44 und GER-11 in Kapitel 6.3.2 durchgeführt. Zudem beziehen sich alle dargestellten globalen Bodenkarten auf die Werte der obersten Bodenschicht (Layer 1), da sich diese aufgrund der benötigten Angleichung der Tiefen (vgl. Kapitel 4.1.1) am wenigsten von den Werten der originalen Datensätze unterscheiden.

Schaut man sich zunächst die generelle Sandanteil-Verteilung (SAND1) laut FAOn weltweit an, werden die höchsten Anteile auf größeren Flächen in Nordkanada und Grönland sowie in der Sahelzone und in Zentralaustralien erreicht. Im Allgemeinen lässt sich feststellen, dass es kaum Flächen mit weniger als 40% Sandanteil gibt (vgl. Abb. 26A). Anhand der Häufigkeitsverteilung (vgl. Abb. 27B) bestätigen sich die bereits aufgezeigten Unterschiede der beiden Datensätze. Insgesamt ist der Wertebereich von Soilgrid kleiner als der von FAOn, wobei jedoch alle Quantile von Soilgrid etwas höher liegen (vgl. Abb. 27A). Die Verteilung folgt mit wenigen Ausnahmen einer klassischen Normalverteilung mit einem Mittelwert von ca. 45%. Auffällig ist beim FAOn-Datensatz eine Häufung der Werte bei 82% Sandanteil, welche sich auch in der kumulierten Häufigkeitslinie durch einen unnatürlichen Sprung ausdrückt (vgl. Abb. 27C).

(A)



(B)

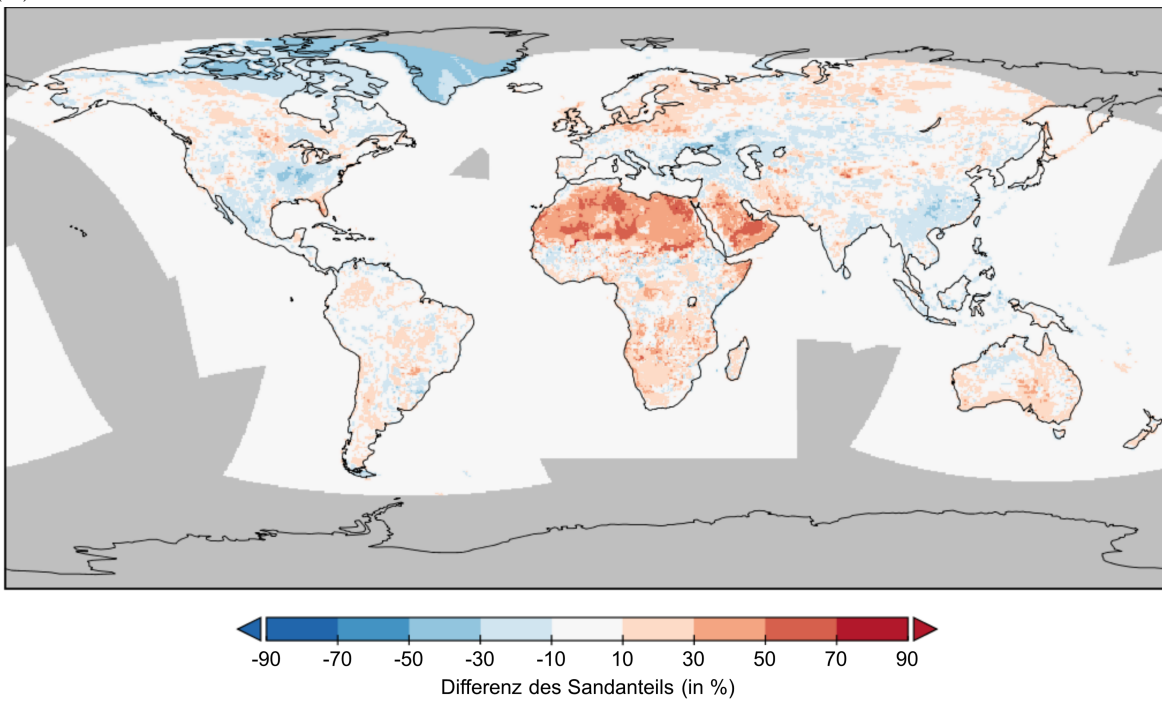


Abb. 26: Vergleich des Sandanteils des ersten Layers (SAND1) von FAOn- und SoilGrids-Datensatz

(A) FAOn

(B) SoilGrids-FAOn

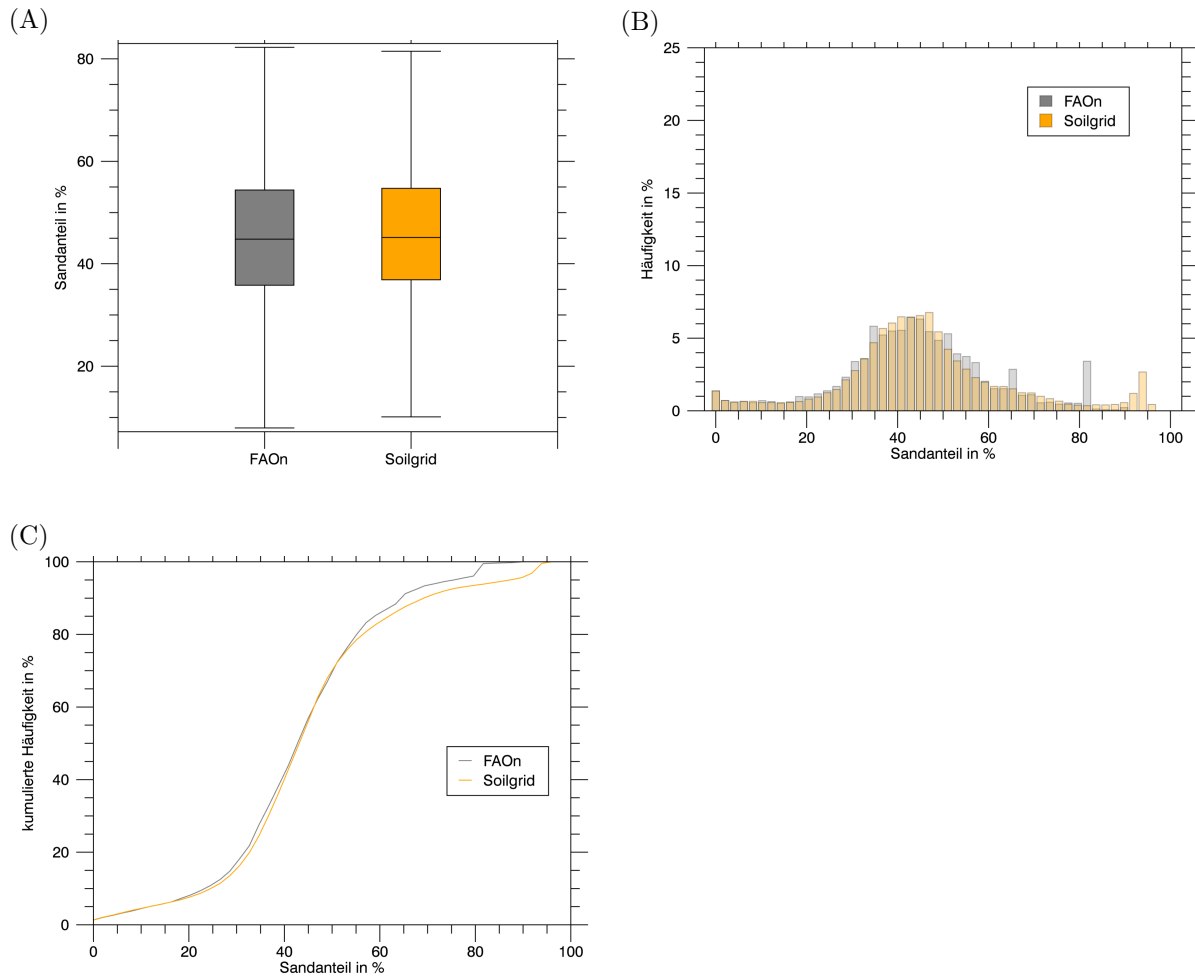


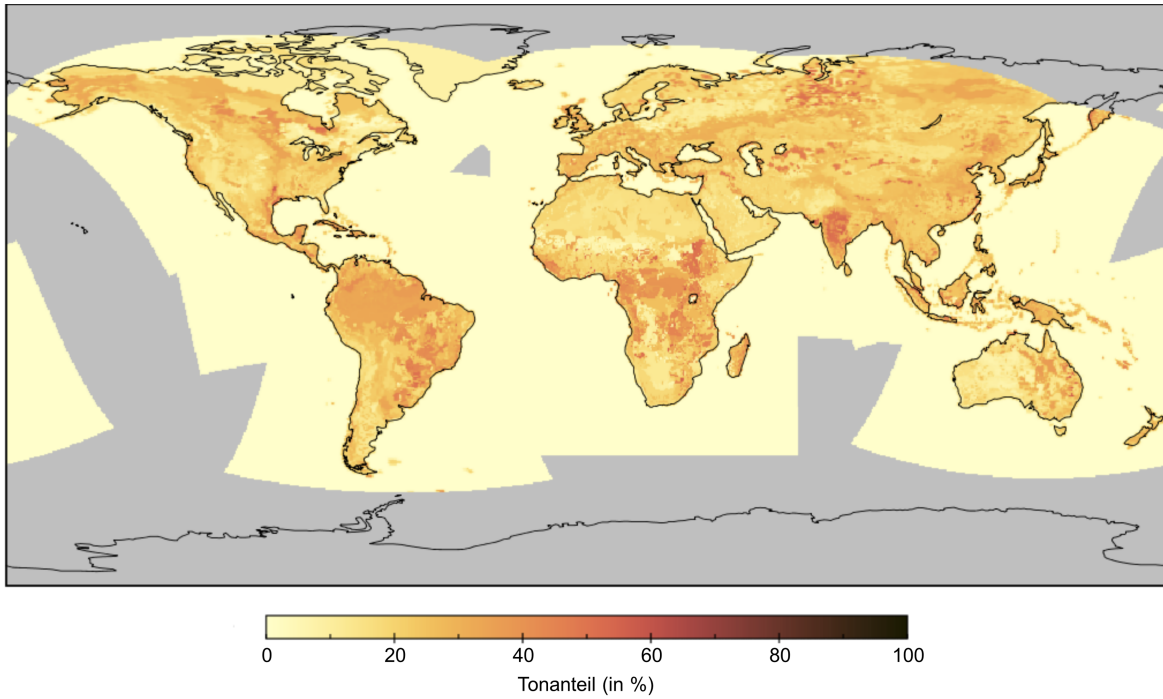
Abb. 27: Statistischer Vergleich von FAOn und Soilgrid für SAND in 0.44° Auflösung

- (A) Boxplot
- (B) Häufigkeitsverteilung
- (C) kumulierte Häufigkeitsverteilung

Beim in Abbildung 28 dargestellten globalen Tonanteil (CLAY1) liegen zumeist kleinere Anteile als beim Sand vor, da diese sich zusammen mit dem nicht dargestellten Schluffanteil zu 100% addieren und Sand bereits sehr hohe Anteile „belegt“. Es fällt auf, dass Soilgrid wieder einen kleineren Wertebereich als FAOn hat, beim Ton allerdings alle Quantile unterhalb der FAOn-Werte liegen (vgl. Abb. 29A). Außerdem gibt es größere Unterschiede vor allem im unteren Tonanteilsbereich, sodass die kumulierte Häufigkeitsverteilungskurven beim Tonanteil fast im gesamten Wertebereich weiter auseinander liegen als beim Sandanteil (vgl. Abb. 29C).

#### 4 Aufbereitung und Vergleich der neuen Oberflächendatensätze und Validationsdaten

(A)



(B)

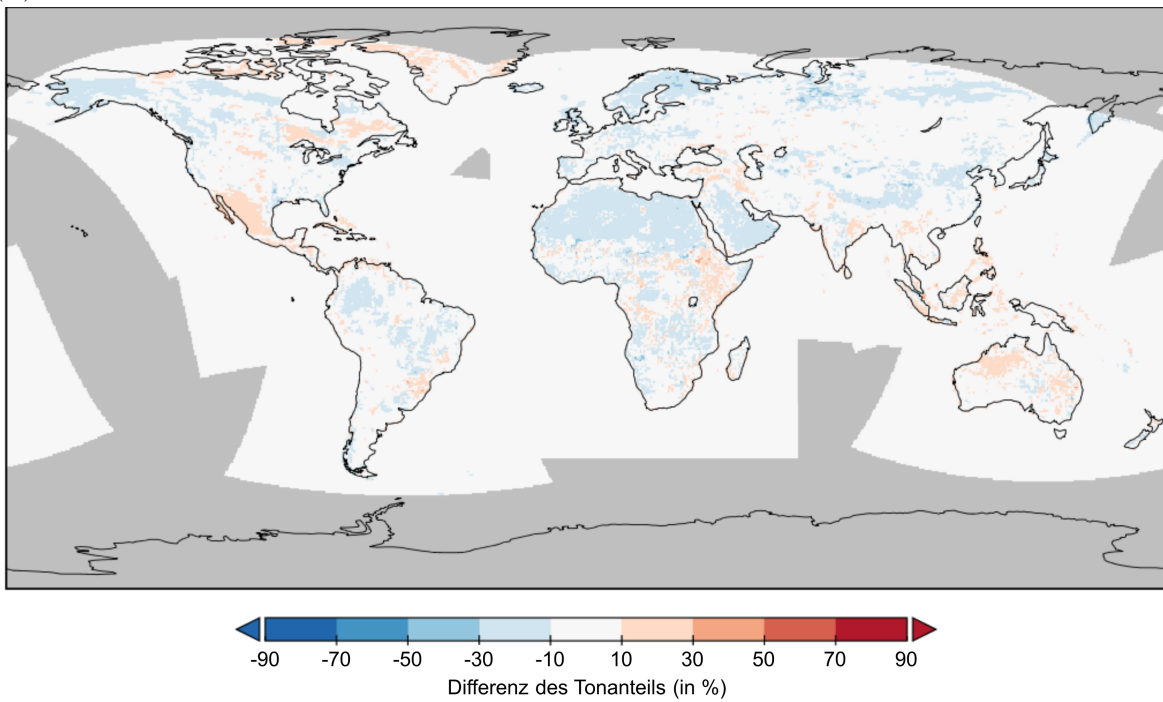


Abb. 28: Vergleich des Tonanteils des ersten Layers (CLAY1) von FAOn- und SoilGrids-Datensatz  
(A) FAOn  
(B) SoilGrids-FAOn

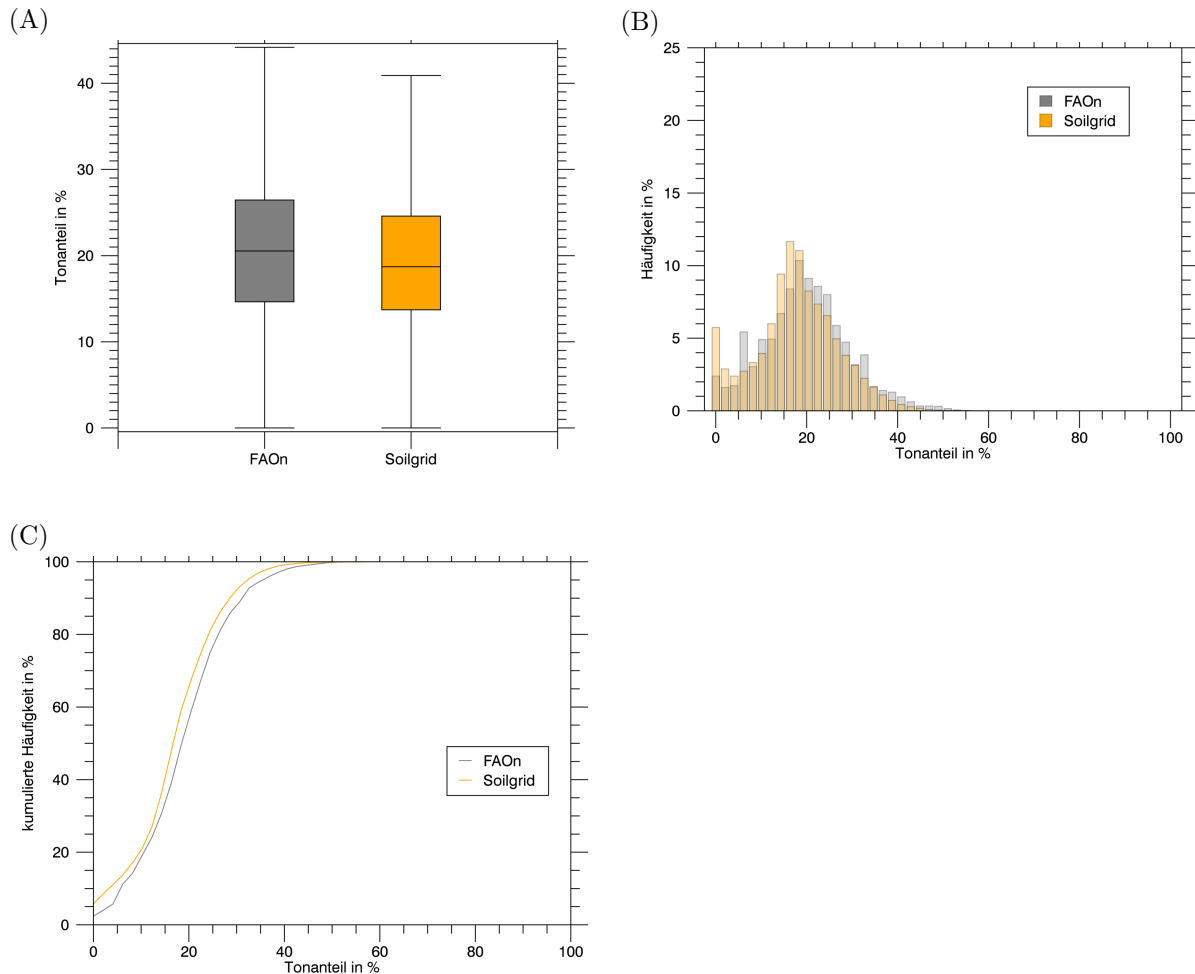


Abb. 29: Statistischer Vergleich von FAOn und Soilgrid für CLAY in 0.44° Auflösung

- (A) Boxplot
- (B) Häufigkeitsverteilung
- (C) kumulierte Häufigkeitsverteilung

Den größten Unterschied zwischen beiden Bodendatensätzen liefert der organische Anteil (FORG1). Dieser ist sowohl aus der Weltkarte (vgl. Abb. 31) als auch aus den Häufigkeitsanalysen (vgl. Abb. 30) abzulesen. Während beim FAOn-Datensatz nur wenige Pixel einen hohen organischen Anteil von über 0.3 besitzen, sieht man besonders gut an Abbildung 30B, dass Soilgrid hier wesentlich höhere Häufigkeiten hat. Auch bei den Boxplots ist das Wertespektrum von Soilgrid viel breiter. Der obere Whisker von Soilgrid liegt bei über 0.6, wohingegen der des FAOn-Datensatzes nur knapp 0.1 erreicht. Vergleicht man beide Karten miteinander, fällt auf, dass Soilgrid nur an den Stellen viel weniger organischen Anteil aufweist, wo FAOn besonders hohe Anteile hat (vgl. Alaska, Kanada südlich der Hudson-Bay oder Nordwestküste Russland). Ausgenommen ist dabei eine große Fläche in Nordsibirien, die in Soilgrid flächig mit fast 100% organischen Anteil ausgewiesen ist und auch bei FAOn über 0.5 liegt. Weiterhin fällt Grönland auf, das laut FAOn kaum organischen Anteil im Boden besitzt. Bei Soilgrid werden dort jedoch Anteile von 0.1–0.3 angezeigt. Außerdem scheint es, als würden in der Südhemisphäre kleinere Unterschiede zwischen den Datensätzen auftreten als nördlich des Äquators. Zu erklären ist dies mit der Tatsache, dass

hier generell kleinere organische Werte angegeben werden. Eine Ausnahme hiervon bildet die Küstenlinie in Südchile. Dort liegen die Soilgrid-Werte zwischen 0.3–0.5 höher als die von FAOn. Zudem existieren für die Südhemisphäre weniger Daten, sodass die verschiedenen Datensätze auf die gleichen Quellen zurückgreifen müssen und die Unterschiede dementsprechend nicht so groß sind.

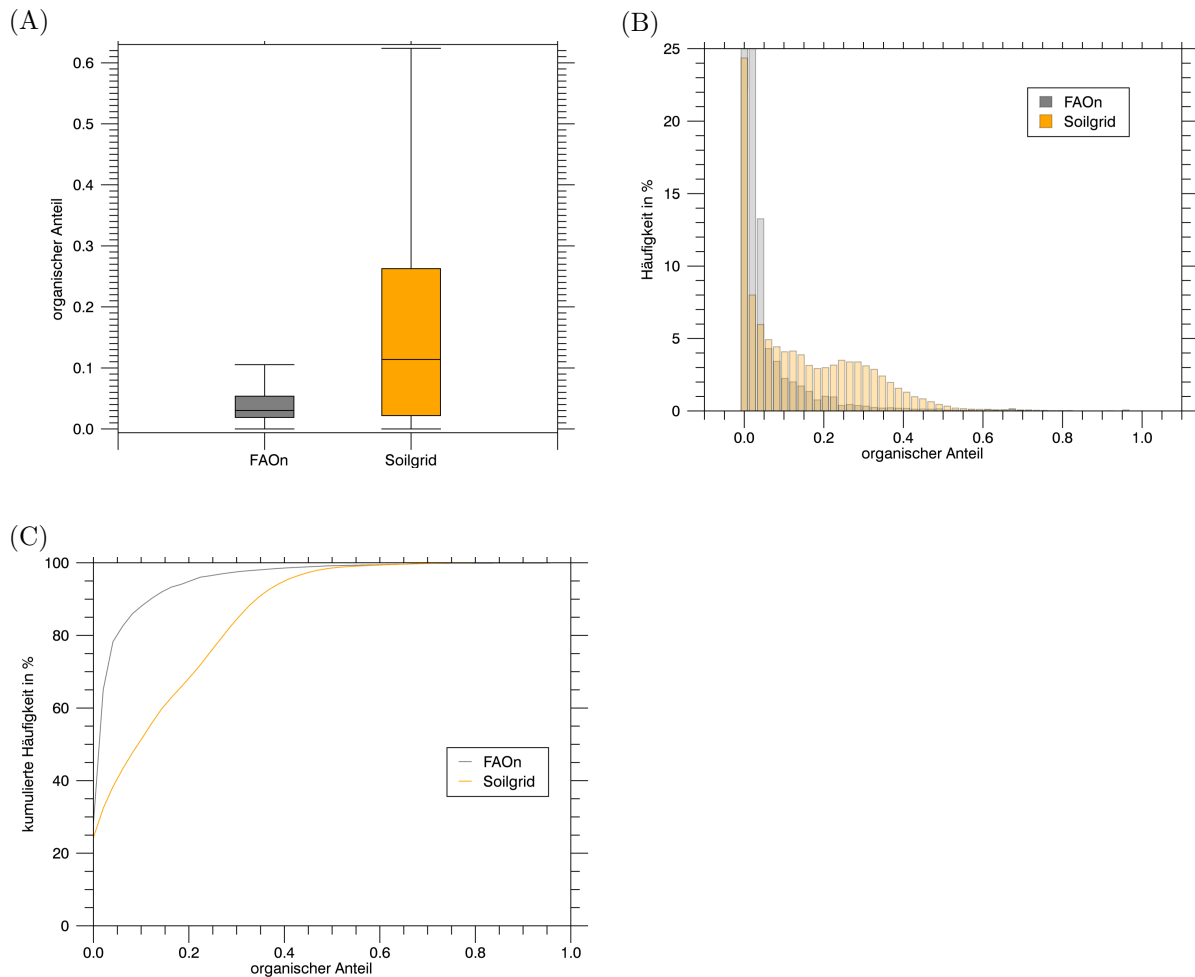


Abb. 30: Statistischer Vergleich von FAOn und Soilgrid für FORG in 0.44° Auflösung

- (A) Boxplot
- (B) Häufigkeitsverteilung
- (C) kumulierte Häufigkeitsverteilung

Schon aufgrund der 10fach bzw. 50fach besseren Auflösung von SoilGrids und HWSD im Vergleich zu den neuen bzw. alten FAO-Daten und im Hinblick auf die 2017 erschienene SoilGrids250m Version (Hengl et al., 2017), die bei Bedarf und ausreichend Speicherkapazität ohne größeren Aufwand wie der 1 km SoilGrids-Datensatz verwendet und in den Präprozess des Modells eingebaut werden kann, wird davon ausgegangen, dass diese Datensätze großes Potential für die Verbesserung der Bodenprozesse in REMO bieten. Eine genauere Analyse der aus diesen unterschiedlichen Datensätzen abgeleiteten Parameter und deren Auswirkungen auf die klimatologischen Variablen des Modells wird in Kapitel 6 vorgenommen.



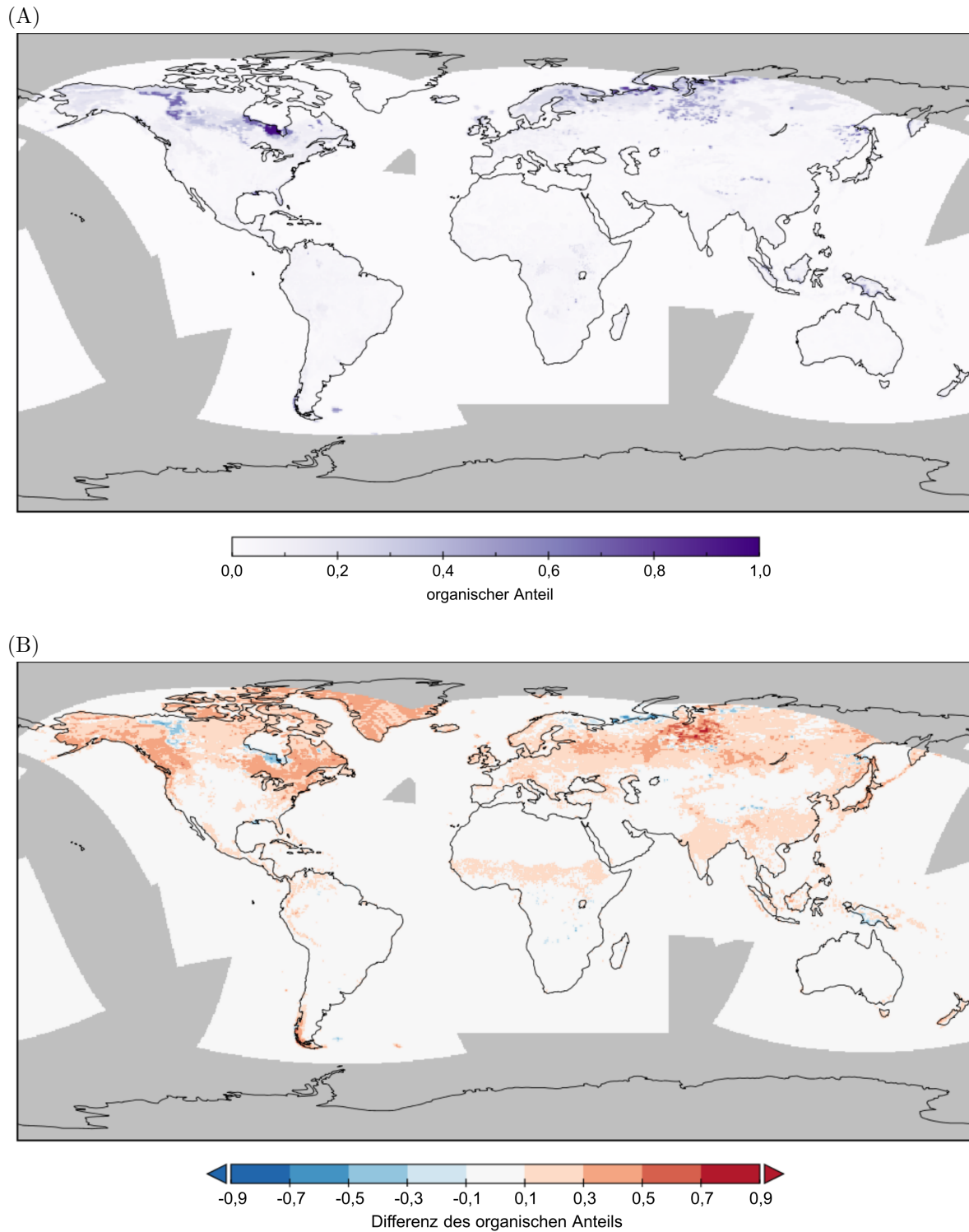


Abb. 31: Vergleich des organischen Anteils des ersten Layers (FORG1) von FAOn- und SoilGrids-Datensatz  
(A) FAOn  
(B) SoilGrids-FAOn

## 4.5 Beobachtungs- und Validierungsdaten

Um die Qualität von Klimamodellen zu testen, werden diese für Zeiträume in der Vergangenheit gerechnet und anschließend mit Beobachtungsdaten verglichen. Unter der Annahme, dass sich die Variabilität der Fehler nicht verändert, lassen sich so Unsicherheiten der Klimamodelle für die Zukunft einschätzen. In den folgenden Kapiteln werden die zur Validierung verwendeten klimatologischen Beobachtungsdatensätze und deren Aufbereitung näher erläutert.

### 4.5.1 Klimatologische Datensätze

Beobachtungsdaten lassen sich grundsätzlich in Raster- und Stationsdaten unterteilen. Stationsdaten sind Daten, die an verschiedenen Orten erhoben werden. Diese Messstationen liegen - außer in extrem kleinen Untersuchungsgebieten - nicht in einem definierten Netz vor, sodass die Dichte der Stationen im Raum und der Abstand zwischen den Stationen sehr stark variiert. In den allermeisten Fällen werden bei klimatischen Variablen die Rasterdatensätze aus Stationsdaten erzeugt, da ein flächendeckendes regelmäßiges Netz an Messungen in der Praxis nicht erreichbar ist. Ausnahmen sind beispielsweise die Oberflächentemperaturmessung oder Niederschlagswerte, die aus Satellitenprodukten erzeugt werden. Aber auch diese Daten müssen aufgrund von Messlücken oder unterschiedlichen räumlichen Abdeckungen aufbereitet werden. Je höher die Modellauflösung wird, desto schwieriger ist es allerdings, flächendeckende Beobachtungsdaten zur Validierung zu finden. Abhilfe schaffen entweder räumliche und zeitliche statistische Interpolation oder dynamische Interpolation. Diese dynamisch interpolierten Daten werden auch Reanalyse-Daten genannt (vgl. Kapitel 3.2.1) und werden mittels physikalischer Atmosphärenmodelle erzeugt. Nachfolgend werden die in dieser Arbeit verwendeten Validationsdaten vorgestellt.

#### 4.5.1.1 E-OBS

Für Europa liegt mit E-OBS ein täglich aufgelöster Rasterdatensatz für Minimum-, Maximum- und Mitteltemperatur (TN, TX, TG), Niederschlag (RR) und Bodendruck (PP) an Beobachtungsdaten vor. Der im EU-FP6 ENSEMBLES und EURO4M Projekt auf Grundlage von ca. 2300 Stationsdaten entwickelte Datensatz wird seit 2008 stetig verbessert. Für die als Ensemble konstruierten Daten werden jeweils der Mittelwert und die Abweichung (Differenz zwischen dem 5. und dem 95. Quantil) aller 100 Simulationen bereit gestellt (vgl. Haylock et al., 2008; Hofstra et al., 2009; van den Besselaar et al., 2011; Cornes et al., 2018). Eine detaillierte Analyse der Stärken und Schwächen des Datensatzes ist in Hofstra et al. (2009) zu finden. In dieser Arbeit wird die im März 2019 veröffentlichte Version E-OBS v19.0e (EOBSV19) und die im Mai 2020 veröffentlichte Version E-OBS v21.0e (EOBSV21) verwendet. Beide Versionen decken den Zeitraum von 1950-2019 ab und sind in einer Auflösung von  $0.1^\circ$  und  $0.25^\circ$  verfügbar. Die Höhendaten von E-OBS stammen aus dem GTOPO30 Datensatz, der ebenfalls bisher in REMO verwendet wird und in Kapitel 3.4.2 genauer vorgestellt wird.

#### 4.5.1.2 CRU

Der vom Climatic Research Unit (CRU) der Universität East Anglia zusammen mit dem Centre for Environmental Data Analysis (CEDA) im Mai 2019 veröffentlichte Datensatz CRU TS4.03 ist ein globaler Datensatz, der in dieser Version Monatswerte für Minimum-, Maximum-, Mitteltemperatur sowie tägliche Temperaturamplitude, Niederschlag, Dampfdruck und Wolkenbedeckung von 1901-2018 in einer Auflösung von  $0.5^\circ$  enthält (University Of East Anglia Climatic Research Unit et al., 2019). Die Daten werden durch eine Winkelabstandsgewichtete Interpolation (ADW) aus (sub-)täglichem Stationsdaten verschiedener nationaler meteorologischer Dienste erzeugt (Harris et al., 2014). In dieser Arbeit werden die drei Temperaturvariablen und die Niederschlagsvariable verwendet. Da jedoch keine topographischen Daten bereitgestellt werden, weil bei der Erzeugung der CRU-Daten keine Höhendaten, sondern Klima-anomalien verwendet werden (Harris et al., 2020), wird bei der Aufbereitung der CRU-Daten auf die Höhenkorrektur verzichtet. Aufgrund der Auflösung von  $0.5^\circ$  werden die CRU-Daten außerdem nur für die Validation der Modelldaten des EUR-44-Gebiets benutzt.

#### 4.5.1.3 ERA-Interim

Neben der Verwendung als Antriebsdaten (vgl. Kapitel 3.4.1) werden die ERA-Interim Daten in dieser Arbeit auch als Validationsdaten für das EUR-44 Untersuchungsgebiet genutzt. Die ERA-Interim Variablen Minimum-, Maximum- und Mitteltemperatur sowie Niederschlag, die nicht als Variablen in den Antriebsdaten von REMO verwendet werden, können beim Vergleich mit anderen Beobachtungsdaten zeigen, welchen Effekt das Betreiben des Klimamodells REMO auf die Vorhersagegüte hat. Wegen der groben Auflösung der Ausgangsdaten ( $0.75^\circ$ ) wird wie bei den CRU-Daten auf die Validation mit dem GER-11-Gebiet verzichtet und nur ein Vergleich mit den EUR-44 Modelldaten durchgeführt. In dieser Arbeit werden die täglichen Werte der vier Variablen (T2MIN, T2MAX, T2M und RR) sowie eine Karte des Geopotentials (in  $\frac{m^2}{s^2}$ ), die vom ECMWF zum Download bereitgestellt werden, verwendet. Für die notwendige Höhenkorrektur wird das Geopotential durch  $g = 9.80665 \frac{m}{s^2}$  dividiert, um die geopotentielle Höhe zu erhalten. Die Daten für das Geopotential werden aus GTOPO30 Daten, die für die Antarktis und Grönland angepasst wurden, erzeugt (vgl. Dee et al., 2011; Berrisford et al., 2011).

#### 4.5.1.4 ERA5

Die Reanalyse-Daten ERA5 sind das Nachfolgerprodukt von ERA-Interim des ECMWF. Die ERA5-Daten werden stündlich ab 1950 zur Verfügung gestellt. Neben der höheren zeitlichen Auflösung erreicht dieser neue Datensatz mit ca. 30 km und 139 Druckleveln auch eine höhere vertikale und horizontale räumliche Auflösung im Vergleich zu ERA-Interim. Zudem wird für die Erzeugung dieser Reanalyse-Daten das LSM Hydrology-Tiled ECMWF Scheme for Surface Exchanges over Land (HTESSEL) (Balsamo et al., 2009) statt der Vorgängerversion TESSEL (vgl. van den Hurk et al., 2000), die bei ERA-Interim verwendet wurde, benutzt. HTESSEL arbeitet

beispielsweise mit einem verbesserten hydrologischen Boden- und einem neuen Schneeschema (vgl. Balsamo et al., 2015). Die für die Erzeugung der ERA5-Daten verwendete ECMWF IFS Parametrisierung (CY41R2) benutzt eine Kombination aus SRTM30, Global Land One-km Base Elevation (GLOBE), Antarctic RAMP2, BPRC (Greenland) und Iceland DEM als topographische Daten (ECMWF, 2016). Wie bei ERA-Interim wird die Höhe als Geopotential zum Download angeboten und muss für die spätere Höhenkorrektur der Daten umgerechnet werden (Hersbach et al., 2018; Hersbach et al., 2020).

In dieser Arbeit werden die drei mittleren Temperaturen, der Niederschlag aller Datensätze sowie die u- und v-Windkomponente in 10 m Höhe von ERA5 als Validationsdaten verwendet. Trotz der Tatsache, dass ERA5 ein Nachfolger von ERA-Interim ist, unterscheiden sich die beiden Datensätze unter anderem aufgrund der beschriebenen Änderungen deutlich, sodass die ERA5-Daten zur Validation genutzt werden können, auch wenn ERA-Interim zum Antreiben der REMO-Modellläufe verwendet wird. In der Literatur gibt es einige Studien, die ERA-Interim und ERA5 miteinander vergleichen. Beispielsweise weisen Tarek et al. (2019) für Nordamerika nach, dass die Temperatur und der Niederschlag von ERA5 beim Vergleich zu Beobachtungsdaten kleinere Bias-Werte erzeugen als ERA-Interim. Auch Tetzner et al. (2019) bescheinigen ERA5 eine Verbesserung sowohl bei den Temperaturen als auch bei den Windregimen für die Antarktische Halbinsel. Zudem werden für diese Arbeit die Bodentemperaturen für mehrere Bodenschichten von ERA5 als Validationsdaten herangezogen, da für diese Variablen nur wenige Messdaten existieren. M. Li et al. (2020) zeigen für China, dass die ERA5-Reanalysedaten dort sehr gut mit Messdaten von Bodenwassergehalt und -temperaturen übereinstimmen. ERA5 schneidet bei dieser Studie im Vergleich zu den anderen vier dort betrachteten Datensätzen am besten ab. Wie bei ERA-Interim liegen Daten für die vier Bodenschichten (7, 28, 100, 289 cm Tiefe als Untergrenze der Schicht) vor, wobei die ausgewiesenen Bodentemperaturen den Temperaturen entsprechen, die in der Mitte jeder Schicht erreicht werden. Des Weiteren stellt ERA5 als Reanalyse-Datensatz Informationen für die verschiedenen Wärmeflüsse bereit. Für diese Arbeit werden einerseits die latenten und sensiblen Wärmeflüsse (surface latent/ sensible heat flux) sowie die kurzwellige und langwellige Nettostrahlung (surface net solar/ thermal radiation) an der Erdoberfläche für die Validation von REMO herangezogen. Da die verwendeten Werte aller Wärmeflussvariablen als stündliche Mittelwerte und in der Einheit  $\frac{\text{J}}{\text{m}^2}$  vorliegen, müssen sie durch 3600 s dividiert werden, um die gleiche Einheit ( $\frac{\text{W}}{\text{m}^2}$ ) wie die Modellvariablen von REMO zu erhalten.

### 4.5.1.5 DWD

Um nicht nur Reanalyse-Daten zur Validation der Bodenvariablen zu verwenden, werden außerdem die Bodentemperaturdaten des DWD herangezogen. Diese liegen als Rasterdaten in einer horizontalen Auflösung von etwa 1 km vor und stellen die tägliche Bodentemperatur in 5 cm Tiefe bei unbewachsenem sandigen Lehmboden dar. Die Bodentemperaturdaten werden dabei mittels des agrarmeteorologischen Modells AMBETI (Braden, 1995), welches ein physikalisch-mathematisches Modell des Systems Boden-Pflanze-Atmosphäre ist, anhand einiger Eingangsdaten wie Lufttemperatur, relativer Luftfeuchte, Niederschlag, Strahlungsgrößen und

Windgeschwindigkeit, initialen Bodentemperaturen und -wassergehalten sowie konstanter Boden- und Vegetationsinformationen berechnet. Das Modell wird für die Standorte der Wetterstationen (ca. 280 Stationen) betrieben. Anschließend wird mit Hilfe mehrerer gebietsabhängiger multipler linearer Regressionen, die die jeweiligen Stationshöhe, -longitude und -latitude als Prädiktoren verwenden, einer Interpolation unter Berücksichtigung der vier umliegenden Gebiete und einer Triangulation der Differenzen der modellierten und interpolierten Werte ein flächendeckender Datensatz für Deutschland erstellt. Es wird jedoch angemerkt, dass der resultierende Datensatz wegen der Interpolation auf Basis des relativ kleinen Stichprobenumfangs und vor allem im Winter durch die Isolationswirkung einer Schneedecke größere Ungenauigkeiten aufweisen kann (DWD, 2018).

## 4.5.2 Aufbereitung der Datensätze

Durch die Vielfalt an Beobachtungs- und Reanalysedaten, die in unterschiedlichen räumlichen und zeitlichen Auflösungen zur Verfügung gestellt werden, müssen diese vor der Verwendung als Validationsdaten des Klimamodells wie folgt aufbereitet werden.

**Räumliche Interpolation:** Wegen der unterschiedlichen räumlichen Auflösung von Beobachtungs- und Modelldaten muss eine räumliche Interpolation angewendet werden, um die Daten pixelweise miteinander vergleichen zu können. Für diese Art von Interpolation existieren viele verschiedene Methoden. In dieser Arbeit werden die verschiedenen Validationsdatensätze mittels der sogenannten Inverse Distance Weighted (IDW) Methode jeweils auf die Auflösung der beiden Testuntersuchungsgebiete EUR-44 und GER-11 gebracht. Dabei wird ein distanzgewichteter Mittelwert von Werten aus dem Ausgangsdatsatz für den Zielpunkt verwendet. Je nach Methode kann zum einen sowohl die maximale Distanz zwischen den Eingangspixeln und dem Zielpixel definiert als auch die maximale Anzahl der nächsten Eingangspixel festgelegt werden und zum anderen der Einfluss der Distanz durch den Exponenten variiert werden. In dieser Arbeit werden ausgehend vom jeweiligen Pixelmittelpunkt die nächsten zehn Eingangspixel für die Mittelwertbildung sowie die quadratische Distanzgewichtung verwendet (vgl. Formel 4.13, Shepard, 1968). Als Distanzmaß dient die Euklidische Distanz (vgl. Formel 4.14; Wilks, 2006).

$$x'_j = \begin{cases} \frac{\sum_{i=1}^n (w_{i,j} \cdot x_i)}{\sum_{i=1}^n w_{i,j}} & , \text{ wenn } d_{i,j} \neq 0 \text{ für alle } i \\ x_i & , \text{ wenn } d_{i,j} = 0 \text{ für ein } i \end{cases} \quad (4.13)$$

mit:

$i$  = Index des jeweiligen Pixels

$n$  = Anzahl der Pixel, die für die Interpolation verwendet werden

$x_i$  = Wert des Pixel  $i$

$x'_j$  = interpolierter Wert für Pixel  $j$

$w_{i,j} = \frac{1}{d_{i,j}^p}$  = Gewichtungsfaktor für den Pixel  $i$  und  $j$   
 $p$  = Distanzgewichtungsfaktor (hier: 2)

$$d_{i,j} = \sqrt{(lon_j - lon_i)^2 + (lat_j - lat_i)^2} \quad (4.14)$$

mit:

$lon_j$  = Longitude des Mittelpunkts des Zielpixels  $j$

$lat_j$  = Latitude des Mittelpunkts des Zielpixels  $j$

$lon_i$  = Longitude des Mittelpunkts des Eingangspixels  $i$

$lat_i$  = Latitude des Mittelpunkts des Eingangspixels  $i$

Zusätzlich wird eine bilineare Interpolation, die von der Software Climate Data Operators (CDO) bereitgestellt wird, zum Vereinigen auf ein gemeinsames Koordinatengitter genutzt (vgl. Schulzweida, 2020). Diese Methode verwendet das SCRIP Software Paket (Los Alamos National Laboratory) (vgl. Jones, 1998) und wird häufig für die Interpolation von klimatischen Variablen verwendet.

**Zeitliche Aggregation:** Die vorliegenden Datensätze werden zum einen auf Monats- und zum anderen auf Saisonwerte gebracht. Für die Temperatur wird der Mittelwert über den jeweiligen Monat bzw. die Saison berechnet. Beim Niederschlag, bei dem die Summe über einen definierten Zeitraum gebildet wird, ist es sehr wichtig, fehlende Werte zu identifizieren, diese zu interpolieren oder den resultierenden Wert aus der Analyse auszuschließen. Aus dem gleichen Grund werden bei der Ermittlung der Minimum- und Maximumtemperatur Fehlwerte im Aggregationszeitraum als Fehlwert im aggregierten Wert ausgewiesen. Da die in dieser Arbeit verwendeten Daten jedoch entweder aus Modell- oder Reanalysedaten, welche keine Fehlwerte enthalten, oder aus bereits aufbereiteten Beobachtungsdaten stammen, ist die Anzahl solcher Fehlwerte sehr gering.

**Höhenkorrektur:** Der sogenannte atmosphärische Temperaturgradient oder thermische Höhengradient (engl. lapse rate) beschreibt die Abnahme der Lufttemperatur mit zunehmender Höhe. Sie variiert nicht nur für unterschiedliche Höhenlagen, sondern auch zeitlich. Geeignete Höhenkorrekturmethode zu finden und verschiedene Methoden miteinander zu vergleichen, ist Gegenstand zahlreicher Forschungen (vgl. Minder et al., 2010 oder z. B. L. Gao et al., 2012 für ERA-Interim Daten). Da mehrere Temperaturdatensätze in dieser Arbeit verwendet werden, die unterschiedliche Topographiedaten als Grundlage voraussetzen, müssen diese Temperaturdaten aneinander sowie an die Modelldaten angeglichen werden. In dieser Arbeit wird (wie in vielen anderen Studien aufgrund der Einfachheit der Methode, vgl. Formel 4.15) für den Höhenkorrekturfaktor  $\Gamma$  der feuchtadiabatische Temperaturgradient von  $0.0065 \frac{\text{K}}{\text{m}}$  angenommen (vgl. Maurer et al., 2002; Stahl et al., 2006; Lundquist & Cayan, 2007).

$$T' = T + \Gamma \cdot \Delta h \quad (4.15)$$

mit:

$T'$  = auf die Höhe von Datensatz A korrigierte Temperatur eines Datensatzes B

$T$  = originale Temperatur eines Datensatzes B

$\Gamma$  = Höhenkorrekturfaktor (hier:  $0.65 \frac{K}{100m}$ )

$\Delta h$  = Höhendifferenz zwischen Datensatz A und B

### 4.5.3 Validationsdatensätze für die Bodenparameter

Neben dem Vergleich der klimatologischen Variablen von Modell- und Beobachtungsdaten zur Beurteilung der Qualität des Klimamodells können auch die „Zwischenschrittparameter“ validiert werden. Allerdings sind hochaufgelöste und flächendeckende Datensätze für Bodenparameter selten (vgl. Kapitel 4.3.2).

In dieser Arbeit wird die aus dem ursprünglichen Datensatz (Dai et al., 2013) weiterentwickelte „new version of the global high-resolution dataset of soil hydraulic and thermal parameters for land surface modeling“ (Dai, Wei et al., 2019; Dai, Xin et al., 2019) zur Validation der neuen REMO Bodenparameter herangezogen, da von diesen auch einige thermische Bodenparameter zur Verfügung gestellt werden. Bei diesem Datensatz (nachfolgend Dai-Daten genannt) werden sowohl das GSDE (Shangguan et al., 2014) als auch der neu in REMO eingebaute SoilGrids1km-Datensatz (vgl. Kapitel 4.3.2) verwendet, um Karten für die verschiedenen hydraulischen Parameter der Campbell (1974) und der van Genuchten (1980) Gleichung zu liefern. Außerdem werden Karten für die volumetrische Wärmekapazität des Feststoffanteils von trockenem Boden ( $csol\_1$  in  $J/m^3K$ ) und für die thermalen Leitfähigkeiten von gefrorenen und ungefroren gesättigten Boden und von trockenem Boden ( $ksat\_u\_1$ ,  $ksat\_f\_1$ ,  $kdry\_1$  in  $W/mK$ ) abgeleitet. Für die Validation der hydraulischen Parameter von REMO wird nur der gesättigte Wassergehalt ( $theta\_s\_1$  in  $cm^3/cm^3$ ) aus diesem Datensatz verwendet. Zusätzlich stehen alle Parameter für unterschiedliche Schichttiefen (angepasst an die Modellschichten von Noah-LSM, JULES oder CoLM/CLM) zur Verfügung. Die ebenfalls angebotene SoilGrids-Version, für die jedoch nur Daten für sechs „Zwischenschichten“ statt der sieben Originalschichten (vgl. Tab. 9) vorliegen, wird in dieser Arbeit zur Validation eingesetzt. Zur Ermittlung der thermischen Leitfähigkeit wird nach Dai, Wei et al. (2019) eine Kombination aus einem modifizierten Modell von Johansen (1975) und Balland & Arp (2005) genutzt. Auch bei den hydraulischen Parametern wurde ein Ensemble aus verschiedenen PTFs verwendet, um den systematischen Bias und die Unsicherheiten, die sich bei der Verwendung einer einzelnen PTF ergeben, zu minimieren (Dai, Xin et al., 2019).

Als weiterer Validationsdatensatz für die hydrologischen Bodenparameter wird die 2016 veröffentlichte Version des HiHydroSoil v1.2 verwendet. Dieser im Projekt FutureWater erstellte Datensatz berechnet mit verschiedenen Pedotransferfunktionen ebenfalls ausgehend vom SoilGrids1km verschiedenste bodenhydrologische Variablen. Fehlwerte, die in den originalen SoilGrids-Daten enthalten sind, werden mit der Global Soil Map of Hydraulic Properties von Droogers (2011), die auf den HWSD-Daten basieren, ersetzt. Alle Karten liegen daher wie Soilgrid in einer horizontalen Auflösung von ca. 1 km vor (de Boer, 2016). Für die Berechnung werden die ursprünglichen

sieben Soilgrid-Schichten verwendet (vgl. Tab. 9). Zur Verfügung gestellt werden aber nur zwei Schichtwerte (0–30 cm (Toplayer) und 30–100 cm (Sublayer)), die aus den drei bzw. zwei jeweiligen Schichtwerten als tiefengewichtete Mittelwerte berechnet werden. Für diese Arbeit werden die Feldkapazität (WCpF2 in  $\frac{\text{m}^3}{\text{m}^3}$ ) und der permanente Welkepunkt (WCpF4.2 in  $\frac{\text{m}^3}{\text{m}^3}$ ) (jeweils abgeleitet aus der van Genuchten Formel mit einem Matrixpotential von -100 bzw. -16 000 cm) sowie der gesättigte Wassergehalt ( $\theta_s$ , WCsat in  $\frac{\text{m}^3}{\text{m}^3}$ ) als Validationsdaten herangezogen. Diese werden für den HiHydroSoil wie folgt berechnet:

$$\theta_s = 0.83080 - 0.28217 \cdot p_b + 0.0002728 \cdot CLAY + 0.000187 \cdot SILT \quad (4.16)$$

mit:

$p_b$  = Trockenrohddichte in  $\frac{\text{g}}{\text{cm}^3}$   
 $CLAY$  = Tonanteil in %  
 $SILT$  = Schluffanteil in %

Zur einfacheren und konsistenten Handhabung wurden sowohl die Dai-Daten als auch die HiHydroSoil-Daten in netCDF Format konvertiert und mit den in Kapitel 4.1 und 4.5.2 beschriebenen Methoden auf die vertikale und horizontale Auflösung des jeweiligen Untersuchungsgebiets interpoliert.



# 5 Implementierung neuer Parametrisierungen und Prozesse sowie daraus resultierende Veränderungen im Programmcode für die Bodenparameter

Nachfolgend wird die Weiterentwicklung von REMO, die zusätzlich zu den in Kapitel 4.3 beschriebenen Datenänderungen im Rahmen der Arbeit stattgefunden hat, beschrieben. Es werden die Hintergründe, welche die Änderungen notwendig gemacht haben, die Theorie und die genutzten Methoden sowie die Implementierung der neuen Prozesse dargestellt.

## 5.1 Neue Berechnung der abgeleiteten Bodenparameter

Wie bereits in Kapitel 4.3.2 beschrieben, existieren weltweit Bodendatenbanken, auf deren Grundlage viele verschiedene Studien erstellt werden, um Prozesse innerhalb des Bodens sowie Interaktionen zwischen Boden und Atmosphäre zu untersuchen. Allerdings enthalten die erwähnten globalen Bodendatenbanken nicht alle thermalen und hydrologischen Parameter, die für ein LSM oder Klimamodell benötigt werden (Dai, Shangguan, Wei et al., 2019), sodass diese Parameter aus den existierenden Parametern mittels PTFs berechnet oder mittels „Lookup“-Tabellen zugewiesen werden müssen. Laut Bouma (1989), der den Begriff Pedotransferfunktion eingeführt hat, transformiert eine PTF Daten, die man hat, in Daten, die man braucht. Zum Erstellen einer PTF wird dabei auf unterschiedliche physikalische, chemische oder biologische Prädiktoren zurückgegriffen, um einen anderen Bodenparameter zu schätzen (Padarian et al., 2018). PTFs können in zwei Gruppen unterteilt werden. Zum einen gibt es klassenbasierte und zum anderen kontinuierliche PTFs, welche sich wiederum in punkt- und parameterbasierte PTFs gliedern lassen. Bei den klassenbasierten PTFs handelt es sich um Werte in Lookup-Tabellen, die im Gegensatz zu den kontinuierlichen PTFs, welche als Funktionen definiert sind, nur für spezielle Klassen definiert sind. Eine andere Unterteilung der PTFs beruht auf der zur Erzeugung der PTF verwendeten mathematischen Methode (z. B. Regressionsanalyse oder Artificial Neural Networks (ANN)) (Wösten et al., 1995; Wösten et al., 2001; Vereecken et al., 2010; van Looy et al., 2017; Al Majou et al., 2018). Aufgrund der Einfachheit und des vermehrten Verwendens der klassenbasierten PTFs in der Anfangszeit der PTFs finden sich diese auch heute noch in vielen LSMs und Klimamodellen (z. B. Bodenparameter in REMO aus Tab. 3). Weitere Informationen und einen guten Überblick über die Entstehung und Entwicklung von verschiedenen PTFs bieten Pachepsky & Rawls (2004) und van Looy et al. (2017). Im folgenden Abschnitt werden die in dieser Arbeit getesteten Methoden zur Generierung von PTFs genauer beschrieben, sowie deren Vor- und Nachteile diskutiert. Trotz der Tatsache, dass einige der dargestellten Methoden nicht in die in dieser Arbeit vorgestellten REMO-Versionen eingebaut wurden, werden sie zum

Verständnis des Vorgehens und zur Erklärung der getroffenen Entscheidungen erläutert.

### **5.1.1 Pedotransferfunktion für Bodenparameter**

Durch die Änderung der Eingangsdaten für den Boden und der damit einhergehenden Veränderung der Anzahl der Texturklassen (vgl. Kapitel 4.3.2) werden für REMO neue Parameterwerte beispielsweise für die Feldkapazität, das Porenvolumen oder die thermale Diffusivität benötigt. Die Suche nach geeigneten Werten für die neuen Texturklassen beginnt mit einem Vergleich der bei der Literaturrecherche gefundenen klassenbasierten und kontinuierlichen PTFs.

#### **Klassenbasierte PTF aus der Literatur**

In der Tabelle in Anhang G wurden verschiedene Werte für die jeweiligen Parameter aus der Literatur zusammengetragen. Zum Vergleich wurden auch die bisherigen Werte von REMO (vgl. Tab. 3) in diese Tabelle in die Texturklasse mit dem gleichen Namen eingetragen. Die dargestellten Werte wurden entweder direkt aus den angegebenen Quellen übernommen oder indirekt durch Bildung des Mittelwerts sowie durch Berechnung ermittelt (vgl. Formel 2.6 bzw.  $DLAMBDA = \lambda_{sat} - \lambda_{dry}$ ). Hierbei ist zu beachten, dass sich die thermalen Parameter abhängig vom Wassergehalt nicht linear verhalten. Daher sind Angaben zur Wärmekapazität und zur thermalen Diffusivität bei mittlerem Wassergehalt (vgl. Parameter 1 und 3) in der Literatur schwer zu finden und kaum aus Werten mit anderem Wassergehalt abzuleiten. Insgesamt ist zu erkennen, dass die jeweiligen Werte große Intervalle aufweisen und teilweise unvollständig sind. Ein Teil der Abweichung kann durch die starke Abhängigkeit von Wasser- und Luftgehalt vor allem der thermalen Parameter erklärt werden, da hierzu in den meisten Quellen unvollständige Angaben gemacht werden. Ein weiteres Beispiel für die schlechte Vergleichbarkeit von kategorialen Angaben sowie für einen weiteren Einflussfaktor auf die Wertebildung pro Texturklasse liefert der in F. Chen & Dudhia (2001) ausgewiesene hydrologische Wert für organisches Material, welcher hier zunächst der Klasse Torf zugeordnet wurde. Für Porenvolumen und Feldkapazität mit 0.439 bzw. 0.329 sind die Werte nur halb so groß wie andere Werte für Torf. Es ist davon auszugehen, dass diese niedrigen hydrologischen Parameter nur für weniger organisches Material als Torf repräsentativ sind. Gleichzeitig ist der volumetrische Wärmekapazitätswert für trockenen Torfboden von Guimberteau et al. (2018) mehr als dreimal so hoch wie in allen anderen Quellen und würde eher den Werten für die Wärmekapazität von Boden mit mittlerem Wassergehalt entsprechen. Der Wärmekapazitätswert unterscheidet sich auch deutlich von dem entsprechenden Wert bei Pielke (1984), obwohl diese beiden Quellen ansonsten in allen Texturklassen die gleichen thermischen Werte angeben. Ein weiterer Grund für die großen Unterschiede ist, dass die Angaben in der Literatur teilweise aus Stichproben von Laboranalysen stammen und sich demnach auf ein gewisses Untersuchungsgebiet beziehen sowie teilweise mit Hilfe von idealisierten PTFs berechnet wurden.

Außerdem fällt auf, dass sich für die hydrologischen Parameter Porenvolumen und Feldkapazität mehr Quellen finden, welche die einzelnen Texturklassen ausweisen, als für die thermalen Para-

meter. Bei den hydrologischen Parametern lässt sich zudem die Reihenfolge der Texturklassen hinsichtlich der Parametergröße meist erkennen, während bei den thermalen Parametern die Reihenfolge je nach Quelle variiert. Besonders auffällig ist der volumetrische Wärmekapazitätswert von trockenem Boden. Zu beachten ist, dass F. J. Braun (2002) und Ochsner et al. (2001) die Wärmekapazität der trockenen Bodensubstanz pro Texturklasse angeben, welche mit dem Wert  $(1 - \theta_s)$  multipliziert werden muss, um den Wert für den benötigten Feststoffanteil von trockenem Boden zu erhalten. Als Porenvolumen  $\theta_s$  wird (wie bei F. J. Braun, 2002) jeweils die Parametrisierung nach Rawls & Brakensiek (1982) gewählt. Während der Wert für die Klasse Cl bei Braun mit  $1.1 \cdot 10^6 \frac{\text{J}}{\text{m}^3 \text{K}}$  zu den Texturklassen mit sehr niedriger Wärmekapazität gehört (wie bei Pielke, 1984), weisen die anderen Quellen diese Tonklasse mit höheren Werten aus. Bei Ochsner wird der geringste bzw. höchste Kapazitätswert für die Klasse SaLo bzw. SiClLo erreicht, bei allen anderen Quellen sind die Werte genau entgegengesetzt. Des Weiteren liegt laut Markert (2012) und Márquez et al. (2016) kein Unterschied in der Kapazität von Sand, Ton und Schluff vor (je  $1.6 \cdot 10^6 \frac{\text{J}}{\text{m}^3 \text{K}}$ ). Bei den anderen Quelle ergeben sich zumeist die größten Unterschiede zwischen den Werten der Klassen Sand und Ton.

Für den Vergleich zu den bisher in REMO verwendeten Werten wurden Mittelwert, Standardabweichung sowie Minimum und Maximum der Literaturwerte in Tabelle 11 pro Texturklasse und Parameter aufgetragen. Es fällt auf, dass für die volumetrische Wärmekapazität für einen Boden mit mittlerem Wassergehalt und auch für einen trockenen Boden jeweils der REMO-Wert für die Texturklasse Cl deutlich über dem Maximum der Literaturwerte und für Sa knapp unter dem Minimum liegt. Die Standardabweichung dieser Texturklassen ist jedoch sehr hoch, verglichen mit den nur für trockenen Boden vorliegenden Werten für ClLo, Lo und LoSa. Von diesen Texturwerten befindet sich nur der REMO-LoSa-Wert innerhalb des durch die Literaturrecherche definierten Wertebereichs. Besonders auffällig ist außerdem die große Variationsbreite von RGCGN (trockener Boden) für Torf. Hier entspricht der bisherige REMO-Wert ( $0.58 \cdot 10^6 \frac{\text{J}}{\text{m}^3 \text{K}}$ ) dem Minimum der Literatur, der Mittelwert der Literatur ist jedoch fast doppelt und der Maximumwert mehr als viermal so groß. Auch bei der thermalen Diffusivität weicht der Torfwert sehr stark ab, wobei berücksichtigt werden muss, dass hier nur ein Vergleichswert vorliegt. Für die Texturklasse Sa dagegen liegt der bisherige Diffusivitätswert nur knapp über dem Maximum der Literaturwerte. Beim Parameter Wärmeleitfähigkeit von trockenem Boden entsprechen die REMO-Werte außer bei der Texturklasse Cl relativ gut den mittleren Literaturwerten. Die Amplitude der Wärmeleitfähigkeit wird dagegen vom bisherigen REMO-Wert in allen drei angegebenen Texturklassen überschätzt. Bei den hydrologischen Parametern (PORVOL und FCAP) fällt auf, dass der bisherige Wert immer am oberen bzw. knapp über dem Maximum der Literaturwerte liegt. Die Standardabweichung und der Wertebereich sind für alle Texturklassen relativ klein. Eine Ausnahme hiervon ist die Torfklasse mit jeweils größeren Variationen, wobei sich der bisherige REMO-Wert für das Porenvolumen und die Feldkapazität jeweils innerhalb des Wertebereichs befindet.

## 5 Implementierung neuer Parametrisierungen und Prozesse sowie daraus resultierende Veränderungen im Programmcode für die Bodenparameter

Tab. 11: Vergleich der verschiedenen Bodenparameter von REMO mit Literaturwerten pro Texturklasse (Zusammenfassung von Anhang G)

- 1: bisheriger REMO-Wert (vgl. Tab. 3)
- 2: Mittelwert und Standardabweichung der Literaturwerte
- 3: Minimum der Literaturwerte
- 4: Maximum der Literaturwerte

Parameter		vol. Wärmekapazität des Bodens bei mittlerem Wassergehalt	vol. Wärmekapazität des trockenen Bodens	thermale Diffusivität bei mittlerem Wassergehalt	Wärmeleitfähigkeit von trockenem Boden	Amplitude der Wärmeleitfähigkeit	Porenvolumen	Feldkapazität
		$RCCGN$ ( $LWDIF$ = $FALSE$ )	$RCCGN$ ( $LWDIF$ = $TRUE$ )	$TLAMBDA$ ( $LWDIF$ = $FALSE$ )	$TLAMBDA$ ( $LWDIF$ = $TRUE$ )	$DLAMBDA$	$PORVOL$	$FCAP$
		$\frac{10^6 J}{m^3 K}$	$\frac{10^6 J}{m^3 K}$	$\frac{10^{-6} m^2}{s}$	$\frac{W}{mK}$	$\frac{W}{mK}$	$\frac{m^3}{m^3}$	$\frac{m^3}{m^3}$
Cl	1	2.48	1.63	0.67	0.18	1.50	0.51	0.46
	2	2.35 ( $\pm 0.14$ )	1.32 ( $\pm 0.22$ )	0.53	( $\pm 0.14$ )	1.27 ( $\pm 0.08$ )	0.45 ( $\pm 0.05$ )	0.40 ( $\pm 0.04$ )
	3	2.25	1.09		0.25	1.20	0.38	0.35
	4	2.45	1.60		0.50	1.33	0.50	0.45
ClLo	1	2.36	1.50	0.71	0.21	1.55	0.48	0.37
	2		1.24 ( $\pm 0.04$ )		0.23		0.45 ( $\pm 0.03$ )	0.32 ( $\pm 0.05$ )
	3		1.18				0.41	0.27
	4		1.29				0.48	0.38
Lo	1	2.25	1.42	0.74	0.25	1.58	0.46	0.34
	2		1.23 ( $\pm 0.02$ )		0.22		0.44 ( $\pm 0.01$ )	0.26 ( $\pm 0.07$ )
	3		1.21				0.43	0.17
	4		1.25				0.45	0.33
LoSa	1	2.10	1.35	0.80	0.28	2.40	0.45	0.26
	2		1.37 ( $\pm 0.04$ )		0.23		0.41 ( $\pm 0.02$ )	0.17 ( $\pm 0.09$ )
	3		1.32				0.39	0.07
	4		1.41				0.43	0.28
Sa	1	1.93	1.28	0.87	0.30	2.40	0.36	0.20
	2	2.39 ( $\pm 0.32$ )	1.43 ( $\pm 0.13$ )	0.79 ( $\pm 0.08$ )	0.32 ( $\pm 0.08$ )	1.95 ( $\pm 0.06$ )	0.39 ( $\pm 0.03$ )	0.13 ( $\pm 0.08$ )
	3	2.12	1.28	0.73	0.22	1.90	0.34	0.05
	4	2.75	1.60	0.85	0.40	2.00	0.43	0.24
Torf	1	2.59	0.58	0.65	0.06	0.50	0.86	0.76
	2	2.18	1.07 ( $\pm 0.81$ )	0.13	0.06 ( $\pm 0.01$ )	0.44 ( $\pm 0.00$ )	0.78 ( $\pm 0.18$ )	0.60 ( $\pm 0.39$ )
	3		0.58		0.05	0.44	0.44	0.33
	4		2.50		0.06	0.44	0.92	0.88

Dieser Vergleich zu den Parametern der bisherigen fünf Texturklassen zeigt, dass es ein großes Spektrum an Variabilität gibt und sich die Literaturwerte teils stark von den bisherigen Parame-

terwerten unterscheiden. Sie differieren in manchen Bereichen so stark, dass sowohl die einzelnen Texturklassenwerte je nach Quelle unterschiedlichen Texturklassen entsprechen als auch der jeweilige Wertebereich stark variiert. Beispielsweise ist der Wert der volumetrischen Wärmekapazität von Boden mit mittlerem Wassergehalt bei Witt (2018) für Sand mit  $2.75 \cdot 10^6 \frac{\text{J}}{\text{m}^3 \text{K}}$  so hoch, dass der Wert nicht einmal für die Klasse Ton bzw. Torf (höchste volumetrische Wärmekapazität) der bisherigen Parametrisierung reicht, obwohl Sand den kleinsten Wert haben sollte.

Anhand der Tabelle in Anhang G, die nur einen kleinen Einblick über die in der Literatur ausgewiesenen Werte darstellt, ist außerdem zu erkennen, dass zwar für einige Studien die gleichen Parametrisierungen verwendet, aber häufig sehr unterschiedliche Werte für Analysen genutzt werden. Eine einheitliche Parametrisierung oder ein viel genutztes Standardwerk existiert nicht. Aufgrund dieser Vielzahl an sehr unterschiedlichen und unvollständigen Werten für die neuen Texturklassenwerte und wegen der Tatsache, dass Klimamodelle stark von ihren jeweiligen Parametrisierungen geprägt werden, wird vom Einbau von klassenbasierten PTFs aus der Literatur abgesehen.

Ein erster Ansatz, um dennoch die neuen Bodendatensätze für REMO verwenden zu können, basiert darauf, zunächst nur die Verbesserung der räumlichen Auflösung zu nutzen. Dafür müssen zu Beginn die Texturkarten der neuen Datensätze aufbereitet werden, denn alle neuen besser aufgelösten Bodentexturdatensätze (FAOn, SoilGrids und HWSD) verwenden die 12 Texturklassen der USDA-Klassifizierung. Damit die jeweils 12 benötigten Parameterwerte so konsistent wie möglich zu den bisherigen fünf Werten passen, werden die neuen Parameterwerte an den bisherigen Werten angelehnt. Im Folgenden werden mehrere Möglichkeiten vorgestellt, die Werte für die neuen „Texturmischklassen“ zu generieren.

## Eigene PTFs aus bisherigen Stützwerten

### 1. Ansatz: Konfusionsmatrix

Für einen besseren Überblick und einen räumlichen Vergleich des alten und der neuen Datensätze auch im Hinblick auf die einfache Übernahme der bisher genutzten Parameterwerte der Texturklassen, wurde als erster Schritt eine räumlich Analyse der alten und neuen globalen Texturklassenwertkarten anhand einer Konfusionsmatrix durchgeführt. Der Vorteil dieser Methode ist der geringe analytische und programmiertechnische Aufwand. Die bisherigen Klassenwerte können für die neuen Texturklassen mit der jeweils höchsten flächenmäßigen Übereinstimmung übernommen werden, sodass der Wertebereich der bisherigen Parameter im Modell gleich bleibt, die räumliche Auflösung jedoch deutlich besser werden würde. Gleichzeitig ist dies auch der Nachteil dieser Methode, da es keine inhaltliche Verbesserung der genutzten Bodeneigenschaften liefert und auch die sehr viel höhere Auflösung der neuen Daten nicht voll ausgenutzt werden würde. Zudem zeigt die erstellte Konfusionsmatrix keine eindeutigen Ergebnisse und eine nachvollziehbare Zuweisung der Klassen, die zumindest grob durch die Klassenbezeichnungen überprüfbar ist, ist nicht möglich. Daher wurde dieser erste Ansatz verworfen.

## 2. Ansatz: Regression mit durchschnittlichen Sand- und Tonanteilen

Der zweite Ansatz macht sich zunutze, dass es in der Boden-Community mehrere Möglichkeiten gibt, einen Boden mit entsprechendem Sand-, Ton- und Schluffanteil einer Texturklasse zuzuweisen. Die Unterteilung von Sand, Ton und Schluff wird anhand des Durchmessers der Partikel vorgenommen. Eine Korngröße von 0.05–2 mm wird der Klasse Sand, von 0.0002–0.05 mm Schluff und kleiner als 0.0002 mm Ton zugeordnet. Die am häufigsten verwendete Klassifikation ist die USDA-Klassifikation. Eine gängige Darstellung der Anteile Sand, Ton und Schluff von Bodenproben ist das gleichseitige Texturdreieck (Texture Triangle) (Gee & Or, 2002; Blott & Pye, 2012). Da Sand-, Ton- und Schluffanteil zusammen 1 bzw. 100% ergeben müssen, werden lediglich zwei der drei Komponenten zur eindeutigen Definition benötigt. Für gewöhnlich sind das Sand als x-Koordinate und Ton als y-Koordinate (vgl. mittleres Schema in Abb. 32). Um einen beliebigen Punkt im speziellen USDA-Triangle, welches in der vollständigen Form im Dreieck Flächen für die einzelnen Texturklassen ausweist (vgl. Abb. 33), darstellen zu können, müssen der Sand- und Tonanteil umgerechnet werden.

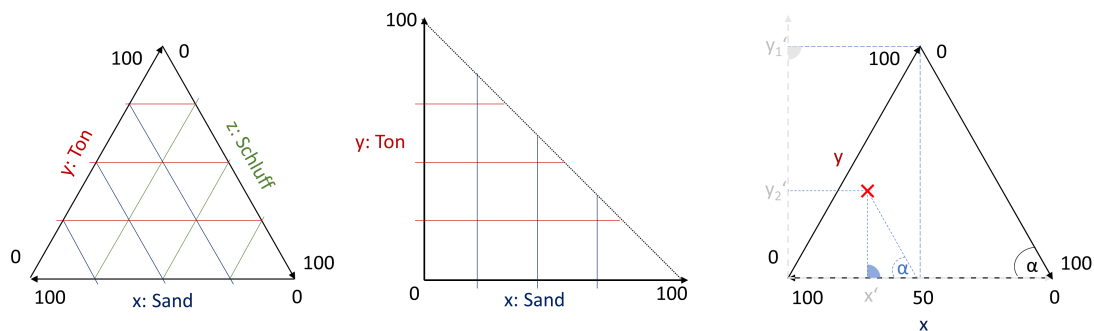


Abb. 32: Schaubild zur Herleitung der Koordinaten für das Texturdreieck  
 links: gleichseitiges Zieldreieck mit Sand-, Ton- und Schluffachse  
 Mitte: gleichschenkliges Dreieck mit Sand- und Tonachse  
 rechts: geometrische Herleitung der Umrechnungsfaktoren der Koordinaten  
 (x in der rechten Teilabbildung entspricht einem Beispielboden und dient zur Veranschaulichung der Herleitung seiner Koordinaten durch die Formeln 5.1-5.2.)

Da die y-Achse im gleichseitigen Dreieck um  $60^\circ$  geneigt ist, gleichzeitig aber der Länge 100 entsprechen soll, wird der Umrechnungsfaktor für die vertikale  $y'$ -Achse wie folgt definiert:

$$y' = \sqrt{y^2 - (100 - x'(y))^2} \quad (5.1a)$$

mit  $y = 100$  folgt:

$$x'(100) = 50 \text{ und damit:}$$

$$y' = \sqrt{100^2 - 50^2} = 50 \cdot \sqrt{3} \quad (5.1b)$$

und da gilt:

$$\frac{y}{100} = \frac{y'}{50 \cdot \sqrt{3}} \quad (5.1c)$$

oder auch:

$$y' = \cos(30) \cdot y \quad (5.1d)$$

resultiert:

$$y' = \frac{\sqrt{3}}{2} \cdot y \quad (5.1e)$$

mit:

$y'$  = neue Tonkoordinate

$y$  = tatsächlicher Tonanteil

Dabei wird der Satz des Pythagoras ausgenutzt, da in dem Hilfsdreieck (vgl. rechte Teilabbildung in Abb. 32) zwischen der  $y'$ - und der  $x$ -Achse ein rechter Winkel vorliegt und die Spitze des Zieldreiecks bei  $x = 50$  liegen soll. Außerdem ist wegen des Strahlensatzes das Verhältnis von  $y'$ - zu  $y$ -Achse immer gleich. Um die Werte für die neue  $x'$ -Achse, die zwar wie die ursprüngliche  $x$ -Achse senkrecht auf der  $y'$ -Achse steht, aber deren Ablesen im  $60^\circ$ -Winkel und nicht horizontal erfolgen soll, zu berechnen, wird folgende Formel verwendet:

$$(x' - x) = \frac{y'}{\tan(60)} \quad (5.2a)$$

durch Umstellen ergibt sich:

$$x' = x + \frac{y'}{\sqrt{3}} \quad (5.2b)$$

oder trigonometrisch formuliert:

$$x' = x + \sin(30) \cdot y' \quad (5.2c)$$

damit resultiert wegen der gespiegelten Achse und mit  $y'$  aus Formel 5.1:

$$x' = (-1) \cdot \left(x + \frac{y}{2}\right) \quad (5.2d)$$

Für einen beliebigen Punkt kann immer ein rechtwinkliges Dreieck gebildet werden, sodass der Punkt als Schnittpunkt von Hypothenuse und Gegenkathete des Winkels  $\alpha$  gilt. Mittels des Satzes von Pythagoras und mittels einer der beiden trigonometrischen Funktionen Tangens oder Sinus kann  $x'$  in Abhängigkeit von  $y'$  berechnet werden, weil die Gegenkathete  $y'$  entspricht

und  $\alpha$  im Hilfsdreieck als Stufenwinkel ebenfalls  $60^\circ$  beträgt (vgl. Abb. 32). Da der Sandanteil von rechts nach links zunehmen soll und nicht wie in einem numerischen Koordinatensystem von links nach rechts, wird die  $x'$ -Achse mit dem Faktor  $-1$  multipliziert, die Beschriftung dieses Anteils bleibt jedoch positiv (vgl. Formeln 5.2).

Alle Inter- und Extrapolationen, die zur Erzeugung der eigenen PTFs verwendet werden, wurden auf Basis der auf 2D konvertierten Sand- und Tonanteile durchgeführt. Das fertige Triangle mit der dazugehörigen USDA-Unterteilung in die 12 Texturklassen und deren Bezeichnung ist in Abbildung 33 dargestellt.

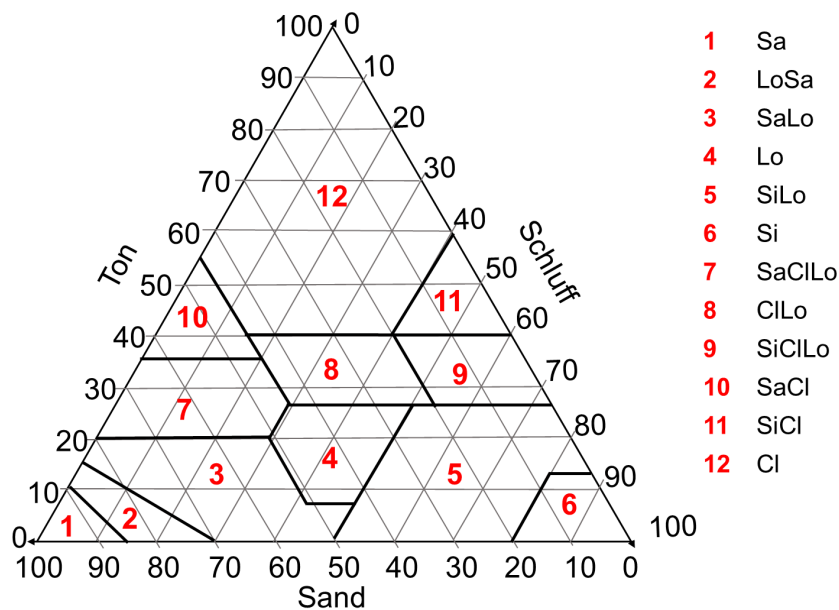


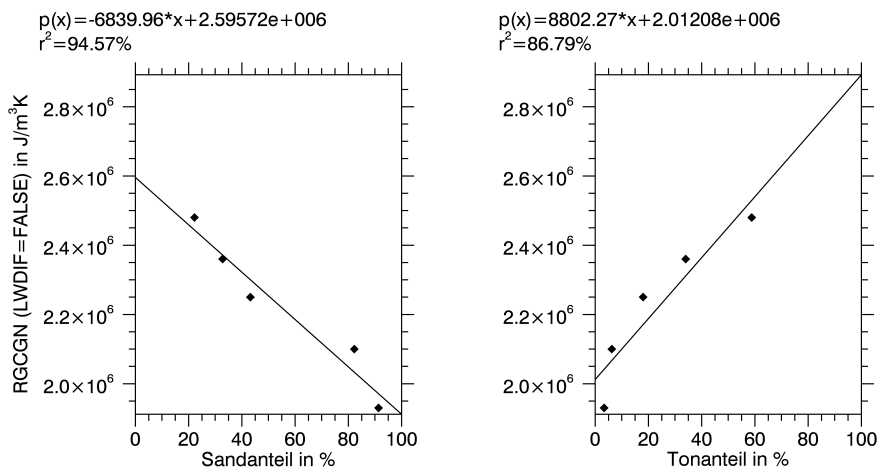
Abb. 33: Textur-Triangle mit den dazugehörigen 12 USDA-Texturklassen (Die jeweiligen Achsenangaben kennzeichnen den Anteil eines Bodens an Ton, Sand und Schluff in %.)

Durch die Zuweisung der neuen Texturklassen zu den bisherigen Texturklassen anhand ihrer Klassenbezeichnung und der Ermittlung der entsprechenden Sand- und Tonanteile des jeweiligen Texturklassenmittelpunktes kann eine numerische Approximation der Bodenparameterwerte für die neuen Texturklassen vorgenommen werden. In dieser Arbeit wurde sowohl die IDW als auch eine bilineare Extrapolation (mit Gewichtung) getestet. Für die IDW wurden alle fünf Stützwerte mit in die Berechnung einbezogen. Der Nachteil dieser IDW-Methode ist, dass sich die Klassenmittelpunkte der bisherigen Klassen eher im mittleren Wertebereich der Sand- und Ton-Achse befinden und somit die Interpolation auch nur in diesem Bereich zuverlässige Werte für die neuen Texturklassen liefert. Dass die bisherigen Parameter von Sand- und Tonanteil abhängig sind bzw. sich durch diese zu einem gewissen Teil herleiten lassen, zeigt Abbildung 34. In dieser Abbildung sind alle sieben in REMO verwendeten Bodenparameter (vgl. Tabelle in Anhang G) für jede der fünf Texturklassen und dem jeweils einer Klasse zugewiesenen mittleren Sand- bzw. Tonanteil aufgetragen. Es ist zu erkennen, dass die Punkte bei allen Parametern sehr nahe an den beiden jeweils konstruierten linearen Regressionsgeraden  $p(x)$  (Methodenbeschreibung in Kapitel 6.1.4) liegen. Dies spiegelt sich auch in den sehr hohen Determinationskoeffizienten  $r^2$  wider. Sie liegt bei den meisten Gleichungen über 80% und bei acht

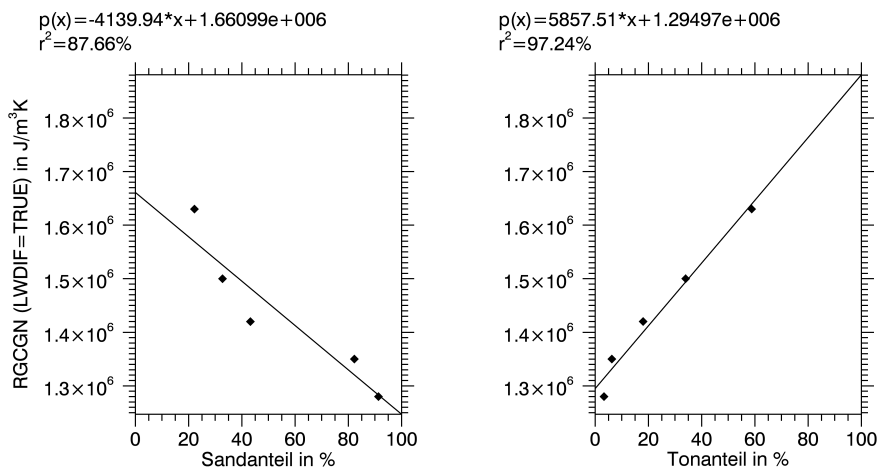


der 14 Gleichungen sogar über 90%. Ausnahmen sind bei der Amplitude der Wärmeleitfähigkeit (DLAMBDA) für die Tonanteil-Gleichung und bei FCAP zu erkennen. Die Aussage der hohen Koeffizienten wird jedoch durch den extrem kleinen Stichprobenumfang relativiert. Nur bei DLAMBDA scheint eine Aufteilung der gegebenen Werte in zwei Gruppen und die anschließende lineare Approximation ein noch besseres Annäherungsergebnis liefern zu können. Dies wurde allerdings aufgrund des Stichprobenumfangs und der Gleichbehandlung aller Parameter verworfen. Diese auf die dargestellte Art aber mit den mit Formel 5.1 und 5.2 konvertierten Sand- und Tonanteil erstellten Gleichungen und Korrelationskoeffizienten der jeweiligen Geraden bildeten die Grundlage für die bilineare Extrapolation mit Gewichtung, die in Kapitel 4.1.1 beschrieben ist. Durch die Kombination beider Geradengleichungen ist es möglich, auch für die Mischtexturklassen Parameterwerte zu erzeugen, die nicht so stark wie die jeweiligen Werte der IDW-Methode streuen. Daher sind die Ergebnisse dieser bilinearen Extrapolation als realistischer einzustufen und wurden für die 12 Texturklassen ins Modell eingebaut.

(A)

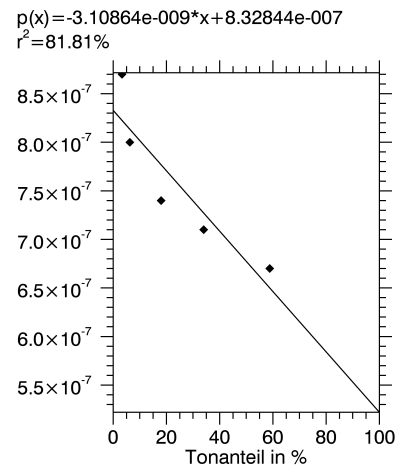
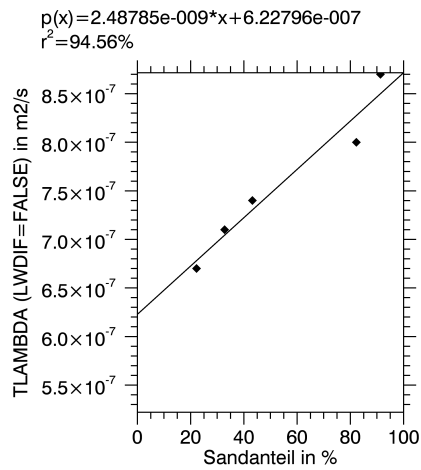


(B)

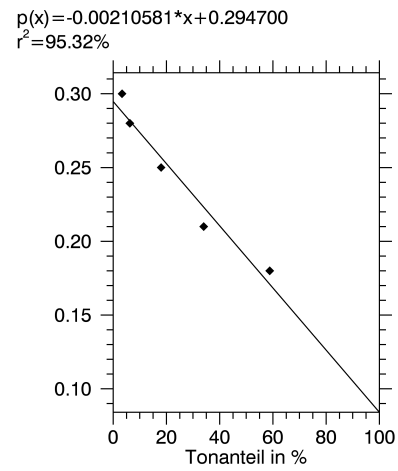
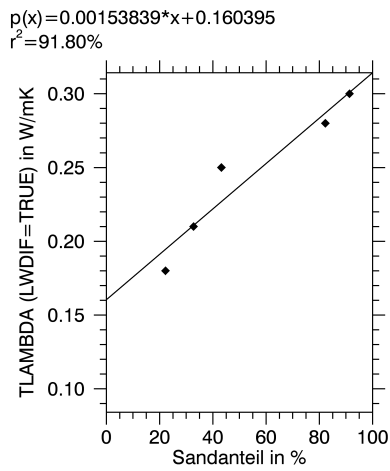


5 Implementierung neuer Parametrisierungen und Prozesse sowie daraus resultierende Veränderungen im Programmcode für die Bodenparameter

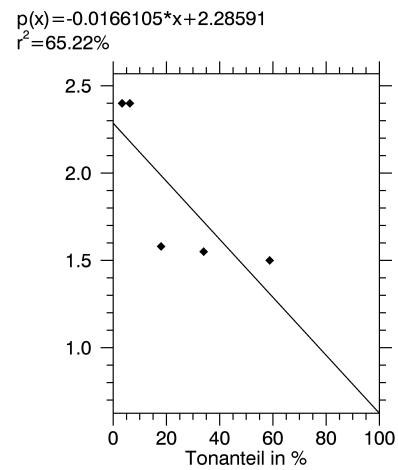
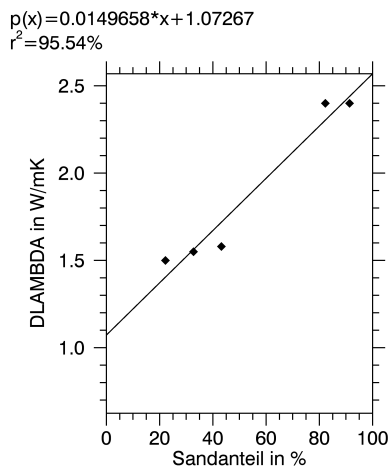
(C)



(D)



(E)



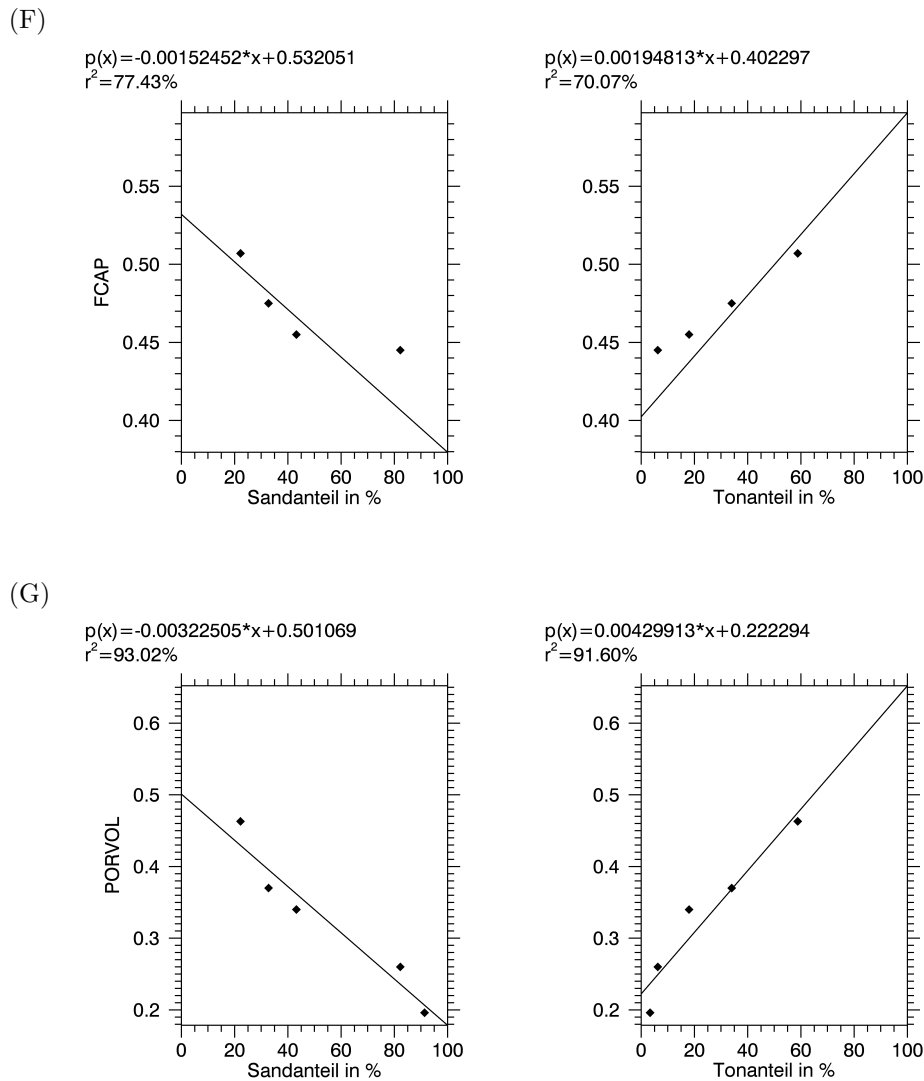


Abb. 34: Analyse der in REMO verwendeten sieben Bodenparameter abhängig vom jeweils zugewiesenen Texturklassenmittelwert für Sand- und Tonanteil

(A) volumetrische Wärmekapazität für einen Boden mit mittleren Wassergehalt (LWDIF=FALSE)

(B) volumetrische Wärmekapazität für einen trockenen Boden (LWDIF=TRUE)

(C) thermische Diffusivität (LWDIF=FALSE)

(D) Wärmeleitfähigkeit für einen trockenen Boden (LWDIF=TRUE)

(E) Amplitude der Wärmeleitfähigkeit

(F) Feldkapazität

(G) Porenvolumen

links: abhängig vom Sandanteil

rechts: abhängig vom Tonanteil

### **Ableitung einer Torfklasse**

Da vor allem der Anteil des organischen Materials einen großen Einfluss auf die benötigten Bodenparameter hat und alle neu implementierten Texturdatensätze keine separate Ausweisung einer „Torf-Klasse“ besitzen, wurde auf Basis des HWSD ein „Torflayer“ erstellt. Der HWSD stellt eine globale Karte zum prozentualen Anteil an organischem Material zur Verfügung. Diese wurde so reklassifiziert, dass alle Pixel mit einem organischen Anteil größer 30% als Torf identifiziert werden, da dieser Grenzwert die besten gebietsmäßigen Übereinstimmungen mit der bisherigen Torfklasse des FAO-Datensatzes lieferte. Diese neue Torfkarte überschreibt die jeweilige Texturklasse in den drei Ausgangsdatensätzen, bei einer entsprechenden durch den Benutzer festgelegten Einstellung in der Bodenbibliothek.

### **3. Ansatz: Regression mit kontinuierlichen Sand- und Tonanteilen**

Um das Potential der neuen Daten noch besser zu nutzen und die Ausgangslage der Bodenmodellierung im Modell weiter zu verbessern, können nach dem Aufstellen der bilinearen Extrapolationsgleichung für die benötigten Parameter auch die tatsächlichen prozentualen Anteile der Korngrößen statt der mittleren Anteile pro Texturklasse verwendet werden. In allen drei neuen Datensätzen (FAOn, Soilgrid und HWSD) sind Werte zum prozentualen Anteil von Sand, Ton und Schluff pro Pixel enthalten. Statt der Texturklassenmittelwerte von Sand und Ton (2. Ansatz), können mit Hilfe der Datensätze auch genauere Angaben zur jeweiligen Korngrößenverteilung gemacht werden. Dadurch erhält man einen kontinuierlichen Wertebereich für die einzelnen Parameter statt wie bisher nur fünf bzw. bei der neuen Methode 12 diskrete Abstufungen der Eigenschaften.

### **Vergleich der drei Ansätze**

Die Auswirkung auf die Werte dieser drei verschiedenen Ansätze (Texturklassenzuweisung, eigene PTF für Texturklassenmittelwerte, eigene PTF mit kontinuierlichen Sand-/Tonanteil) zeigt Abbildung 35 exemplarisch für die Parameter volumetrische Wärmekapazität (RGCGN), Amplitude der Wärmeleitfähigkeit (DLAMBDA) und Porenvolumen (PORVOL). Die in der oberen Reihe der Abbildung dargestellten Werte für die drei Bodenparameter zeigen klar auf, dass teilweise große Sprünge zwischen benachbarten Klassen auftreten. Besonders stark sieht man eine solche Differenz bei der Amplitude der Wärmeleitfähigkeit (vgl. Abb. 35B). Die beiden Werte für Sand und sandigen Lehm ähneln sich stark und definieren die obere Grenze des Wertebereichs. Während die anderen drei gegebenen Werte für Lehm, toniger Lehm und Ton ebenfalls sehr ähnlich sind, sich aber stark von den sandigen Werten unterscheiden.

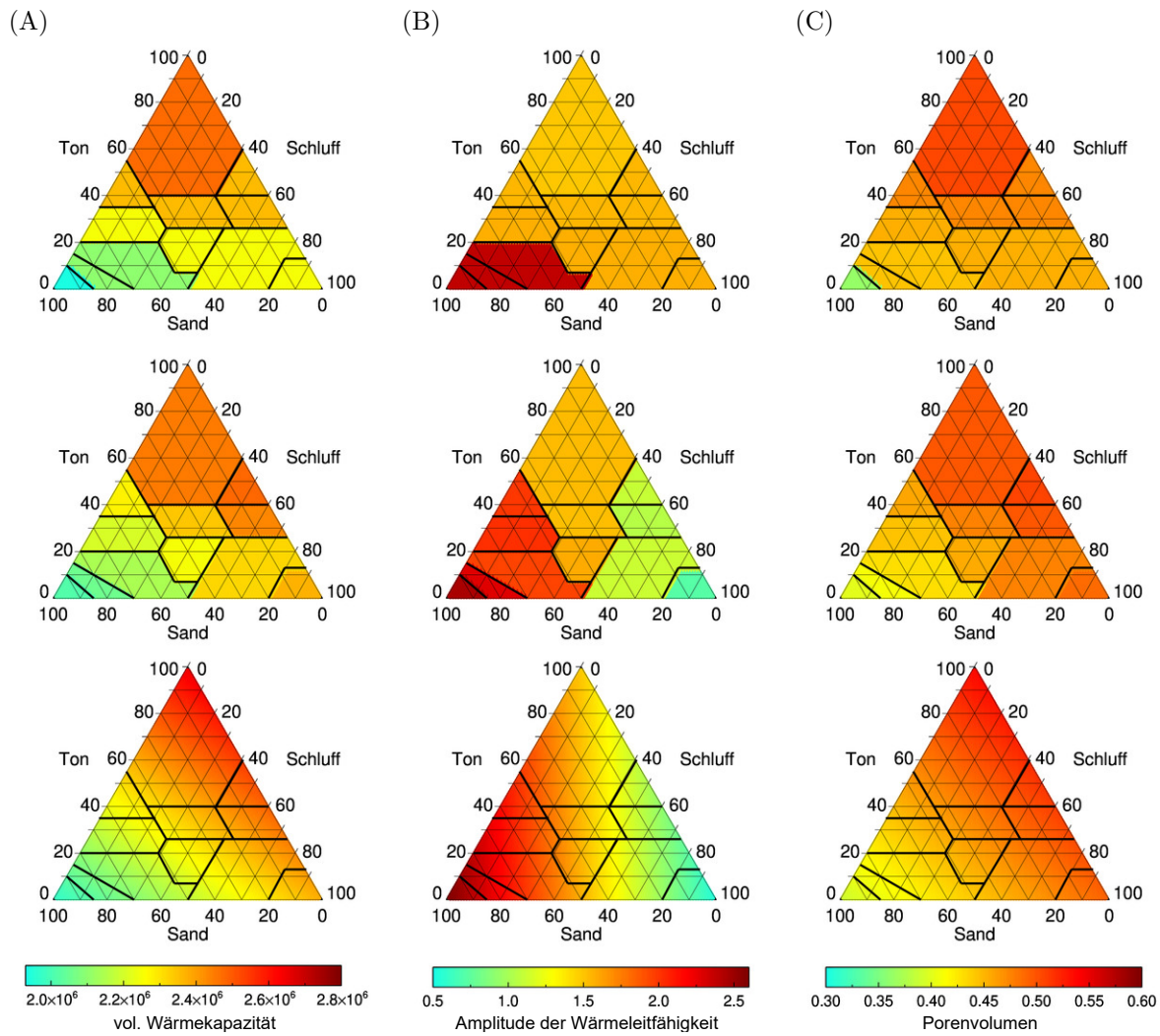


Abb. 35: Vergleich der Werte der Bodenparameter anhand des USDA-Triangles  
 oben: bisherige Werte aus dem EM/DM Manual zugewiesen anhand des Klassennamen  
 Mitte: bilineare Approximation anhand des Sand- und Tonanteils der 12 Klassenmittelpunkte (LSOIL=T)  
 unten: bilineare Approximation für alle Sand- und Tonanteile (LSOIL=F)  
 (A) volumetrische Wärmekapazität für einen mittleren Wassergehalt (in  $J/m^3K$ )  
 (B) Amplitude der Wärmeleitfähigkeit (in  $W/mK$ )  
 (C) Porenvolumen (in  $m^3/m^3$ )

Ein weiterer Nachteil für die Zuweisung über die Namen ist das Fehlen eines Stützwertes für Schluff. Die gesamte „Schluff-Achse“ ist in den gegebenen fünf Texturwerten nicht repräsentiert. Durch die in der Mitte und unten dargestellten Werte der bilinearen Approximation ist allerdings zu erkennen, dass diese Achse je nach Parameter einen Anteil an der Ausprägung des Bodenparameters hat. Während beim Porenvolumen der Schluffanteil kaum einen Einfluss auf den Parameterwert zu haben scheint, unterscheiden sich sowohl bei RGCGN als auch bei DLAMBDA die Werte für hohe bzw. niedrige Schluffanteile stark.

Der Vorteil der kontinuierlichen Parameter (Abb. 35 unten) durch die Berechnung mittels des Sand- und Tonanteils im Vergleich zu den diskreten Werten durch die Texturklassenmittelpunkte zeigt sich vor allem bei den Texturklassen, die eine größere Fläche im Dreieck abbilden. Besonders in der flächenmäßig größten Texturklasse 12 (Clay) reicht der Wertebereich des Tons von 60-100% und die Parameterwerte umfassen dementsprechend fast die Hälfte des jeweiligen Wertespektrums bei der kontinuierlichen Methode, während bei den anderen beiden Methoden nur ein Wert in der Mitte des Spektrums für die Klasse angenommen wird.

### **Vergleich der PTFs aus der Literatur**

In einem weiteren Schritt kann nach dem Einbau des Sand-, Ton- und organischen Anteils in REMO auch auf bekannte kontinuierliche PTFs für die benötigten Parameter (RGCGN, TLAMBDA, FCAP und PORVOL) aus der Literatur oder aus anderen LSMs oder RCMs zurückgegriffen werden. Vor allem für FCAP und PORVOL finden sich in der Literatur viele verschiedene PTFs. Neben FCAP und PORVOL werden in REMO noch weitere Bodenparameterwerte speziell für die hydrologische Bodenmodellierung benötigt, welche jedoch bisher von FCAP abhängen. Der Wert für den PWP wird auf 35% des Feldkapazitätswertes pro Pixel und der für den residualen Wassergehalt  $\theta_r$  auf 15% der Feldkapazität gesetzt. In vielen anderen Studien werden jedoch jeweils verschiedene PTFs für FCAP und PWP angegeben. Eine Auswahl an Entwicklungs- und Vergleichsstudien von PTFs für FCAP und PWP liefern die Tabellen in Anhang H und I. Sehr wichtig bei der Auswahl und Verwendung von PTFs für FCAP und PWP ist das Beachten der Einheiten der Eingangs-, aber auch der Zielvariablen. Wie bereits in Kapitel 2 beschrieben, kann der jeweilige Wassergehalt entweder gravimetrisch oder volumetrisch angegeben werden. Eine Umrechnung ist mit Formel 2.10 möglich, doch die als Umrechnungsfaktor notwendige Trockenrohddichte steht nicht immer als Messwert zur Verfügung. Teilweise wird ein konstanter Wert für die Trockenrohddichte des Bodens verwendet, was jedoch zu Ungenauigkeiten führt, da auch die Trockenrohddichte stark von der Bodenzusammensetzung abhängt. In REMO wird die volumetrische Form benötigt. Zudem sollen keine weiteren Variablen eingeführt werden, sodass PTFs, die beispielsweise die Kationenaustauschkapazität (z. B. Dharumarajan et al., 2019) oder Werte zum geometrischen Durchmesser der Bodenteilchen (z. B. Ghanbarian & Millán, 2010 oder Ostovari et al., 2015) benötigen, ebenfalls nicht berücksichtigt werden.

Eine geeignete PTF zu finden, wird außerdem dadurch erschwert, dass die einzelnen PTFs meist nur für kleine Stichproben aus kleinräumigen Gebieten und somit nur für spezielle Bodenzusammensetzungen entwickelt und getestet werden. Teilweise werden innerhalb der Studien noch Aufteilungen der Stichproben vorgenommen, sodass verschiedene PTFs für unterschiedliche Bedingungen erstellt werden. Dieser Umstand macht den Vergleich und die Kombination der Erkenntnisse auf der Grundlage verschiedener Studien schwierig. Beispielsweise wird Arruda et al. (1987) in unterschiedlichen Quellen mit verschiedenen Formeln (vgl. ID 3, 4 und 5 in Anhang H) zitiert. Ferner ist anhand der in den beiden Tabellen aufgelisteten Studien zu erkennen, dass es unterschiedliche numerische Ansätze gibt. Der Großteil der ausgewählten Studien bildet einfache lineare Gleichungen ab, die mittels einer einfachen oder multiplen linearen Regression aus Stichproben erzeugt werden. Aber auch exponentielle Zusammenhänge (vgl. ID 4) und

gebrochen-rationale (vgl. ID 8) oder höher dimensionale Methoden (vgl. ID 3 und ID 4 für FCAP) werden für die Erstellung der PTFs verwendet. Des Weiteren spielt die räumliche Skala der Daten, auf deren Basis die PTF entwickelt wurde, eine entscheidende Rolle und beeinflusst die resultierenden Koeffizienten in der PTF (Vereecken et al., 2016).

Zusätzlich ist der Vergleich von PTFs verschiedener Studien wegen der unterschiedlichen Definitionen von FCAP und PWP erschwert. Zumeist wird ein Matrixpotential von  $-33$  bzw.  $-1500$  kPa verwendet (vgl. Viterbo, 2002). Jedoch gibt z. B. Pidgeon (1972) neben der Formel für den Wassergehalt bei  $-33$  kPa Matrixpotential noch eine von diesem Wert abhängige Formel für die Feldkapazität an. Die in Anhang H und I aufgelisteten Funktionen zeigen des Weiteren, dass sich die Koeffizienten verschiedener PTFs deutlich unterscheiden. Für den Einsatz in REMO wird in dieser Arbeit zunächst eine Art Ensemble-Analyse der FCAP und PWP-Werte für verschiedene Sand-, Ton- und organische Anteile vorgenommen. Anschließend wird die PTF ausgewählt, die bei allen Bodenzusammensetzungen dem Ensembledurchschnitt am ähnlichsten ist. Da die gravimetrischen Werte beim Vergleich trotz des Einsatzes eines mittleren Wertes für die Trockenrohddichte als Umrechnungsfaktor von  $1.4 \frac{\text{g}}{\text{cm}^3}$  (vgl. Hillel, 2003) deutlich von den volumetrischen Werten abweichen, scheiden diese für den Einsatz in REMO aus. Außerdem ist auch die PTF von Pidgeon (1972) ungeeignet, da nur eine Formel für FCAP, nicht aber für PWP in der Studie aufgestellt wird. Die PTFs von Adhikary et al. (2008) (ID 1), die auf Grundlage einer Stichprobe aus Indischen Böden entwickelt wurden, sind die resultierenden Funktionen und werden für die Modellläufe id10 und id11 (vgl. Experimentplan 3, Abb. 46) in REMO getestet. Diese PTFs schnitten auch in den Vergleichsstudien von Rustanto et al. (2017), Santra et al. (2018) und Qiao et al. (2019) sowohl für FCAP als auch für PWP gut ab. Nachteil dieser PTF ist jedoch die Nichtberücksichtigung des organischen Anteils, von dem bekannt ist, dass er die Wasserhaltefähigkeit vor allem in sandigen Böden erhöht (vgl. Rawls et al., 2003).

Auch für das Porenvolumen, welches ebenfalls als spezifischer Wassergehalt aufgefasst werden kann (vgl. Definition in Kapitel 2.2), existieren verschiedene PTFs. Weit verbreitet ist die Annäherung von Cosby et al. (1984) (vgl. Dai et al., 2013; van Looy et al., 2017, Bonan, 2019). Diese ist nur vom Sandanteil abhängig und wird auch im CLM verwendet. Da das Porenvolumen von organischem Material jedoch um ein Vielfaches höher ist als das von mineralischen Böden, erweitern Lawrence & Slater (2008) die Cosby-PTF durch eine Gewichtung mit dem organischen Anteil (vgl. Formel 5.3). Diese Erweiterung wird in dieser Arbeit in den Testläufen id09, id10 und id11 (vgl. Abb. 46) verwendet.

$$PORVOL = (1 - f_{org}) \cdot (0.489 - 0.00126 \cdot SAND) + f_{org} \cdot 0.9 \quad (5.3)$$

mit:

$f_{org}$  = organischer Anteil

$SAND$  = Sandanteil in %

PORVOL (entspricht dem gesättigten Wassergehalt  $\theta_s$ ), FCAP, PWP und  $\theta_r$  sind Kennziffern der WRC. Die oben genannten PTFs bilden jedoch jeweils nur einzelne Werte (sogenannte

„pressure heads“) auf dieser Kurve ab. Laut van Looy et al. (2017) sollten diese pressure heads jedoch zueinander passen, um eine in sich konsistente Bodenwassermodellierung zu erhalten. Die Verwendung von verschiedenen PTFs für die einzelnen Kennwerte ist daher zu vermeiden.

Diese WRC werden ebenfalls mit verschiedenen Ansätzen parametrisiert. Die bekanntesten Funktionen sind Brooks & Corey (1964), Campbell (1974) und van Genuchten (1980). Von diesen drei Ansätzen gibt es zudem verschiedene Abwandlungen. Beispielsweise ist die ebenfalls häufig zitierte Clapp-Hornberger-Methode (nach Clapp & Hornberger (1978)) eine Vereinfachung der Campbell-Kurve. Die Clapp-Hornberger-Parametrisierung wurde nach dem Ansatz von Hagemann & Stacke (2015) bereits in REMO für die hydraulische Konduktivität und Diffusivität getestet. Weitere Informationen und ein Vergleich der Modellansätze sind z. B. in F. J. Braun (2002) und Dai et al. (2013) zu finden. Neben der Modellperformanzeinschätzung ist speziell im Hinblick auf den Einsatz im Klimamodell zu beachten, dass z. B. Shao & Irannejad (1999) in ihrer Studie feststellten, dass der Clapp-Hornberger- und der Brooks-Corey-Ansatz dem van Genuchten-Ansatz bei fast gleicher Genauigkeit numerisch deutlich überlegen ist (fast 6x schneller). Eine Übersicht, welche Parametrisierung in den verschiedenen LSMs und hydrologischen Modellen eingesetzt werden, liefern unter anderem F. J. Braun & Schädler (2005) und van Looy et al. (2017).

Auch diese die WRC abbildenden Funktionen benötigen von der jeweiligen Bodenzusammensetzung abhängige Kurvenparameter (wie z. B. den residualen und gesättigten Wassergehalt oder andere Kurvenparameter). Diese Parameter sind in verschiedenen Studien anhand von Bodenstichproben optimiert worden und können dementsprechend ebenfalls als PTF gelten. Während Vereecken et al. (2010) verschiedenste Parametrisierungen der van Genuchten-Methode miteinander vergleichen und deren Performanz diskutieren, liefern Dai et al. (2013) eine Übersicht über fünf verschiedene Clapp-Hornberger-Parametrisierungen. Neben dem Vergleich der Methoden werden auch mittels verschiedener PTFs generierte Parameterkarten als Datensätze veröffentlicht (z. B. Montzka et al., 2017; Dai, Wei et al., 2019 oder Y. Zhang et al., 2020) und verschiedene PTFs werden als Software Pakete, wie z. B. „Rosetta“ von Schaap et al. (2001) (bzw. neuere Version Y. Zhang & Schaap, 2017), oder SOILPAR von Acutis & Donatelli (2003) zur Verfügung gestellt.

Die in der Studie von Dai enthaltene und auch z. B. im globalen Ökosystemmodell Lund-Potsdam-Jena managed Land (LPJmL) (vgl. Lutz et al., 2019) verwendete Parametrisierung nach Saxton & Rawls (2006) wird in dieser Arbeit als Alternative für die punktbasierten PTFs (ab Modellversion id12 im Experimentplan 3) getestet. Ein weiterer Vorteil dieser neuen Parametrisierung nach Saxton & Rawls (2006) für die pressure heads ist, dass sie zur bereits in REMO enthaltenen hydrologischen Modellierung der hydraulischen Konduktivität und Diffusivität nach Clapp-Hornberger „passt“. Diese hydrologische Modellierung wird jedoch erst bei aktiviertem L5LAY-Schalter verwendet, welcher in dieser Arbeit jedoch immer ausgeschaltet ist. Die Werte für die pressure heads werden hingegen auch für das in dieser Arbeit verwendete Bucket-Schema benötigt. Es ergeben sich folgende Formeln für PWP (5.4), FCAP (5.5) und PORVOL (5.6):



$$\begin{aligned} PWP &= \theta_{1500t} + (0.14 \cdot \theta_{1500t} - 0.02) \\ &= 1.14 \cdot \theta_{1500t} - 0.02 \end{aligned}$$

mit:

$$\begin{aligned} \theta_{1500t} &= -0.024 \cdot SAND + 0.487 \cdot CLAY + 0.006 \cdot SOM + \\ &0.005 \cdot (SAND \cdot SOM) - 0.013 \cdot (CLAY \cdot SOM) \\ &+ 0.068 \cdot (SAND \cdot CLAY) + 0.031 \end{aligned} \quad (5.4)$$

$$\begin{aligned} FCAP &= \theta_{33t} + (1.283 \cdot \theta_{33t}^2 - 0.374 \cdot \theta_{33t} - 0.015) \\ &= 1.283 \cdot \theta_{33t}^2 + 0.626 \cdot \theta_{33t} - 0.015 \end{aligned}$$

mit:

$$\begin{aligned} \theta_{33t} &= -0.251 \cdot SAND + 0.195 \cdot CLAY + 0.011 \cdot SOM + \\ &0.006 \cdot (SAND \cdot SOM) - 0.027 \cdot (CLAY \cdot SOM) + \\ &0.452 \cdot (SAND \cdot CLAY) + 0.299 \end{aligned} \quad (5.5)$$

$$\begin{aligned} PORVOL &= FCAP + (1.636 \cdot \theta_{(sat-33)t} - 0.107) - 0.097 \cdot SAND + 0.043 \\ &= FCAP + 1.636 \cdot \theta_{(sat-33)t} - 0.097 \cdot SAND - 0.064 \end{aligned}$$

mit:

$$\begin{aligned} \theta_{(sat-33)t} &= 0.278 \cdot SAND + 0.034 \cdot CLAY + 0.022 \cdot SOM \\ &- 0.018 \cdot (SAND \cdot SOM) - 0.027 \cdot (CLAY \cdot SOM) \\ &- 0.584 \cdot (SAND \cdot CLAY) + 0.078 \end{aligned} \quad (5.6)$$

mit:

$PWP$  = permanenter Welkepunkt in  $\frac{m^3}{m^3}$

$FCAP$  = Feldkapazität in  $\frac{m^3}{m^3}$

$PORVOL$  = Porenvolumen in  $\frac{m^3}{m^3}$

$SAND$  = Sandanteil in %

$CLAY$  = Tonanteil in %

$SOM$  = organischer Anteil in %

Die Studie von Alaya et al. (2017) für Böden in Tunesien bescheinigt den originalen Saxton & Rawls-PTFs für FCAP und PWP gute Ergebnisse (je  $r^2 = 72\%$ ) beim Vergleich mit Beobachtungsdaten. Auch Sung & Iba (2010) stellen fest, dass besonders FCAP und PWP durch die Funktionen mit einem sehr kleinen Bias geschätzt werden kann. PORVOL dagegen wird in der Stichprobe häufig unterschätzt. Sie erreichen jedoch durch eine Neukalibrierung der PTFs noch bessere Ergebnisse für Böden in Malaysia. Diese Neukalibrierung stellt allerdings eine Anpassung an die verwendeten Stichproben dar und eignet sich somit nicht zum Einsatz in einem regionalen Klimamodell, das weltweit eingesetzt wird.

Für die thermalen Parameter RGCGN, TLAMBDA und DLAMBDA und der speziellen Aufteilung der Werte nach trockenem Boden und Boden mit mittlerem Wassergehalt, die für die initialen in REMO enthalten Formeln 3.2 und 3.3 benötigt werden, finden sich sehr viel weniger PTFs. In Kapitel 5.1.3 werden deshalb andere Methoden zur Ermittlung der volumetrischen Wärmekapazität und der thermischen Diffusivität des Bodens vorgestellt. Eine Übersicht der in den Testläufen dieser Arbeit verwendeten Kombinationen der Bodenparameter und deren jeweilige Auswirkung auf abgeleitete Parameter und auf die Modellperformanz wird in Kapitel 6.2 und 6.4 gezeigt.

### 5.1.2 Ausweitung der Bodenparameter in die Tiefe

Durch die Verwendung der neuen Bodendatensätze, die für mehrere Bodenschichten Daten bereitstellen (vgl. Kapitel 4.3.2), müssen die Bodenparameter in REMO nicht mehr als konstant über alle fünf Modellbodenschichten angenommen werden. Die je nach Datensatz unterschiedlichen Schichten wurden mittels der in Kapitel 4.1.1 dafür definierten Interpolationen für die Variablen Sand-, Ton- und organischen Anteil auf die fünf REMO-Schichten (vgl. Abb. 10) vertikal angepasst. Beim Ausführen von REMO kann der Anwender in der Namelist durch den neuen Schalter L5LAYER\_TOTAL entscheiden, ob wie bisher nur mit einem über alle fünf Schichten konstanten Wert (FALSE) oder mit den neuen variablen Werten (TRUE) gerechnet werden soll. Abbildung 36 zeigt exemplarisch den Mittelwert pro Schicht über alle Landpunkte innerhalb des EUR-44-Gebiets für Sand- und Tonanteil sowie für den organischen Anteil für den HWSD- und den SoilGrids-Datensatz. Zum einen sind die Werte für die originalen Tiefen angegeben und zum anderen die interpolierten Werte für die fünf REMO-Schichten. Für HWSD stehen nur zwei Stützwerte für die Interpolation zur Verfügung, während bei Soilgrid sieben Werte verwendet werden können.

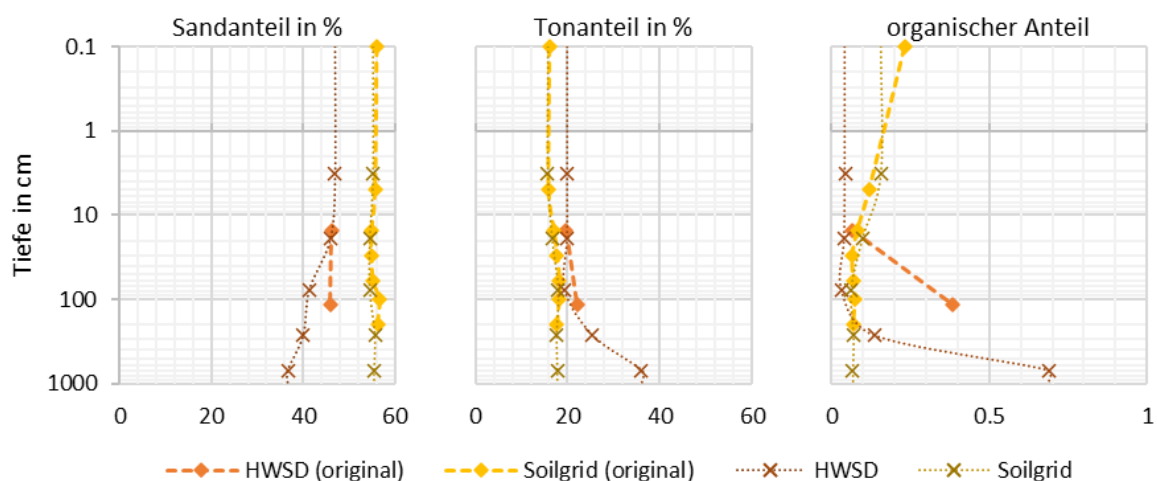


Abb. 36: Mittelwert über die Landpunkte des EUR-44-Gebiets für die Variablen Sand-, Ton- und organischen Anteil für die originalen Tiefen sowie für die REMO-Schichten für HWSD- und SoilGrids-Datensatz

Die Soilgrid-Werte von Sand- und Tonanteil für die obersten drei REMO-Schichten liegen auf den Splines der originalen Werte. Die Werte für REMO-Schicht 4 und 5 unterscheiden sich für alle drei Variablen kaum vom Wert der dritten Schicht. Bei den Werten des HWSD weichen diese beiden Schichtwerte sehr deutlich von den oberen Werten ab, wobei die HWSD-Werte insgesamt im Vergleich zu Soilgrid einen jeweils größeren Wertebereich haben. Besonders stark treten diese Effekte beim organischen Anteil auf, weil dieser Anteil, im Gegensatz zur linearen Interpolationsmethode für Sand- und Tonanteil, exponentiell von der Tiefe abhängt. Folglich nimmt die Differenz zwischen HWSD und Soilgrid mit steigender Tiefe zu. Es fällt auf, dass der HWSD für alle Schichten einen geringeren Sandanteil und dafür einen höheren Tonanteil ausweist. Beim organischen Anteil liefert der HWSD für die oberen drei Schichten kleinere Werte als Soilgrid, weil auch der originale Wert des HWSD für 0–30 cm niedriger ist als der Wert bei Soilgrid für 15 cm und 30 cm. Da der organische Wert des HWSD für 30–200 cm dagegen viel höher liegt (0.38 bei 115 cm statt 0.07 für Soilgrid bei 100 cm), wird durch die Interpolationsmethode eine untypische starke Zunahme des organischen Anteils mit der Tiefe angenommen. Wie in Kapitel 6.3.2 anhand der Karten für die oberste Schicht gezeigt wird, gilt dieser dargestellte Versatz zwischen HWSD und Soilgrid in allen drei Variablen nicht im gesamten Gebiet, sondern fällt lokal sehr unterschiedlich aus.

Aufgrund der großen Abweichung des organischen Anteils zwischen beiden Datensätzen in den unteren beiden Schichten, dem allgemeinen Verständnis folgend, dass in den tieferen Bodenschichten kaum organisches Material vorhanden ist und wegen des großen Einflusses von organischem Material auf die Wertebildung der Bodenparameter wurde neben einer Parametrisierung mit den berechneten organischen Anteilen für Schicht 4 und 5 auch eine Parametrisierung getestet, bei welcher der organische Anteil für diese beiden Schichten überall auf Null gesetzt wird (vgl. Modellläufe ab id08 im Experimentplan 3, Abb. 46). Der Unterschied, der sich durch diese Änderung für die Bodenparameter ergibt, wird in Kapitel 6.4 und die daraus folgende Auswirkung auf verschiedene Klimavariablen wird in Kapitel 6.5 gezeigt.

### 5.1.3 Verbesserung und alternative Formeln für die Wärmeeigenschaften des Bodens

#### Wärmeleitfähigkeit

Da für die Wärmeeigenschaften des Bodens, die zur Berechnung der Bodentemperaturänderungen benötigt werden (vgl. Formel 2.14), keine flächendeckenden Messdaten oder Beobachtungsdaten vorliegen, sind diese nur approximierbar. Hierfür existieren besonders für die Wärmeleitfähigkeit eine Vielzahl an Methoden sowie deren Abwandlungen (de Vries, 1952; de Vries, 1975; Johansen, 1975; Farouki, 1981b; Campbell, 1985; Lawrence & Slater, 2008; Y. Lu et al., 2014; u. a.) und viele Studien (z. B. S. X. Chen, 2008; Fröb, 2011; Haigh, 2012; Likos, 2015; J. Chen et al., 2018), die diese analysieren, in verschiedenen Modellen verwenden oder miteinander vergleichen. Eine gute Übersicht und einen Vergleich dieser Methoden, die in unterschiedlichen Modellen zum Einsatz kommen, liefern vor allem Dai, Wei et al. (2019), He et al. (2019), He et al. (2020) und J. Wang et al. (2020). Je nach Methode werden unterschiedliche Eingangsdaten benötigt. Um

REMO so einfach wie möglich zu halten und die benötigten externen Daten auf ein Minimum zu beschränken, werden nur Methoden benutzt, die Sand-, Ton- und organischen Anteil sowie Porenvolumen oder daraus ableitbare Variablen verwenden. Der teilweise in den Methoden zur Berechnung benötigte Quarzgehalt  $q$  kann durch den Sandanteil angenähert werden (vgl. Peters-Lidard et al., 1998), wenngleich einige Studien (z. B. S. Lu et al., 2007) darauf hinweisen, dass dies ein Fehler ist (vgl. F. J. Braun, 2002; Bonan, 2019).

In dieser Arbeit werden als Alternative zu der bisherigen Parametrisierung der volumetrischen Wärmekapazität und der Wärmeleitfähigkeit des Bodens (vgl. Formel 3.2 und 3.3) die folgenden Ansätze getestet: Johansen (1975), Farouki (1981b) und Lawrence & Slater (2008). Bei diesen drei Methoden wird auf die Wärmeleitfähigkeit des trockenen Bodens ein Term addiert (vgl. Formel 5.7). Dieser vom Wassergehalt abhängige Term wird durch die Amplitude der gesättigten ( $\lambda_{sat}$ ) und trockenen ( $\lambda_{dry}$ ) Wärmeleitfähigkeit skaliert und kennzeichnet den Anstieg der thermischen Konduktivität bei erhöhter Sättigung. Er wird als Kerstenzahl bzw. Kerstenfaktor  $K_e$  oder Kerstenfunktion (nach Kersten, 1949) bezeichnet. In den drei folgenden Ansätzen wird der Zusammenhang zwischen dem Grad der Sättigung und der Wärmeleitfähigkeit logarithmisch berechnet. Außerdem werden je nach Methode unterschiedliche Fälle abhängig vom Zustand des Bodens (gefroren/ungefroren) und der Sättigung ( $S_e = \frac{\theta}{\theta_{sat}}$ ) unterschieden. Wie in Tabelle 12 zu erkennen ist, unterscheiden sich je nach Methode sowohl die Berechnung des Werts für die Wärmeleitfähigkeit des trockenen Bodens ( $\lambda_{dry}$ ) als auch die Wärmeleitfähigkeit des Feststoffanteils des Bodens ( $\lambda_{sol}$ ), die zur Berechnung der gesättigten Wärmeleitfähigkeit benötigt wird. Bei Lawrence & Slater (2008) wird neben dem Quarzanteil (vgl. Johansen) bzw. dem Sand- und Tonanteil (vgl. Farouki) der organische Anteil im Boden sowohl bei  $\lambda_{sol}$  als auch bei  $\lambda_{dry}$  berücksichtigt. Im Hinblick auf die sich stark unterscheidenden Wärmeeigenschaften von organischem Material im Vergleich zu den anderen beteiligten Materialien (vgl. z. B. Tab. 2) ergeben sich dadurch starke Änderungen für organische Böden.

$$\lambda = (1 - K_e) \cdot \lambda_{dry} + K_e \cdot \lambda_{sat} = (\lambda_{sat} - \lambda_{dry}) \cdot K_e + \lambda_{dry} \quad (5.7)$$

Tab. 12: Verschiedene Berechnungsmethoden der Wärmeleitfähigkeit des Bodens im Vergleich

	Johansen (1975)	Farouki (1981b)	Lawrence & Slater (2008)
$K_e(\theta)$ für ungef. Boden	$\begin{cases} 0.7 \cdot \log_{10} \left( \frac{\theta}{\theta_{sat}} \right) + 1, & \text{wenn } \frac{\theta}{\theta_{sat}} > 0.05 \\ \log_{10} \left( \frac{\theta}{\theta_{sat}} \right) + 1, & \text{wenn } \frac{\theta}{\theta_{sat}} > 0.1 \end{cases}$	$\begin{cases} 0, & \text{wenn } \frac{\theta}{\theta_{sat}} \leq 10^{-7} \\ \log_{10} \left( \frac{\theta}{\theta_{sat}} \right) + 1, & \text{wenn } \frac{\theta}{\theta_{sat}} > 10^{-7} \end{cases}$	$\log_{10} \left( \frac{\theta}{\theta_{sat}} \right) + 1$
$K_e(\theta)$ für gefr. Boden		$\frac{\theta}{\theta_{sat}}$	
$\lambda_{dry}$		$\lambda_{m,dry}$	$(1 - f_{org}) \cdot \lambda_{m,dry} + f_{org} \cdot \lambda_{org,dry}$
$\lambda_{sat}$		$\lambda_{sol}^{1-\theta_{sat}} \cdot \lambda_w^{\theta_{sat} \cdot f_u} \cdot \lambda_i^{\theta_{sat} \cdot (1-f_u)}$	
$\lambda_{sol}$	$\lambda_q \cdot \lambda_o^{(1-q)}$	$\lambda_{m,sat}$	$(1 - f_{org}) \cdot \lambda_{m,sat} + f_{org} \cdot \lambda_{org,sat}$

mit:

$f_u$  = Wasseranteil des ungefrorenen Wassers im Boden

$f_{org}$  = Anteil an organischen Material im Boden

$\lambda_w = 0.57 \frac{W}{mK}$  = Wärmeleitfähigkeit von Wasser

$\lambda_i = 2.29 \frac{W}{mK}$  = Wärmeleitfähigkeit von Eis

$\lambda_{org,dry} = 0.05 \frac{W}{mK}$  = Wärmeleitfähigkeit von organischem Material nach Farouki (1981b)

$\lambda_{org,sat} = 0.25 \frac{W}{mK}$  = Wärmeleitfähigkeit von organischem Material nach van Wijk (1963)

$\lambda_{m,dry} = \frac{0.135p_b + 64.7}{2700 - 0.947p_b}$  = Wärmeleitfähigkeit von trockenen Bodenmineralien

$\lambda_{sol}$  = Wärmeleitfähigkeit des Feststoffanteils im Boden

$q$  = Quarzanteil

$\lambda_q = 7.7 \frac{W}{mK}$  = Wärmeleitfähigkeit von Quarz

$$\lambda_o = \begin{cases} 2.0 \frac{W}{mK}, & \text{= Wärmeleitfähigkeit von Mineralien mit } q > 0.2 \\ 3.0 \frac{W}{mK}, & \text{= Wärmeleitfähigkeit von Mineralien mit } q \leq 0.2 \end{cases}$$

$\lambda_{m,sat} = \frac{8.80 \cdot SAND + 2.92 \cdot CLAY}{SAND + CLAY}$  = Wärmeleitfähigkeit der Bodenmineralien mit  $SAND$  = Sandanteil in % und  $TON$  = Tonanteil in %

$p_b = 2700 \cdot (1 - \theta_{sat})$  = Annäherung der Trockenrohdichte nach Bonan (2019) in  $\frac{kg}{m^3}$

Um den Einfluss der verschiedenen Anteile (Sand, Ton, Wasser, Eis, organisches Material) auf die effektive Wärmeleitfähigkeit zu demonstrieren und den Unterschied der drei ausgewählten Modelle auch hinsichtlich der bisher verwendeten Werte zu zeigen, sind in Abbildung 37 einige Fallbeispiele dargestellt. Abhängig vom jeweiligen Wassergehalt  $\theta$  werden die Wärmeleitfähigkeiten von verschiedenen Bodenzusammensetzungen für unterschiedliche Anteile an Eis und organischem Material gezeigt.

In Abbildung 37A werden die thermalen Leitfähigkeiten für einen ungefrorenen Boden ohne organisches Material für die drei beschriebenen neuen Methoden sowie für die durch die bilineare Extrapolation der bisherigen Stützwerte (vgl. Kapitel 5.1.1) und der bisherigen Formel (vgl. Formel 3.3) pro Wassergehalt dargestellt. Gut zu erkennen sind die zwei linearen Geraden der bisherigen REMO-Methode. Die erste Gerade im unteren Wassergehaltbereich modelliert eher höhere Werte im Vergleich zu den anderen Methoden. Die zweite Gerade liegt unterhalb der Sättigungsgrenze eher im unteren Wertebereich. Des Weiteren fällt in der Abbildung auf, dass sich die Kurven besonders bei ungefrorenem Boden mit geringem Wassergehalt in ihrem Verhalten unterscheiden. Bei der Methode nach Farouki werden sogar in der Realität nicht mögliche negative thermale Leitfähigkeiten modelliert (in der Abbildung aufgrund der gewählten y-Achse nicht zu erkennen), da hier die Kerstenzahl aufgrund des Logarithmuswertes negativ wird. Sowohl die Johansen- als auch die Lawrence & Slater-Methode ergeben in diesem unteren Wassergehaltbereich kleine Ausschläge nach oben. Dies ist durch die jeweils unterschiedlichen Berechnungsmethoden für die Kerstenzahl anhand der Sättigungsbedingungen (vgl. Tab. 12) zu erklären.

Trotz der höheren Sättigungsgrenze bei 100% organischem Anteil (vgl. Abb. 37C und 37D) werden sowohl im gefrorenen als auch im ungefrorenen Boden geringere Wärmeleitfähigkeiten erreicht. Auch bei den Methoden von Johansen und Farouki, bei denen der organische Anteil nicht explizit in die Berechnungsformel eingeht, wird die thermale Leitfähigkeit durch den hohen Porenvolumenwert und der daraus resultierenden niedrigen Wärmeleitfähigkeit für den trockenen und gesättigten Boden stark verringert. Anhand der Diagramme ist zudem zu erkennen, dass der Eisanteil des Bodens in REMO bisher für die Berechnung der Wärmeleitfähigkeit nicht berücksichtigt wird. Im Modell wird lediglich eine konstante thermale Diffusivität für Eis ( $12 \cdot 10^{-7} \frac{\text{m}^2}{\text{s}}$ ) für Pixel der Gletschermaske angegeben. Generell ist zu erkennen, dass je nach Methode unterschiedliche Maximalwerte erreicht werden. Besonders groß ist die Differenz bei 0% organischem Anteil und gefrorenem Boden. Hier werden im Beispielboden an der Sättigungsgrenze für Johansen ca. 2.75, für Farouki schon 3.75 und für Lawrence & Slater mit  $4.75 \frac{\text{W}}{\text{mK}}$  fast doppelt so hohe Werte erreicht.

Außerdem lässt sich feststellen, dass sich die Methoden nicht in eine Reihenfolge bringen lassen. Je nach vorherrschenden Bedingungen erreicht eine Methode höhere Werte als eine andere und bei anderen Verhältnissen dreht sich die Reihenfolge um. Aus diesen Werten lässt sich nicht ableiten, welche Approximationsmethode die Realität am besten nachbildet. Der Vergleich mit Messdaten in unterschiedlichen Studien liefert diesbezüglich ebenfalls keine eindeutigen Ergebnisse. Für gefrorene Böden empfiehlt Farouki (1981a) beispielsweise die Johansen-Methode, da durch diese

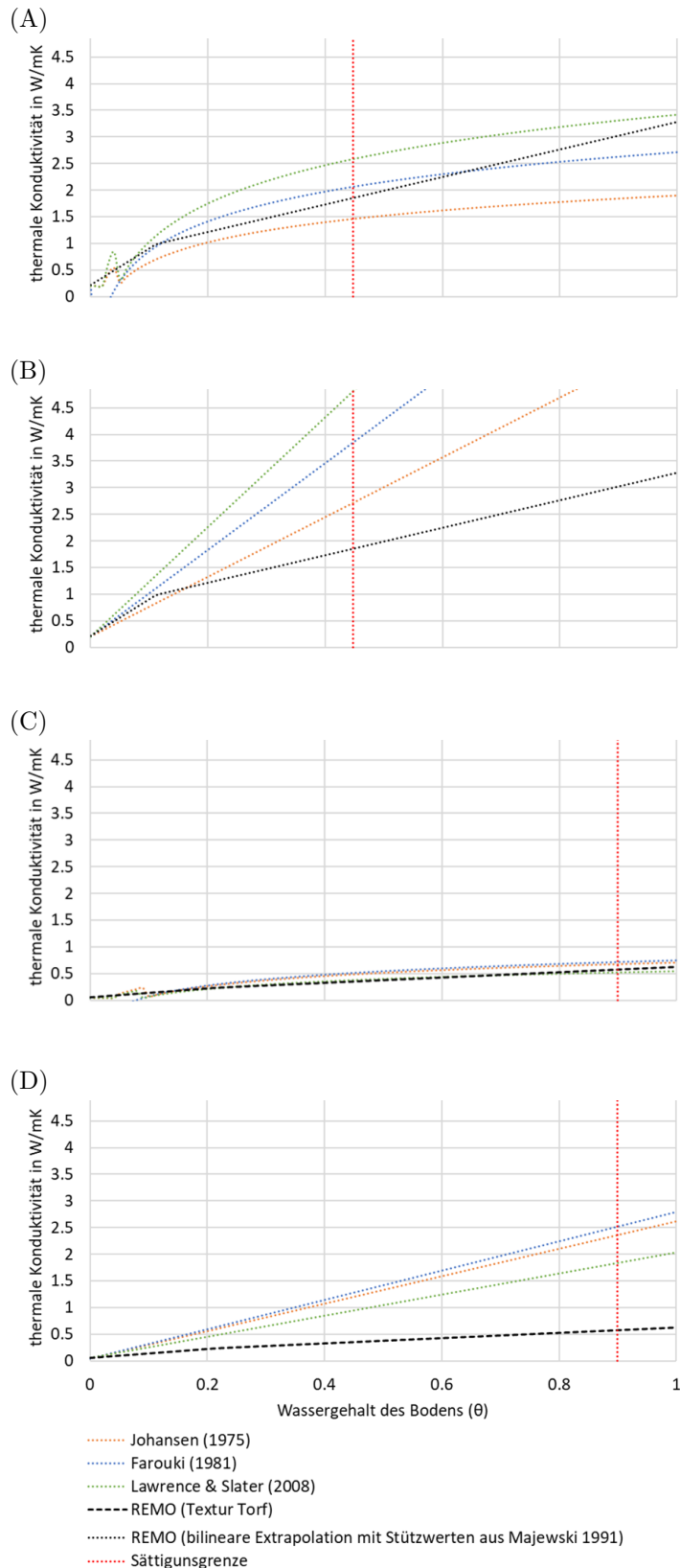


Abb. 37: Vergleich verschiedener Methoden zur Berechnung der Wärmeleitfähigkeit für einen Boden mit 33% Sand- und 33% Tonanteil, sowie mit:  
 (A) 0% organischem Anteil / ungefroren ; (B) 0% organischem Anteil / gefroren  
 (C) 100% organischem Anteil / ungefroren ; (D) 100% organischem Anteil / gefroren  
 Die vertikale Sättigungsgrenze zeigt den nach Lawrence & Slater (2008) berechneten Wert des Porenvolumens (Formel 5.3) für den Beispielboden an.

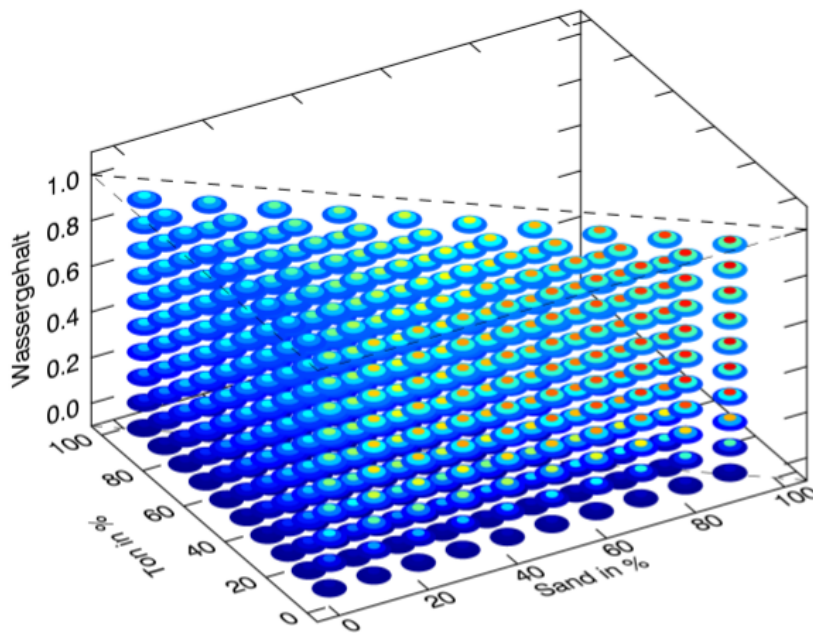
bessere Approximationen erreicht wurden. Zudem schnitt die überarbeitete Johansen-Formel laut Peters-Lidard et al. (1998) auch für ungefrorene Böden besser ab als die ebenfalls oft zitierte Methode von Mccumber & Pielke (1981), die einen exponentiellen statt logarithmischen Verlauf modelliert. Auch Dai, Wei et al. (2019) bescheinigen der Johansen-Methode und besonders den daraus abgeleiteten Methoden (v. a. Balland & Arp, 2005) eine größere Genauigkeit im Vergleich zu Beobachtungsdaten als z. B. die de Vries- oder die Farouki-Methode. Hierbei ist allerdings zu beachten, dass in dieser Studie nicht die originalen Methoden, sondern jeweils durch den organischen Anteil und durch einen Grobmaterialanteil ergänzte Methoden verwendet wurden. Bei Ren et al. (2019) hingegen konnte die Johansen-Methode beim Vergleich von insgesamt neun unterschiedlichen Wärmeleitfähigkeitsmodellen nicht so gut abschneiden. Laut Dong et al. (2015) unter- bzw. überschätzt die Johansen-Formel die Wärmeleitfähigkeit bei niedriger bzw. hoher Sättigung. Einschränkungen bezüglich der Gültigkeit einer bestimmten Methode sind in fast allen Studien zu finden. Jede Methode bildet Teilbereiche der Kurve besser und andere Teilbereiche schlechter ab als andere Methoden.

Da es keine überlegene Methode zu geben scheint, ist es nicht verwunderlich, dass verschiedene Modelle unterschiedliche Parametrisierungen nutzen. Die von Peters-Lidard abgewandelte Methode von Johansen befindet sich beispielsweise im Organizing Carbon and Hydrology In Dynamic Ecosystems (ORCHIDEE) (Krinner et al., 2005; Dufresne et al., 2013; F. Wang et al., 2016), im Interaction Soil-Biosphere-Atmosphere (ISBA) (Decharme et al., 2011), in JULES (Best et al., 2011), im Met Office Unified Model (UM) und im NOAA (Ek et al., 2003) LSM. Die originale Johansen-Methode wird dagegen in NOAA (Ek et al., 2003), im Community Atmosphere Biosphere Land Exchange (CABLE) (Kowalczyk et al., 2006), im Canadian Land Surface Scheme (CLASS) (Verseghy, 1991) Modell und im Purdue Regional Model (PRM) (W.-Y. Sun & Chern, 2005) (vgl. He et al., 2019) verwendet. Die ebenfalls vorgestellte Parametrisierung nach Lawrence und Slater, die zusätzlich den organischen Anteil berücksichtigt, wird im Community Land Model (CLM) (Oleson et al., 2010; Oleson & Lawrence, 2013) und in der für die hohen Breiten angepasste Version ORCHIDEE-MICT (Guimberteau et al., 2018) verwendet. J. Chen et al. (2018) nutzen für ihre Analyse der thermalen und hydraulischen Eigenschaften im CLM4.5 jedoch eine Abwandlung der Gewichte der Sand- und Tonanteile.

In dieser Arbeit wurde aufgrund der Berücksichtigung von organischem Material die Lawrence & Slater-Methode in REMO eingebaut und ab Modellversion id16 (vgl. Abb. 46) in den Sensitivitätsstudien verwendet. Abbildung 38 zeigt für diese Methode die Wärmeleitfähigkeit in Abhängigkeit von Sand- und Tonanteil sowie von organischem Anteil und vom Wassergehalt. Anders als in der vorherigen Abbildung 37 wurde die Sättigung auf 1 begrenzt, sodass die thermale Leitfähigkeit nicht mehr ansteigt, wenn der Wassergehalt das Porenvolumen erreicht.



(A)



(B)

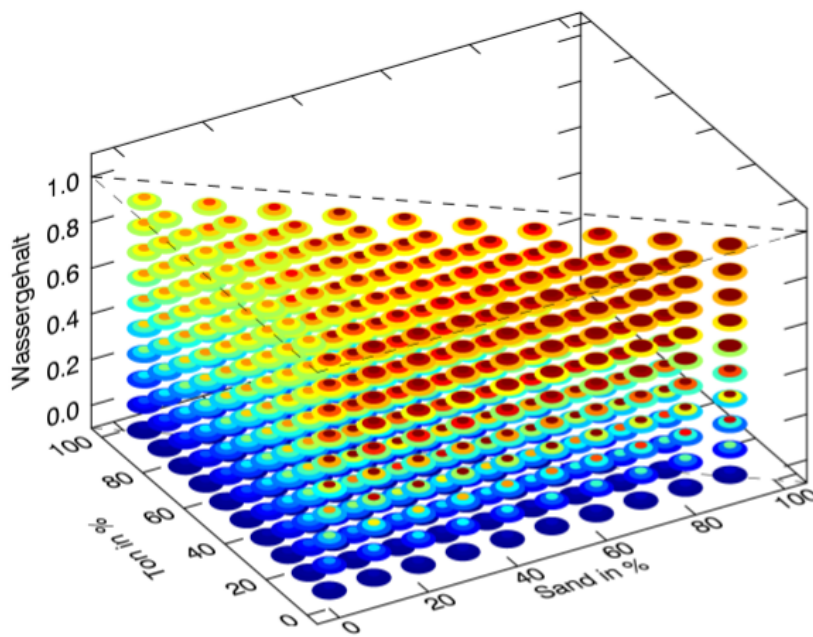


Abb. 38: Vergleich der Wärmeleitfähigkeit nach Lawrence & Slater (2008) in verschiedenen Böden  
 (A) ungefroren  
 (B) gefroren

Der Wassergehalt (z-Achse) stellt nicht die Sättigung, sondern den Anteil der Wassersäule an der gesamten Bodensäule dar. Erreicht der Wassergehalt das Porenvolumen, das von der Bodenzusammensetzung abhängt (hier berechnet nach Clapp & Hornberger (1978) und Cosby et al. (1984)), verändert sich die Wärmeleitfähigkeit nicht mehr.

Die Definitionslücke bei dieser Methode, die sich aufgrund des Nenners bei der Berechnung des  $\lambda_{sol}$  Wertes für einen Boden mit 100% Schluffanteil ergibt, wird durch den Mittelwert der Wärmeleitfähigkeit von Sand- und Tonanteil angenähert. Dieser Fall tritt jedoch durch die räumliche Interpolation in beiden Testuntersuchungsgebieten nicht auf und ist daher nur für die Abbildung relevant.

Wie bereits die unterschiedlichen Wärmeleitfähigkeiten der Bestandteile des Bodens (Tab. 2) zeigen, variiert die effektive Wärmeleitfähigkeit abhängig vom jeweiligen Anteil sehr deutlich. Der Unterschied zwischen gefrorenem und ungefrorenem Boden ist besonders groß. Es ist auch zu erkennen, dass der Anstieg der thermalen Leitfähigkeit stärker bei einer Erhöhung des Sandanteils als bei einer Erhöhung des Tonanteils ausfällt. Vergleicht man außerdem die Leitfähigkeiten für einen gewissen Sand-/Ton- und Wassergehaltspunkt, zeigt sich, dass je mehr organisches Material im Boden vorhanden ist, die Wärmeleitfähigkeit desto geringer ist.

### **Wärmekapazität**

Anders als die Wärmeleitfähigkeit ist die Wärmekapazität linear vom Wassergehalt abhängig. Zudem sind die Unterschiede zwischen verschiedenen Böden nicht so groß wie bei der Leitfähigkeit. Die Unterschiede basieren hauptsächlich auf den resultierenden Porenvolumina der Bodenarten (vgl. Bonan, 2019). Wegen des nachgewiesenen linearen Zusammenhangs existieren für die Berechnung der Wärmekapazität auch nicht so viele verschiedene Methoden zur Approximation. An et al. (2016) berechnen sie beispielsweise unter der Voraussetzung, dass im Boden weniger organisches Material vorhanden ist, als die Summe des Anteils von Wasser ( $4.18 \cdot \theta$ ) und dem 0.85-fachen der Trockenrohddichte des Bodens. Auch Abu-Hamdeh (2003) verwendet die Trockenrohddichte sowie den Wasser-, Sand- und Tonanteil, um mittels einer Geraden die volumetrische Wärmekapazität zu berechnen. Jedoch wird besonders das de Vries-Modell (vgl. Formel 2.5) aufgrund der guten allgemeinen Approximationseigenschaft wegen der Berücksichtigung von Wasser- und Eisgehalt in vielen LSMs verwendet (z. B. CLM, ORCHIDEE, JSBACH, ISBA, NOAH). Das LSM JULES verwendet dagegen den Ansatz von Cox et al. (1999), welcher zusätzlich die vorherrschende Temperatur berücksichtigt (vgl. F. Wang et al., 2016). Unterschiede zwischen den einzelnen Modellen ergeben sich vor allem durch die Parametrisierung des Wertes für die volumetrische Wärmekapazität des trockenen Bodenfeststoffanteils ( $c_{sol,dry}$ ). Während bei einigen Ansätzen ein für alle Bodenarten konstanter Wert eingesetzt wird, weisen andere Formeln Werte abhängig von z. B. Sand-, Ton- und/oder organischem Anteil aus - ähnlich dem Vorgehen bei der Ermittlung des Wertes für die Wärmeleitfähigkeit von trockenem Boden.

Die in dieser Arbeit vorgenommenen Änderungen an der volumetrischen Wärmekapazität werden zum einen durch neue Werte für diese trockenen Bodenfeststoffe (mit Berücksichtigung des organischen Anteils) (vgl. Vorgehen von Tran et al. (2017) für CLM4.5) und zum anderen durch die Berücksichtigung des Luftvolumens erzeugt. Des Weiteren ändert sich die Gesamtwärmekapazität des Bodens durch veränderte Porenvolumenwerte, deren verschiedene Parametrisierungen in Kapitel 5.1.1 beschrieben sind.

Wie Formel 2.5 zeigt, ergibt sich die volumetrische Wärmekapazität aus der mit dem jeweiligen Volumenanteil gewichteten Summe einzelner Bodenkomponenten (Feststoff, Wasser, Eis und Luft). Durch die Tatsache, dass das Porenvolumen den Wasser- und Luftanteil im Boden ausdrückt, lässt sich diese Formel zu Formel 5.8 umformulieren. Dies entspricht der von de Vries (1963) aufgestellten Formel, welcher für die volumetrische Wärmekapazität des Feststoffanteils einen konstanten Wert von  $1.926 \frac{\text{MJ}}{\text{m}^3\text{K}}$  angibt. Der erste Summand der Formel entspricht der volumetrischen Wärmekapazität des Feststoffanteils in einem trockenen Boden, der in REMO bisher direkt aus Tabelle 3 ohne den Multiplikationsfaktor eingesetzt wurde (vgl. bisherige Formel 3.2 zur Berechnung der volumetrischen Wärmekapazität). Lawrence & Slater (2008) machen diesen Kapazitätswert (wie bei der thermischen Konduktivität) vom jeweiligen Sand-, Ton- und organischen Anteil abhängig (vgl. Formel 5.9). Die Kombination aus beiden Formeln ergibt mit zusätzlicher Berücksichtigung des Luftanteils Formel 5.10.

$$c_v = (1 - \theta_{sat}) \cdot c_{sol,dry} + (\theta \cdot f_u) \cdot c_{wat} + (\theta \cdot (1 - f_u)) \cdot c_{ice} \quad (5.8)$$

mit:

$$c_{sol,dry} = 1.926 \cdot 10^6 \frac{J}{m^3K}$$

$$c_{sol,dry} = (1 - f_{org}) \cdot \frac{2.128 \cdot SAND + 2.385 \cdot CLAY}{SAND + CLAY} \cdot 10^6 + f_{org} \cdot c_o \quad (5.9)$$

mit:

$$c_o = 2.5 \cdot 10^6 \frac{J}{m^3K}$$

$$c_v = (1 - \theta_{sat}) \cdot c_{sol,dry} + (\theta \cdot ((1 - f_u) \cdot c_{ice}) + (f_u \cdot c_{wat})) - \theta \cdot c_a \quad (5.10)$$

mit:

$\theta$  = aktueller Wassergehalt in  $m$  (ZWQ)

$f_u$  = ungefrorener Anteil am aktuellen Wassergehalt (ZWI)

$\theta_{sat}$  = Porenvolumen in  $\frac{m^3}{m^3}$  (PORVOL)

SAND = Sandanteil des Bodens in %

CLAY = Tonanteil des Bodens in %

$f_{org}$  = organischer Anteil des Bodens

$c_o$  = vol. Wärmekapazität von organischem Material in  $\frac{J}{m^3K}$

$c_{sol,dry}$  = vol. Wärmekapazität des trockenen Bodens in  $\frac{J}{m^3K}$

$c_{wat} = c_{s,w} \cdot \rho_w = 4186 \frac{J}{kgK} \cdot 1000 \frac{kg}{m^3} = \text{vol. Wärmekapazität von Wasser (CLW} \cdot \text{RHOH2O)}$

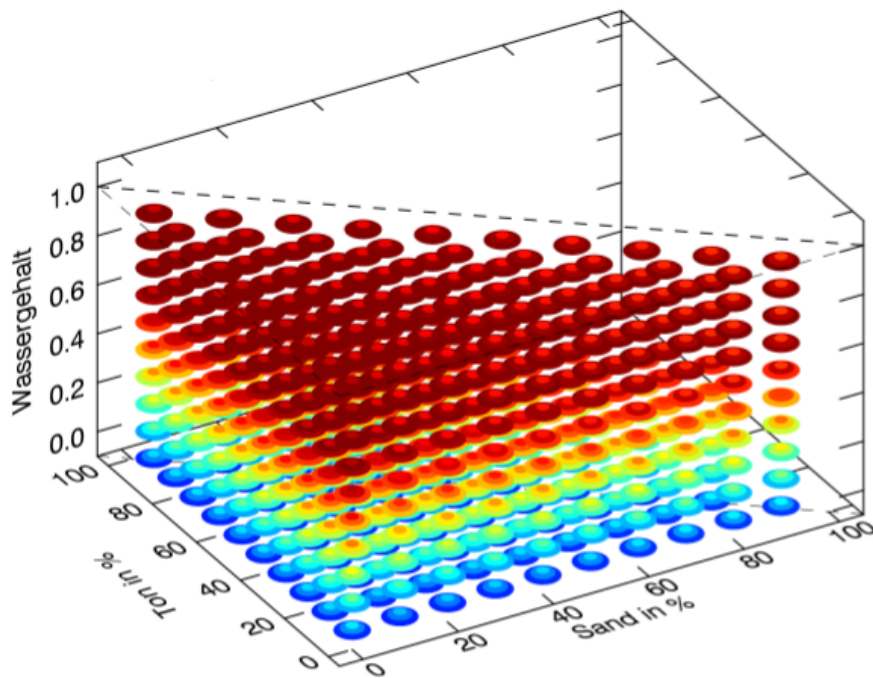
$c_{ice} = c_{s,i} \cdot \rho_i = 2093 \frac{J}{kgK} \cdot 917 \frac{kg}{m^3} = \text{vol. Wärmekapazität von Eis (ZCPICE} \cdot \text{RHOICE)}$

$c_a = c_{s,a} \cdot \rho_a = 1006 \frac{J}{kgK} \cdot 1.2041 \frac{kg}{m^3} = \text{vol. Wärmekapazität von Luft}$

Ab Modelllauf id16 (vgl. Abb. 46) wird diese neue Parametrisierung in REMO verwendet. Anhand der Formel für die trockene volumetrische Wärmekapazität der mineralischen Bestandteile des Bodens (5.9) lässt sich ableiten, dass Ton eine höhere Kapazität besitzt als Sand und damit eine Änderung des Tonanteils stärkere Änderungen der thermalen Kapazität nach sich zieht als eine Änderung des Sandanteils. Da sich die Werte im Vergleich zu den anderen Komponenten nur minimal unterscheiden, ist dieser Effekt in Abbildung 39, welche die mit Formel 5.10 berechneten volumetrischen Wärmekapazitäten für unterschiedliche Bodenzusammensetzungen darstellt, nur sehr schwach erkennbar. Die unterschiedlichen volumetrischen Wärmekapazitäten der einzelnen BodenkompONENTEN (vgl. Tab. 1) erklären ebenfalls den große Unterschied zwischen komplett ungefrorenem (vgl. Abb. 39A) und gefrorenem (vgl. Abb. 39B) Boden. Eis besitzt nur eine etwa halb so große Wärmekapazität wie Wasser. Dennoch ist auch beim gefrorenen Boden zu erkennen, dass die Faktoren Wasser und organischer Anteil die maßgeblichen Einflussfaktoren zur Bildung der Gesamtwärmekapazität sind. Je mehr Wasser im Boden enthalten ist, desto höher wird die Kapazität. Beim organischen Anteil ist der Zusammenhang nicht eindeutig. Während im unteren Wassergehaltsbereich Böden ohne organischen Anteil höhere Wärmekapazitäten erreichen als Böden mit höheren Anteilen, dreht sich das Verhältnis im oberen Wertebereich des Wassergehalts um. Dies hängt mit dem hohen Porenvolumen von organischem Material zusammen. Bei niedrigem organischen Anteil steht nicht so viel Porenvolumen zur Verfügung, das überhaupt mit Wasser (mit hoher thermaler Wärmeleitfähigkeit) gefüllt werden kann.

Diese neu berechneten Wärmekapazitätswerte stimmen gut mit den in der Literatur angegebenen Werten für trockene und gesättigte Böden überein. Beispielsweise wird in Pahud (2002) für trockenen Ton eine volumetrische Wärmekapazität von 1.5–1.6, für trockenen Sand hingegen 1.3–1.6 sowie für gesättigten Ton 1.6–3.4 und für gesättigten Sand 2.2–2.9 bzw. für Torf 0.5–3.8  $\frac{\text{MJ}}{\text{m}^3\text{K}}$  angegeben. Auch die bisher in REMO verwendeten Werte für die trockene Wärmekapazität (vgl. Tab. 3) für die Klassen Sand (1.28) und Ton (1.63) bzw. Torf (0.58) passen zu den neuen Werten, sodass diese eine geeignete Ergänzung darstellen. Wie sich diese neuen Parametrisierungen auf die Werte in den beiden Untersuchungsgebieten auswirken, wird in Kapitel 6.4 dargestellt.

(A)



(B)

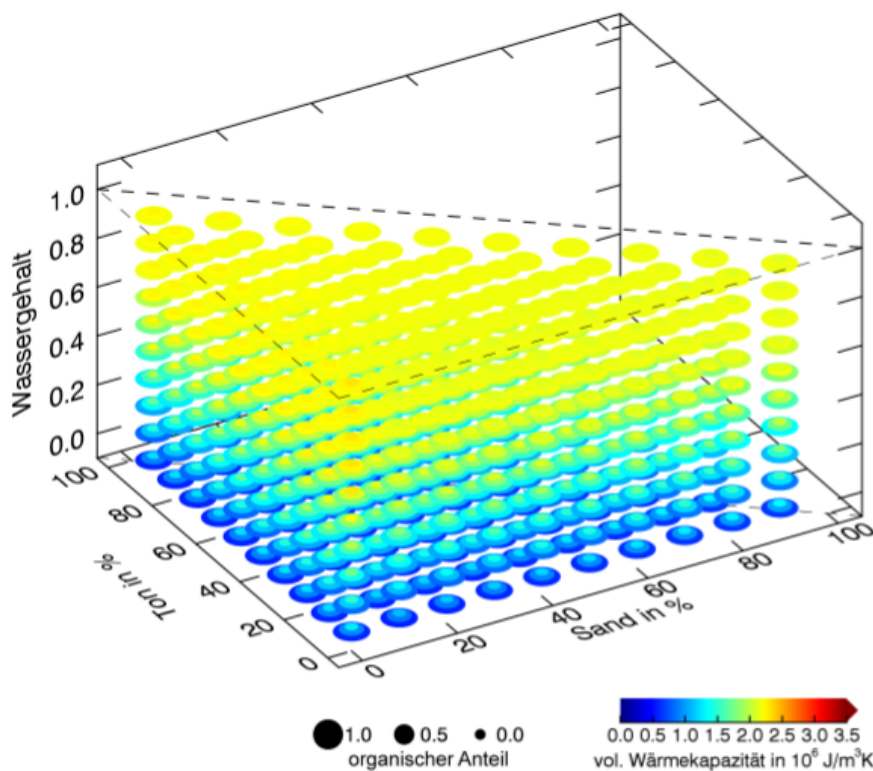


Abb. 39: Vergleich der volumetrischen Wärmekapazität nach Lawrence & Slater (2008) in verschiedenen Böden

(A) ungefroren

(B) gefroren

Der Wassergehalt (z-Achse) stellt nicht die Sättigung, sondern den Anteil der Wassersäule an der gesamten Bodensäule dar. Erreicht der Wassergehalt das Porenvolumen, das von der Bodenzusammensetzung abhängt (hier berechnet nach Clapp & Hornberger (1978) und Cosby et al. (1984)), verändert sich die Wärmekapazität nicht mehr.

## 5.2 Neues Bodentemperaturschema und Anpassung des Bodenwärmestroms

### Neues Bodentemperaturschema

Die in Kapitel 2.2 darstellte Gleichung zur Berechnung der Bodentemperatur (Formel 2.14) kann der Gruppe der eindimensionalen partiellen Differentialgleichungen zugeordnet werden und wird in Klimamodellen meist für mehrere Bodenschichten aufgestellt. Das sich dann ergebende Gleichungssystem kann auf unterschiedliche Weise gelöst werden. Werden  $c_i$  und  $\lambda_i$  als konstant über Zeit und Raum angenommen, ist die thermale Diffusivität  $\alpha_i$  ebenfalls konstant und die Gleichung lässt sich analytisch mit verschiedenen Methoden lösen, indem die Bodentemperatur in jeder beliebigen Tiefe beispielsweise als Variation eine Sinuskurve ausdrückt wird. An et al. (2016) und Mikailsoy & Shein (2019) beschreiben und vergleichen exemplarisch weitere analytische Methoden (je die Amplituden-, die Phasenverschiebungsmethode sowie die Harmonische und die Konduktion-Konvektionsmethode bzw. die Arctangens- und die Logarithmische Methode). Durch die in Kapitel 5.1.2 erläuterten Änderungen unterscheiden sich (bei entsprechender Wahl der Einstellung) die Wärmeeigenschaften der fünf Bodenschichten in REMO, sodass die bisherigen Gleichungen und deren Lösungsmethode zur Berechnung der Bodentemperaturänderungen nicht mehr allgemein gültig und instabil sind (Bonan, 2019).

Den analytischen Methoden gegenüber stehen verschiedene numerische Methoden, um die Diffusionsgleichung zu lösen. Bei der Finite-Differenzen-Methode wird die Gleichung nicht für den ganzen Raum bestimmt, sondern für einzelne Gitterpunkte approximiert, die durch das Einteilen von Raum und Zeit in diskrete Abschnitte entstehen. Dazu wird die Ableitung in Formel 2.14 durch den zentralen Differenzenquotienten angenähert. Je kleiner die Intervalle gewählt werden, desto genauer wird die Lösung approximiert (Rahimi et al., 2013).

Für die Energiebilanz einer Bodenschicht  $i$  gilt damit:

$$\frac{c_i \Delta z_i}{\Delta t} (T_i^{n+1} - T_i^n) = -F_{i-1} + F_i \quad (5.11a)$$

mit:

$i$  = Bodenschicht  $i$

$n$  = aktueller Zeitschritt

$n + 1$  = darauffolgender Zeitschritt

$\Delta t$  = Zeitschritt in s

$\Delta z_i$  = Schichtdicke von Schicht  $i$

$c_i$  = volumetrische Wärmekapazität

$T_i^n$  = Bodentemperatur in Schicht  $i$  zum Zeitschritt  $n$

$F_i$  = Wärmefluss von der Schichtmitte  $i$  zur Schichtmitte  $i+1$  (positiv entspricht nach oben)

und  $F_i$  lässt sich berechnen mit:

$$F_i = -\lambda_i \frac{T_i^n - T_{i+1}^n}{z_{i+1} - z_i} \quad (5.11b)$$

mit:

$\lambda_i$  = Wärmeleitfähigkeit der Bodenschicht  $i$

$z_i$  = Schichtdicke der Bodenschicht  $i$

Formel 5.11a lässt sich mit einer expliziten, impliziten oder einer Mischmethode lösen. In Abbildung 40 sind die sogenannten Differenzensterne (engl. stencil) für drei Lösungsmethoden für den eindimensionalen Fall dargestellt, der für die hier verwendete partielle Differentialgleichung benötigt wird. Die Sterne verdeutlichen den Unterschied der einfließenden Werte zur Bildung der neuen Bodentemperaturwerte. Weitere Ausprägungen und Abwandlungen sowie deren jeweilige Eigenschaften können in Richtmyer & Morton (1994) nachgeschlagen werden. Diese Eigenschaften (Konsistenz, Stabilität und Konvergenz) beschreiben die Güte der Approximation (Munz & Westermann, 2019).

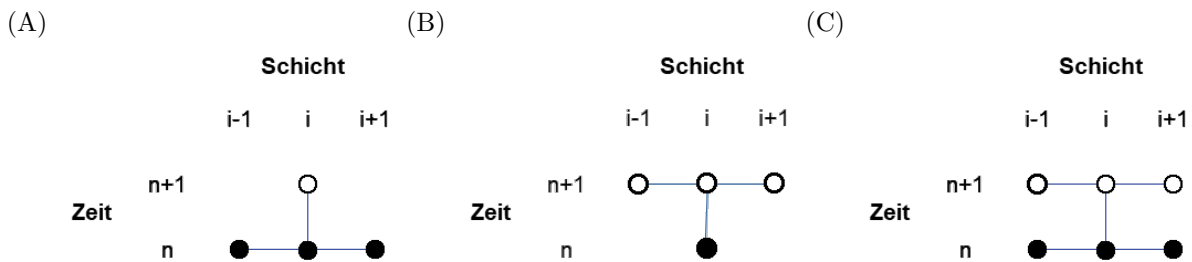


Abb. 40: Differenzensterne zur Darstellung der räumlichen und zeitlichen Zusammenhänge bei verschiedenen Lösungsansätze (nach Bonan, 2019). Nicht ausgefüllte Knoten entsprechen unbekanntem Werten, die berechnet werden.

- (A) explizite Methode
- (B) implizite Methode
- (C) Crank-Nicolson Methode

Bei der expliziten Methode (vgl. Formel 5.12) wird die Temperatur für eine definierte Bodenschicht für den nächsten Zeitschritt abhängig von den Temperaturen der jeweils darüberliegenden, der definierten und der darunterliegenden Bodenschicht des aktuellen Zeitschritts berechnet (vgl. Abb. 40A). Mit je einer Unbekannten für  $N$  Bodenschichten ergibt sich ein lineares Gleichungssystem erster Ordnung. Allerdings ist die Lösung von diesem nur numerisch stabil und konvergiert, wenn  $M \leq 0.5$  (Qin et al., 2002; Bonan, 2019).

$$\frac{\delta T}{\delta t} = \frac{T_i^{n+1} - T_i^n}{\Delta t} = \frac{\lambda}{c_v} \left( \frac{T_{i-1}^n - 2 \cdot T_i^n + T_{i+1}^n}{\Delta z^2} \right) = -F_{i-1}^n + F_i^n \quad (5.12a)$$

was sich vereinfachen lässt zu:

$$T_i^{n+1} = T_i^n + M \cdot (T_{i-1}^n - 2 \cdot T_i^n + T_{i+1}^n) \quad (5.12b)$$

bzw.

$$T_i^{n+1} = M \cdot T_{i-1}^n + (1 - 2 \cdot M) \cdot T_i^n + M \cdot T_{i+1}^n \quad (5.12c)$$

mit:

$$M = \frac{\lambda \cdot \Delta t}{c_v \Delta z^2}$$

Die implizite Methode (vgl. Formel 5.13) nutzt hingegen die drei Schichttemperaturen des nächsten Zeitschritts (vgl. Abb. 40B), um die Differenz der linken Seite zu beschreiben, wodurch sich die Vorzeichen der Koeffizienten und die Formulierung der Gleichung hinsichtlich der Zeitabhängigkeit umdrehen. Das Verfahren ist numerisch stabil und konvergiert für  $M > 0$  (Bonan, 2019).

$$\frac{T_i^{n+1} - T_i^n}{\Delta t} = \frac{\lambda}{c_v} \left( \frac{T_{i-1}^{n+1} - 2 \cdot T_i^{n+1} + T_{i+1}^{n+1}}{\Delta z^2} \right) = -F_{i-1}^{n+1} + F_i^{n+1} \quad (5.13a)$$

woraus folgt:

$$T_i^n = T_i^{n+1} + M \cdot (-T_{i-1}^{n+1} + 2 \cdot T_i^{n+1} - T_{i+1}^{n+1}) \quad (5.13b)$$

bzw.

$$T_i^n = -M \cdot T_{i-1}^{n+1} + (1 + 2 \cdot M) \cdot T_i^{n+1} - M \cdot T_{i+1}^{n+1} \quad (5.13c)$$

mit:

$$M = \frac{\lambda \cdot \Delta t}{c_v \Delta z^2}$$

Bei Mischmethoden werden beide Arten kombiniert, indem eine gewichtete Addition vorgenommen wird:

$$\frac{c_v \Delta z_i}{\Delta t} (T_i^{n+1} - T_i^n) = \alpha (-F_{i-1}^n + F_i^n) + (1 - \alpha) (-F_{i-1}^{n+1} + F_i^{n+1}) \quad (5.14)$$

mit:

$\alpha$  = Gewichtungsfaktor oder Implizitätsfaktor (engl. implicitness factor)

Für  $\alpha = 0.5$  ergibt sich die sogenannte Crank-Nicholson Methode (vgl. Abb. 40C)), die dem Mittelwert der beiden anderen Schemata entspricht. Sie kombiniert die numerische Stabilität sowie die Genauigkeit der Lösung in Raum und Zeit und ist daher eine Methode zweiter Ordnung. Die mathematischen Beweise für die Konvergenz und Stabilität der drei verschiedenen Methoden unter Berücksichtigung der Randbedingungen können bei Oishi et al. (2015) nachgelesen werden. Durch geschicktes Umformulieren und Umsortieren der Gleichung nach den verschiedenen Bodenschichten ergibt sich folgende Gleichung 5.15 zum Berechnen der Bodentemperaturänderung.

$$\frac{T_i^{n+1} - T_i^n}{\delta t} = \frac{\lambda}{c_v} \left( \frac{T_{i-1}^n - 2 \cdot T_i^n + T_{i+1}^n}{2 \cdot \Delta z^2} + \frac{T_{i-1}^{n+1} - 2 \cdot T_i^{n+1} + T_{i+1}^{n+1}}{2 \cdot \Delta z^2} \right) \quad (5.15a)$$

woraus folgt:

$$-\frac{1}{2} \cdot M \cdot T_{i-1}^{n+1} + (1 + M) \cdot T_i^{n+1} - \frac{1}{2} \cdot T_{i+1}^{n+1} = \frac{1}{2} \cdot M \cdot T_{i-1}^n + (1 - M) \cdot T_i^n + \frac{1}{2} \cdot T_{i+1}^n \quad (5.15b)$$

mit:

$$M = \frac{\lambda \cdot \Delta t}{c_v \Delta z^2}$$



Allerdings hängt  $c_i$ , wie in Kapitel 5.1.3 gezeigt, stark vom Wassergehalt ab und variiert daher über die Zeit. Werden die neuen Bodendaten (bei L5LAYER\_TOTAL=TRUE) verwendet, variieren sowohl  $c_i$  als auch  $\lambda_i$  nicht nur über die Zeit, sondern auch über den Raum, sodass  $M$  nicht für alle Schichten konstant ist und sich schichtabhängige Koeffizienten für die Diffusionsformeln ergeben.

Mit:

$$z_i = z_{i+0.5} - z_{i-0.5} = z_{i+0.5} - 0.5 \cdot \Delta z_i = \text{mittlere Tiefe der Schicht } i \quad (5.16a)$$

folgt:

$$z_i - z_{i-1} = \frac{1}{2}(\Delta z_{i-1} + \Delta z_i) \quad (5.16b)$$

und damit gilt:

- für die implizite Methode:

$$\begin{aligned} \frac{c_i \Delta z_i}{\Delta t} (T_i^{n+1} - T_i^n) &= -F_{i-1}^{n+1} + F_i^{n+1} = \\ &= -\frac{-\lambda_{i-1}}{\frac{1}{2}(\Delta z_{i-1} + \Delta z_i)} \cdot (T_{i-1}^{n+1} - T_i^{n+1}) + \frac{-\lambda_i}{\frac{1}{2}(\Delta z_i + \Delta z_{i+1})} (T_i^{n+1} - T_{i+1}^{n+1}) \end{aligned} \quad (5.17)$$

und durch Umformulieren:

$$a_i \cdot T_{i-1}^{n+1} + b_i \cdot T_i^{n+1} + c_i \cdot T_{i+1}^{n+1} = d_i \quad (5.18)$$

mit:

$$\begin{aligned} a_i &= \frac{-\lambda_{i-1}}{\Delta z_{i-1} + \Delta z_i} \\ c_i &= \frac{-\lambda_i}{\Delta z_i + \Delta z_{i+1}} \\ b_i &= \frac{c_i \Delta z_i}{2\Delta t} - a_i - c_i \\ d_i &= \frac{c_i \Delta z_i}{2\Delta t} \cdot T_i^n \end{aligned}$$

- für die explizite Methode:

$$\begin{aligned} \frac{c_i \Delta z_i}{\Delta t} (T_i^{n+1} - T_i^n) &= -F_{i-1}^n + F_i^n = \\ &= -\frac{-\lambda_{i-1}}{\frac{1}{2}(\Delta z_{i-1} + \Delta z_i)} \cdot (T_{i-1}^n - T_i^n) + \frac{-\lambda_i}{\frac{1}{2}(\Delta z_i + \Delta z_{i+1})} (T_i^n - T_{i+1}^n) \end{aligned} \quad (5.19)$$

und durch Umformulieren:

$$a_i \cdot T_{i-1}^n + b_i \cdot T_i^n + c_i \cdot T_{i+1}^n = d_i \quad (5.20)$$

mit:

$$a_i = \frac{\lambda_{i-1}}{\Delta z_{i-1} + \Delta z_i}$$

$$\begin{aligned}
 c_i &= \frac{\lambda_i}{\Delta z_i + \Delta z_{i+1}} \\
 b_i &= \frac{c_i \Delta z_i}{2\Delta t} - a_i - c_i \\
 d_i &= \frac{c_i \Delta z_i}{2\Delta t} \cdot T_i^{n+1}
 \end{aligned}$$

- für die Crank-Nicolson-Methode:

$$\begin{aligned}
 \frac{c_i \Delta z_i}{\Delta t} (T_i^{n+1} - T_i^n) &= \frac{1}{2} \cdot \left( -\frac{\lambda_{i-1}}{z_i - z_{i-1}} \cdot ((T_{i-1}^n - T_i^n) + (T_{i-1}^{n+1} - T_i^{n+1})) \right. \\
 &\quad \left. - \frac{\lambda_i}{z_{i+1} - z_i} \cdot ((T_i^n - T_{i+1}^n) + (T_i^{n+1} - T_{i+1}^{n+1})) \right) \quad (5.21)
 \end{aligned}$$

und damit:

$$\begin{aligned}
 \frac{c_i \Delta z_i}{\Delta t} (T_i^{n+1} - T_i^n) &= -\frac{\lambda_{i-1}}{\Delta z_{i-1} + \Delta z_i} \cdot ((T_{i-1}^n - T_i^n) + (T_{i-1}^{n+1} - T_i^{n+1})) \\
 &\quad - \frac{\lambda_i}{\Delta z_i + \Delta z_{i+1}} \cdot ((T_i^n - T_{i+1}^n) + (T_i^{n+1} - T_{i+1}^{n+1})) \quad (5.22)
 \end{aligned}$$

und durch Umformulieren:

$$a_i \cdot T_{i-1}^{n+1} + b_i \cdot T_i^{n+1} + c_i \cdot T_{i+1}^{n+1} = d_i \quad (5.23)$$

mit:

$$\begin{aligned}
 a_i &= \frac{-\lambda_{i-1}}{\Delta z_{i-1} + \Delta z_i} \\
 c_i &= \frac{-\lambda_i}{\Delta z_i + \Delta z_{i+1}} \\
 b_i &= \frac{c_i \Delta z_i}{\Delta t} - a_i - c_i \\
 d_i &= \frac{c_i \Delta z_i}{\Delta t} \cdot T_i^n - \frac{\lambda_{i-1}}{\Delta z_{i-1} + \Delta z_i} \cdot (T_i^n - T_{i-1}^n) - \frac{\lambda_i}{\Delta z_i + \Delta z_{i+1}} \cdot (T_i^n - T_{i+1}^n)
 \end{aligned}$$

Für die Implementierung in REMO wird aufgrund der erwähnten Vorteile die Crank-Nicolson-Methode (Gleichung 5.23) gewählt. Außerdem wird für die oberste Schicht  $i = 1$  der  $F_{i-1}^n$  Term (wie im bisherigen Schema) auf den Bodenwärmestrom  $TEM1$  (vgl. Formel 3.1 in Kapitel 3.3) (Dirichlet-Randbedingung) und der Term für die unterste Bodenschicht auf Null gesetzt (Neumann-Randbedingung). Bei der Dirichlet-Randbedingung wird die Temperatur am Rand als bekannt angenommen, während bei der Neumann-Randbedingung der Wärmefluss stattfindet und daher der Wert der Ableitung als Bedingung vorgegeben wird.

Wenn Gleichung 5.23 für alle Bodenschichten formuliert wird, wobei folglich  $a_1 = 0$  und  $c_N = 0$  gilt, ergibt sich dadurch ein tridiagonales Gleichungssystem (vgl. Formel 5.24) mit den Koeffizienten  $a_i$  und  $c_i$  als Elemente auf der Sub- bzw. Superdiagonalen und den jeweiligen Koeffizienten  $b_i$  auf der Diagonalen:

$$\begin{pmatrix} b_1 & c_1 & & & 0 \\ a_2 & b_2 & c_2 & & \\ & a_3 & \ddots & \ddots & \\ & & \ddots & \ddots & c_{N-1} \\ 0 & & & a_N & b_N \end{pmatrix} \cdot \begin{pmatrix} T_1^{n+1} \\ T_2^{n+1} \\ \vdots \\ T_{N-1}^{n+1} \\ T_N^{n+1} \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_{N-1} \\ d_N \end{pmatrix} \quad (5.24)$$

Dieses aufgestellte Gleichungssystem kann durch einen sogenannten „Tridiagonal Solve Algorithmus“ gelöst werden und gilt für alle möglichen Koeffizienten als numerisch stabil. In dieser Arbeit wurde der Thomas-Algorithmus (Thomas, 1949) verwendet. Dieser basiert auf dem Gaußschen Eliminationsschema (vgl. Formel 5.25a) und dem anschließenden Rückwärts-Einsetzverfahren (vgl. Formel 5.25b), sodass die Lösung in  $O(N)$  Rechenoperationen gefunden werden kann. Dieser Ansatz wird in REMO in ganz ähnlicher Weise bereits für die Berechnung des Wassergehaltes pro Schicht bei aktiviertem 5Layer-Schema verwendet. Die Koeffizienten  $a, b, c$  und  $d$  werden an ein neues Unterprogramm übergeben, in dem der Algorithmus realisiert wird. Die dann berechneten neuen Bodentemperaturen werden anschließend zurück an das Programm *surf.f* übertragen.

$$c'_i = \begin{cases} \frac{c_1}{b_1}, & \text{für } i=1 \\ \frac{c_i}{b_i - c'_{i-1} \cdot a_i}, & \text{für } i=2, \dots, N-1 \end{cases} \quad (5.25a)$$

$$d'_i = \begin{cases} \frac{d_1}{b_1}, & \text{für } i=1 \\ \frac{d_i - d'_{i-1} \cdot a_i}{b_i - c'_{i-1} \cdot a_i}, & \text{für } i=2, \dots, N \end{cases}$$

$$T_i^{n+1} = \begin{cases} d'_N, & \text{für } i=N \\ d'_i - c'_i \cdot T_{i+1}^{n+1}, & \text{für } i=N-1, 1 \end{cases} \quad (5.25b)$$

### Anpassung der Frequenz des Strahlungsupdates

Abhängig von der Verwendung dieses neuen numerischen Schemas und zusätzlich abhängig von einer Kombination verschiedener Parameteränderungen im Modellcode und dem verwendeten Untersuchungsgebiet, trat beim Testen des neuen Bodentemperaturberechnungsschemas der Fall auf, dass das Modell aufgrund von Oberflächentemperaturen unter 100 Kelvin abgebrochen (vgl. Sicherheitsschalter in *surfhum.f*, TLUCUA) ist. Bei der genaueren Analyse fiel auf, dass nur die Oberflächentemperatur und die Temperatur der ersten Bodenschicht von diesem Phänomen betroffen waren. Ebenso trat die fehlerhafte Temperatur nicht plötzlich ein, sondern äußerte sich als oszillierende Schwankung mit im zeitlichen Verlauf größer werdender Amplitude, was erst durch eine Analyse der tatsächlichen Werte pro Zeitschritt von 240 Sekunden, nicht aber bei den üblicherweise analysierten und im Fall der Temperatur gemittelten Stundenwerten zu beobachten ist. Zusätzlich trat das Phänomen nur an einzelnen Pixeln auf, die je nach verwendeten Parametern

variierten. In Abbildung 41 ist das beschriebene Verhalten der Oberflächentemperatur (TSL) zweier Beispielpixel (Westküste Island und Südküste Grönland) dargestellt.

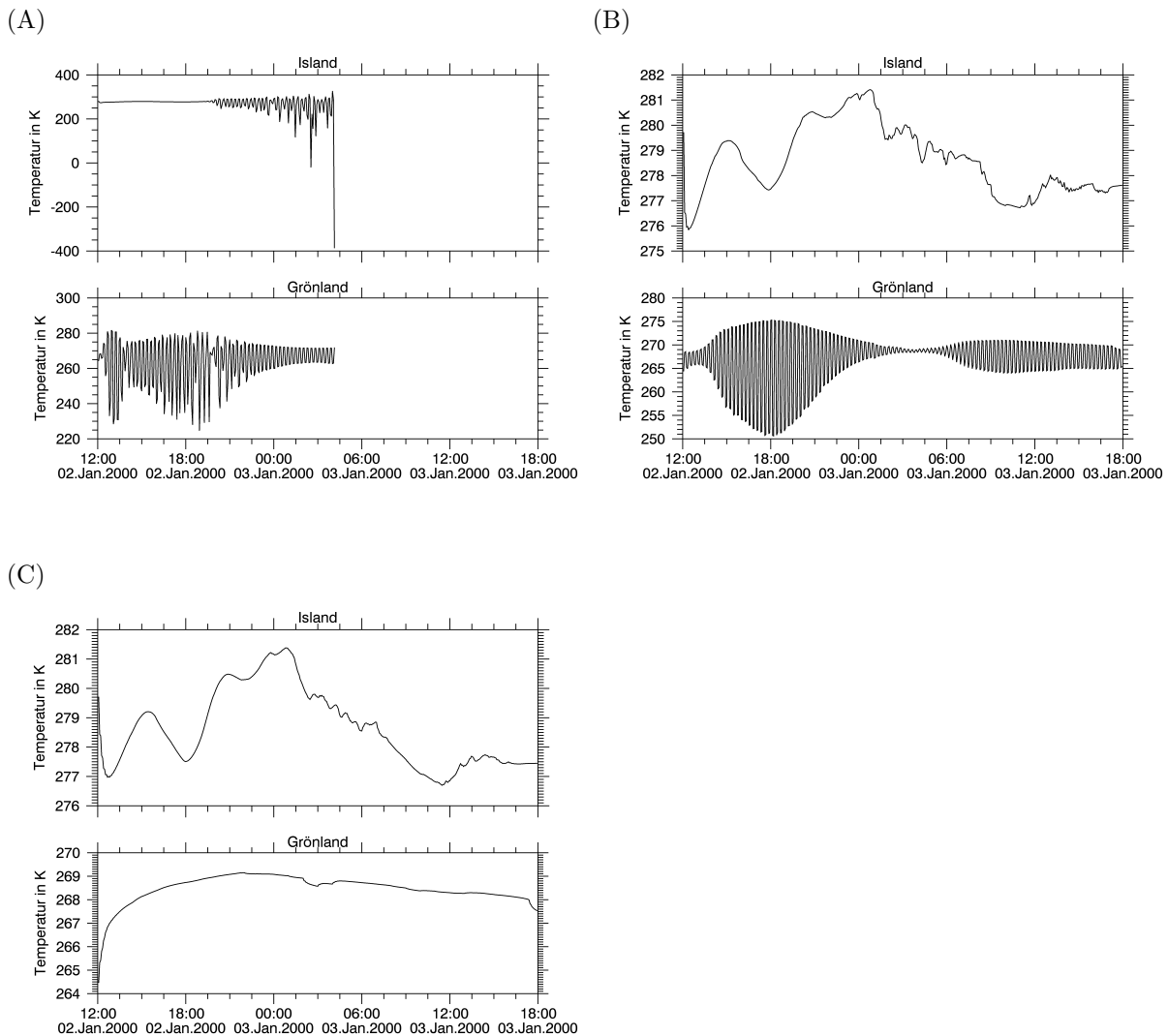


Abb. 41: Auswirkung der veränderten Parametrisierung und des neuen Temperaturschemas auf die Oberflächentemperatur anhand zweier Beispielpixel (oben: Island; unten: Grönland)  
 (A) Modellabbruch aufgrund von zu stark oszillierenden Oberflächentemperaturen  
 (B) kein Modellabbruch nach Zusammenfassung mehrerer Parametrisierungsänderungen und Update des Wassergehalts trotzdem nicht numerisch stabil  
 (C) finale stabile Version mit pro Zeitschritt aktualisiertem Strahlungsinput und allen neuen Parametrisierungen

Durch die Tatsache, dass dieses Verhalten nur in der obersten Bodenschicht, nur an bestimmten Pixeln und nur bei Verwendung des neuen Bodentemperaturschemas auftrat, konnte abgeleitet werden, dass es sich um ein numerisches Problem handelt. Mit der Kombination mehrerer voneinander abhängiger Parametrisierungsschritte bezüglich der Wärmeleitfähigkeit und Wärmekapazität der Bodenschichten sowie der Anpassung des Wassergehalts pro Schicht bei Verwendung des 5Layer-Schemas konnten die Amplituden der Schwankungen stark reduziert und der Modellabbruch somit verhindert werden. Dennoch traten bei genauerem Hinsehen an den „Problempixeln“

und auch teilweise an neuen Pixeln weiterhin das Phänomen von oszillierenden Oberflächentemperaturen auf (vgl. Abb. 41B). Die Besonderheit der obersten Bodentemperaturschicht im Vergleich zu den darunterliegenden Schichten liegt im Energiefluss, der von oben in diese eingeht und von der Strahlung abhängt (vgl. Additionsterm TEM1 in Formel 3.1). Dieser Term stellt die Veränderung des Bodenwärmestroms  $G$  (vgl. Formel 3.1 in Kapitel 3.3) pro Zeitschritt dar und ist direkt von den „Strahlungsvariablen“ abhängig. Mit Blick auf den Calltree (vgl. Anhang E: Aufrufhäufigkeit bei *radint* und dessen Kinder 2x statt wie bei *rad* und *surf* 60x) ist zu erkennen, dass die Strahlungsvariablen im vorliegenden Fall nur jede Stunde statt zu jedem Rechenzeitschritt neu berechnet werden. Da für die Bodentemperaturberechnung nur die Änderung des Bodenwärmeflusses benötigt wird, wird nicht die Aufrufhäufigkeit aller Strahlungsvariablen erhöht, sondern nur der für die Berechnung dieser Variable notwendige Teil aus der entsprechenden Subroutine in die Subroutine *surf* kopiert. Dadurch wird gewährleistet, dass die anderen Wärmeflussvariablen nicht verändert werden, sich keine Aufrufhäufigkeiten der einzelnen Subroutinen ändern und sich damit auch die Rechenzeit nicht signifikant erhöht.

Außerdem befanden sich alle detektierten „Problempixel“ an Küsten und stellten damit Mischpixel (vgl. Abb. 9) dar. Die bisher zur Berechnung von TEM1 verwendeten Strahlungsvariablen bezogen sich auf den Gesamtpixel. Die Oberflächentemperatur und Bodentemperatur sind jedoch nur für den Landanteil des Pixels definiert. Dies wurde in der neuen Version (in der Subroutine *surf*) angepasst, sodass nur noch der Anteil der Strahlung für den Landanteil zur Berechnung benutzt wird. Dadurch kann auch das Oszillieren der Temperaturen unterdrückt (vgl. Abb. 41C) und das neue Bodentemperaturschema mit den erläuterten Anpassungen verwendet werden.

# 6 Auswirkungen der neuen Datensätze und Prozesse

Um die Auswirkungen der neuen Datensätze (vgl. Kapitel 4) und Prozesse (vgl. Kapitel 5) testen zu können, werden zunächst im folgenden Unterkapitel gängige statistische Methoden und Darstellungsformen vorgestellt, die zur Datenanalyse herangezogen werden. Danach wird eine Übersicht der verschiedenen Experimente gezeigt. Anschließend werden die Auswirkungen auf die beiden ausgewählten Untersuchungsgebiete EUR-44 und GER-11 sowohl in Bezug auf die Eingangsdaten als auch in Bezug auf die daraus abgeleiteten Modellparameter dargestellt. Die Validation der Änderungen wird im abschließenden Vergleich zwischen dem bisherigen Modell und den neuen Modellversionen sowie mit Beobachtungsdaten durchgeführt.

## 6.1 Statistische Methoden zur Datenanalyse

Für die Analyse der Unterschiede und Effekte der beschriebenen Daten- und Parametrisierungsänderungen werden die als Stichprobe aufzufassenden Modell- und Beobachtungsdaten in unterschiedlicher Art und Weise gemittelt. Dies kann sowohl auf räumlicher als auch auf zeitlicher Dimension oder gleichzeitig erfolgen. Es wird auch darauf hingewiesen, dass in den Analysen bei der eine räumliche Mittelung stattfindet, die jeweiligen Flächen der einzelnen Pixel durch eine Gewichtung berücksichtigt werden. Zum Angleichen der Koordinatensysteme der Modell- und der Beobachtungsdatensätze werden die gleichen Methoden verwendet, die bereits in Kapitel 4.5.2 beschrieben wurden. In den nachfolgenden Abschnitten sind weitere verwendete statistische Methoden zur Analyse der Daten erläutert. Dabei werden die Standardmethoden nur kurz erwähnt, während auf komplexere Analysemethoden näher eingegangen wird.

### 6.1.1 Univariate Kenngrößen

Die klassischen univariaten Kenngrößen sind Mittelwert  $\bar{x}$  (Formel 6.1), Varianz  $s^2$  (Formel 6.2) und Standardabweichung  $s$  (Formel 6.3). Sie beschreiben die Verteilung und Streuung der Stichprobe einer Variable. Allerdings ist vor allem der Mittelwert stark sensitiv gegenüber Extremwerten und Ausreißern (Wilks, 2006). Die Kenngrößen, wie die verschiedenen Quantile, die durch die Extremwerte nicht so stark beeinflusst werden, beziehen sich auf die Häufigkeiten der Stichprobe und wurden in Kapitel 4.2 erläutert. Dennoch liefern auch diese klassischen Werte einen guten Überblick und Vergleichswert für Stichproben.

$$\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^n x_i \quad (6.1)$$

$$s^2 = \frac{1}{n-1} \cdot \sum_{i=1}^n (x_i - \bar{x})^2 \quad (6.2)$$

$$s = \sqrt{s^2} \quad (6.3)$$

mit:

$i$  = Index eines Stichprobenwertes

$n$  = Stichprobenumfang

$x_i$  = Wert der Variable mit Index  $i$

$\bar{x}$  = Mittelwert einer Variable

$s^2$  = Varianz einer Variable

$s$  = Standardabweichung einer Variable

In dieser Arbeit werden diese Kenngrößen für unterschiedliche räumliche und zeitliche Vergleichsanalysen benutzt. Die Daten werden beispielsweise zu Tages-, Monats-, Saison- oder Jahresmitteln zusammengefasst. Zu beachten ist, dass sich der Saisonmittelwert für den Winter (DJF) jeweils aus dem Wert für Dezember des Vorjahres und den Werten für Januar und Februar des Jahres, für das er dargestellt ist, ergibt. Zu Beginn einer Zeitreihe fehlt somit der Dezemberwert des Vorjahres für den Wintersaisonwert.

Eine Besonderheit bei der Analyse der verschiedenen Variablen stellen die vektoriellen Windkomponenten dar. Sie liegen als  $u$ - und  $v$ -Komponente in den Modell- und in den Validationsdaten vor. Für die Auswertung und Darstellung werden aus diesen beiden Vektoren die Windgeschwindigkeit und -richtung abgeleitet (Papadopoulos et al., 1992; Kraus, 2004; Farrugia & Micallef, 2006):

$$V = \sqrt{u^2 + v^2} \quad (6.4)$$

$$\phi = \arctan\left(\frac{u}{v}\right) \quad (6.5)$$

mit:

$V$  = Windgeschwindigkeit

$\phi$  = Windrichtung

$u$  = Windkomponente von West nach Ost

$v$  = Windkomponente von Süd nach Nord

In Abbildung 42 sind die beiden Komponenten sowie die aus den Formeln 6.4 und 6.5 resultierenden Windwerte schematisch dargestellt. Während die  $u$ -Komponente die Windgeschwindigkeit in West-Ost-Richtung angibt, steht die  $v$ -Komponente für die Geschwindigkeit in Süd-Nord-Richtung. Bei negativen  $u$ - oder  $v$ -Werten ist die Richtung jeweils entgegengesetzt. Die Länge des Vektors, welche dem Betrag beider Komponenten entspricht, ist die „Gesamtwindgeschwindigkeit“ und

der resultierende Winkel  $\phi$  gibt die Richtung an, aus welcher der Wind kommt. Der Windvektor zeigt dagegen in die Richtung, in welche der Wind weht. Umgekehrt können aus den beiden Variablen Windgeschwindigkeit und -richtung die beiden Windkomponenten mit den folgenden Formeln berechnet werden:

$$\begin{aligned} u &= V \cdot \cos\left(\frac{270 - \phi}{\frac{180}{\pi}}\right) \\ v &= V \cdot \sin\left(\frac{270 - \phi}{\frac{180}{\pi}}\right) \end{aligned} \quad (6.6)$$

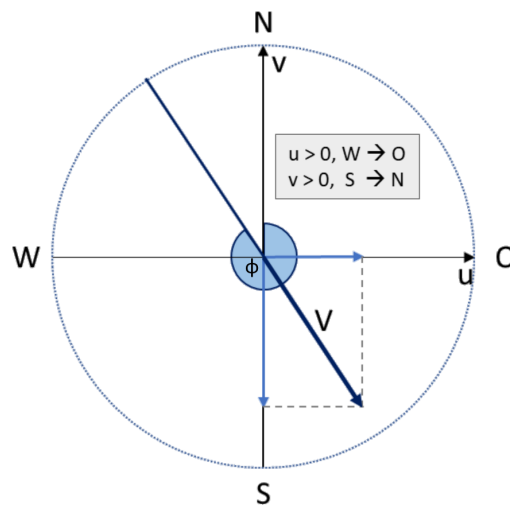


Abb. 42: Schaubild zur Erklärung der Ableitung von Windgeschwindigkeit und -richtung aus der u- und v-Komponente

Um verschiedene Windrichtungen miteinander vergleichen zu können, eignet sich neben der einfachen Differenz, die aufgrund des „kreisförmigen“ Wertebereichs ( $360^\circ = 0^\circ$ ) problematisch sein kann, die Berechnung des Winkels zwischen zwei Windvektoren  $x$  und  $y$ . Da die Windkomponenten  $u$  und  $v$  als Vektor aufgefasst werden können, ist der kleinere Winkel zwischen zwei verschiedenen Windvektoren definiert durch Formel 6.7.

$$\theta = \arccos\left(\frac{x \circ y}{|x| \cdot |y|}\right) \quad (6.7)$$

mit:

$$x = \begin{pmatrix} u_x \\ v_x \end{pmatrix}$$

$$y = \begin{pmatrix} u_y \\ v_y \end{pmatrix}$$

$x \circ y = u_x \cdot u_y + v_x \cdot v_y =$  Skalarprodukt der Vektoren  $x$  und  $y$

$|x| =$  Länge des Vektors  $x =$  Windgeschwindigkeit (vgl. Formel 6.4)

Je größer die Winkeldifferenz, desto größer ist der Unterschied der Windrichtung. Der Wert von Formel 6.7 liegt immer zwischen  $0$  und  $180^\circ$ .



## 6.1.2 Statistische Tests

### Kolmogorov-Smirnov-Test

Viele statistische Hypothesentests haben die Anforderung der Normalverteilung an die Daten. Um diese Annahme für die Daten zu testen, existieren ebenfalls verschiedene statistische Tests. Für die Tests in dieser Arbeit wird der Kolmogorov-Smirnov (oder Kolmogoroff-Smirnoff, KS)-Test verwendet. Dieser liegt in mehreren Varianten vor. Mit dem KS-Anpassungstest kann eine empirische Verteilungsfunktion gegen eine beliebige, aber bekannte Verteilung getestet werden. Sind die Verteilungsparameter der Verteilung nicht bekannt, sondern werden aus der Stichprobe geschätzt, muss auf den sogenannten Lilliefors-Test (Lilliefors, 1967) zurückgegriffen werden, da der KS-Test zu konservative Ergebnisse liefern würde. Der einzige Unterschied zwischen KS- und Lilliefors-Test sind die kleineren kritischen Werte für den Lilliefors-Test, die somit weniger wahrscheinlich anzeigen, dass die Daten normalverteilt sind. Mit dem Zweistichproben KS-Test können dagegen zwei unbekannte Verteilungen auf Gleichheit getestet werden. Bei diesem Test wird der Verteilungstyp allerdings nicht bestimmt.

In allen Fällen lässt sich die Prüfgröße  $T$  als maximale Differenz der beiden Verteilungsfunktionen ermitteln. Zuvor werden die empirischen Daten jedoch standardisiert. Für den benötigten Normalverteilungstest wird die in Formel 6.8 dargestellte  $z$ -Standardisierung verwendet, bei der durch den Einsatz von Mittelwert und Standardabweichung der jeweiligen Stichprobe eine Normierung erfolgt.

$$z_i = \frac{x_i - \bar{x}}{s} \quad (6.8)$$

mit:

$z_i$  = standardisierter Stichprobenwert

$x_i$  = originaler Stichprobenwert

$\bar{x}$  = Mittelwert der Stichprobe

$s$  = Standardabweichung der Stichprobe

Anschließend werden die standardisierten Werte aufsteigend sortiert, um dann die maximale Differenz zwischen der kumulierten empirischen Verteilung  $S$  und der kumulierten Normalverteilung  $N$  mit Mittelwert 0 und Standardabweichung 1 zu ermitteln. Dieser Maximalwert  $T$  (vgl. Formel 6.9) entspricht der Prüfgröße dieses Tests auf Normalverteilung.

$$T = \sup(|N(z_i) - S(z_i)|) \quad (6.9)$$

In dieser Arbeit werden die kritischen Werte für den zweiseitigen Lilliefors-Test benutzt. Der kritische Wert wird für einen Stichprobenumfang  $n$ , der kleiner als 30 ist, in einer Signifikanztabelle (z. B. Conover, 1980) nachgeschlagen. Für größere Stichprobenumfänge wird er durch folgende

Formel 6.10 angenähert:

$$T_{krit} = \frac{krit}{\sqrt{n}} \quad (6.10)$$

mit:

$n$  = Stichprobenumfang

$krit = 1.031$  für  $\alpha = 0.02$

$krit = 0.886$  für  $\alpha = 0.1$

$krit = 0.805$  für  $\alpha = 0.2$

$krit = 0.768$  für  $\alpha = 0.3$

$krit = 0.736$  für  $\alpha = 0.4$

Die Nullhypothese, dass die Daten normalverteilt sind, wird dann verworfen, wenn die Prüfgröße  $T$  größer als der kritische Wert für eine gewählte Irrtumswahrscheinlichkeit  $\alpha$  ist. Um den p-Wert oder kritische Werte für nicht angegebene  $\alpha$  (vgl. Formel 6.10) für den Lilliefors-Test auszurechnen, entwickelten Molin & Abdi (2007) eine polynomische Approximation der Lilliefors-Verteilung:

$$A(T) = \frac{-(b_1 + n) + \sqrt{(b_1 + n)^2 - 4 \cdot b_2 \cdot (b_0 - T^{-2})}}{2 \cdot b_2} \quad (6.11)$$

mit:

$b_0 = 0.37872256037043$

$b_1 = 1.30748185078790$

$b_2 = 0.08861783849346$

und

$$\begin{aligned} P(T) = & -0.37782822932809 + 1.67819837908004 \cdot A(T) - 3.02959249450445 \cdot A(T)^2 \\ & + 2.80015798142101 \cdot A(T)^3 - 1.39874347510845 \cdot A(T)^4 + 0.40466213484419 \cdot A(T)^5 \\ & - 0.06353440854207 \cdot A(T)^6 + 0.00287462087623 \cdot A(T)^7 + 0.00069650013110 \cdot A(T)^8 \\ & - 0.00011872227037 \cdot A(T)^9 + 0.00000575586834 \cdot A(T)^{10} \end{aligned} \quad (6.12)$$

Liegt der errechnete p-Wert  $P(T)$  unter der gewählten Irrtumswahrscheinlichkeit  $\alpha$ , wird die Nullhypothese, dass die Daten normalverteilt sind, verworfen.

### Gepaarter Zweistichproben-t-Test

Liegt eine Normalverteilung der Daten vor, kann mittels des Zweistichproben-t-Test festgestellt werden, ob bzw. um welchen Wert sich zwei Stichproben im Mittel signifikant unterscheiden. Für den ersten Fall liegt ein zweiseitiger Test der Nullhypothese „ $H_0 : \mu_x = \mu_y$  bzw.  $\mu_x - \mu_y = 0$ “ vor. Diese Testhypothese wird auch ungerichtet genannt. Beim zweiten Fall wird die Differenz der beiden Erwartungswerte der Stichproben einseitig gegen einen gewissen Erwartungswert  $E(d)$  getestet (gerichteter Test).

Mit Formel 6.13 kann die Prüfgröße für gepaarte Stichproben (auch verbundene Stichproben genannt) berechnet werden. Dieser Spezialfall liegt vor, wenn die beiden Stichproben voneinander abhängen. Es handelt sich um Beobachtungspaare und daher gilt bei verbundenen Stichproben, dass beide Stichproben die gleiche Anzahl an Merkmalen enthalten. Dies ist beispielsweise der Fall, wenn eine abhängige Variable zu zwei unterschiedlichen Zeitpunkten erfasst wird oder wie im vorliegenden Fall, wenn sich die Werte beider Stichproben jeweils auf den gleichen Zeitpunkt und Pixel beziehen. Die Reihenfolge der Werte innerhalb einer Stichprobe kann hierbei nicht beliebig ausgetauscht werden. Dieser gepaarte Zweistichproben-t-Test entspricht einem Einstichproben-t-Test der Differenzen der beiden Stichproben.

$$T = \left| \sqrt{n} \frac{\bar{d} - E(d)}{s_d} \right| = \left| \sqrt{n} \frac{\frac{1}{n} \sum_{i=1}^n (x_i - y_i) - E(d)}{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (d_i - \bar{d})^2}} \right| \quad (6.13)$$

mit:

$n$  = Stichprobenumfang

$x_i$  = i-ter Wert einer Stichprobe

$y_i$  = i-ter Wert einer anderen Stichprobe

$d_i = x_i - y_i$  = i-ter Differenzwert beider Stichproben

$E(d)$  = Erwartungswert der Differenz gegen den der Hypothesentest prüft (hier: 0)

$\bar{d}$  = Mittelwert der Differenzen

$s_d$  = Standardabweichung der Differenzen

Als Prüfgröße wird der Wert der t-Verteilung, die sich für hohe Freiheitsgrade der Standardnormalverteilung angleicht, für ein gewähltes Signifikanzniveau mit Freiheitsgrad  $df = n - 1$  gewählt. Ist  $T$  größer als die Prüfgröße, wird die Nullhypothese (Differenz = 0, was der Gleichheit der beiden Stichproben entspricht) auf Basis des gewählten Signifikanzniveaus abgelehnt.

Für unverbundene Stichproben würde der Nenner der Prüfgröße nicht der Standardabweichung der Differenzen entsprechen, sondern dem geschätzten Standardfehler der Mittelwertsdifferenzen. Dieser berechnet sich aus der Wurzel der Summe beider Varianzen geteilt durch den jeweiligen Stichprobenumfang. Außerdem ändert sich die Anzahl der Freiheitsgrade auf  $n_x + n_y - 2$  (Bortz & Schuster, 2010).

### Wilcoxon-Test

Liegt dagegen keine Normalverteilung der Daten vor, wird in dieser Arbeit der sogenannte Vorzeichen-Rang-Test von Wilcoxon (Wilcoxon-Test) verwendet (Wilcoxon, 1945), um die Gleichheit der verbundenen Stichproben zu testen. Mittels dieses Tests wird geprüft, ob die Differenzen der beiden Stichproben um den Median, der bei dem ungerichteten Test gleich Null sein sollte, symmetrisch angeordnet sind und beide Stichproben die gleiche Verteilung haben. Da es sich im Gegensatz zum t-Test um einen nichtparametrischen Test handelt, werden für die Berechnung der Prüfgröße zunächst die Ränge der absoluten Differenzen benötigt. Dabei werden nur

die Beobachtungspaare berücksichtigt, deren Differenz nicht gleich Null ist. Die betragsmäßig kleinste Differenz bekommt den Rang 1 zugewiesen. Alle weiteren Rangzahlen werden durch die aufsteigenden Differenzen definiert. Bei mehreren Differenzen mit gleichem Wert (Bindung) wird jedem Differenzwert der Mittelwert aller entsprechender Ränge zugewiesen. Zusätzlich wird jeweils beachtet, ob es sich um eine positive oder negative Differenz handelt, denn für beide Fälle wird jeweils die Summe über die Ränge gebildet. Die Prüfgröße ist dann der kleinere Wert der beiden Rangsummen (vgl. Formel 6.14). Je größer der Unterschied der beiden Rangsummen, desto weniger wahrscheinlich ist die Nullhypothese. Für kleine Stichprobenumfänge ( $n < 25$ ) sind die kritischen Werte einer vorgefertigten Tabellen (z. B. McCornack, 1965) zu entnehmen. Ansonsten lassen sie sich mit der angegebenen Formel 6.14 ableiten. Die Nullhypothese wird verworfen, wenn die Prüfgröße kleiner oder gleich dem kritischen Wert ist (Bortz & Schuster, 2010, Hedderich & Sachs, 2018).

$$\begin{aligned}
 R_i &= \text{rang}(|d_i|), \text{ für } d_i \neq 0 \\
 R_p &= \sum_{i=1}^{n0} R_i, \text{ für } d_i > 0 \\
 R_n &= \sum_{i=1}^{n0} R_i, \text{ für } d_i < 0 \\
 R_{\text{prüf}} &= \min(R_p, R_n) \\
 R_{\text{krit}} &= \frac{R_{\text{prüf}} - \frac{n0 \cdot (n0 + 1)}{4}}{\sqrt{\frac{n0 \cdot (n0 + 1) \cdot (2 \cdot n0 + 1)}{24}}}, \text{ für } n0 > 25
 \end{aligned} \tag{6.14}$$

mit:

$d_i = x_i - y_i =$  Differenz des i-ten Wertepaares

$n0 =$  Anzahl der Wertepaare, deren Differenz ungleich Null ist

$R_i =$  Rang des i-ten Wertes

$R_p =$  Rangsumme aller Werte mit positiver Differenz

$R_n =$  Rangsumme aller Werte mit negativer Differenz

## ANOVA

Eine Analysis of Variance (ANOVA) ist ein weiterer statistischer Test, mit dessen Hilfe man die Unterschiede im Mittelwert von mehr als zwei Datensätzen anhand von Variablen miteinander vergleichen kann. Sie basiert auf dem F-Test, bei dem paarweise die Varianz der verschiedenen Gruppenmittel im Vergleich zur Gesamtvarianz verglichen wird. Dabei wird ein hoher F-Testwert erreicht und somit eine höhere Wahrscheinlichkeit, dass sich die Gruppen wirklich unterscheiden, wenn die Varianz innerhalb der Gruppe kleiner als die Varianz zwischen den verschiedenen Gruppen ist. Eine Mehrwege-ANOVA, die nur bei mehr als zwei unabhängigen Variablen durchgeführt wird, analysiert, wie viel von der Gesamtvarianz der Daten von den einzelnen sogenannten

Faktoren und der Interaktion der Faktoren erklärt werden kann. Die Gesamtvarianz  $QS_{total}$  setzt sich wie bei der einfaktoriellen ANOVA aus der Variation zwischen und innerhalb der einzelnen Gruppen zusammen. Bei mehrfaktoriellen Varianzanalysen besteht die Variation  $QS_{zwischen}$  neben der Quadratsumme der einzelnen Faktoren noch aus der Variation aller möglichen Interaktionen der Faktoren. Für eine zweifaktorielle ANOVA gilt damit Formel 6.15.

$$QS_{total} = QS_{zwischen} + QS_{innerhalb} = (QS_A + QS_B + QS_{AxB}) + QS_{innerhalb} \quad (6.15)$$

Mit den jeweiligen Quadratsummen und ihren jeweils dazugehörigen Freiheitsgraden ergeben sich die mittleren Quadratsummen:

$$\begin{aligned} MS_A &= \frac{QS_A}{df_A} = \frac{n \cdot q \cdot \sum_{i=1}^p (\bar{x}_i - \bar{x})^2}{p - 1} \\ MS_B &= \frac{QS_B}{df_B} = \frac{n \cdot p \cdot \sum_{j=1}^q (\bar{x}_j - \bar{x})^2}{q - 1} \\ MS_{AxB} &= \frac{QS_{AxB}}{df_{AxB}} = \frac{n \cdot \sum_{j=1}^q \sum_{i=1}^p (\bar{x}_{ij} - \bar{x}_i - \bar{x}_j + \bar{x})^2}{(p - 1) \cdot (q - 1)} \\ MS_{innerhalb} &= \frac{QS_{innerhalb}}{df_{innerhalb}} = \frac{\sum_{i=1}^p \sum_{j=1}^q \sum_{k=1}^n (x_{ijk} - \bar{x}_{ij})^2}{N - (p \cdot q)} = \frac{\sum_{i=1}^p \sum_{j=1}^q \sum_{k=1}^n (x_{ijk} - \bar{x}_{ij})^2}{p \cdot q \cdot (n - 1)} \end{aligned} \quad (6.16)$$

mit:

$N = p \cdot q \cdot n$  = Gesamtstichprobengröße

$p$  = Anzahl der Faktorstufen von Faktor A

$q$  = Anzahl der Faktorstufen von Faktor B

$n$  = Anzahl der Stichprobengröße einer Gruppe

$\bar{x}_i$  = Mittelwert der Gruppe i von Faktor A

$\bar{x}_j$  = Mittelwert der Gruppe j von Faktor B

$\bar{x}_{ij}$  = Mittelwert der Gruppe i von Faktor A und Gruppe j von Faktor B

$x_{ijk}$  = k-ter Stichprobenwert in Gruppe i von Faktor A und Gruppe j von Faktor B

$\bar{x}$  = Mittelwert der Gesamtstichprobe

Und aus diesen lassen sich anschließend die zum Testen auf Signifikanz benötigten F-Werte wie folgt berechnen:

$$\begin{aligned} F_A &= \frac{MS_A}{MS_{innerhalb}} \\ F_B &= \frac{MS_B}{MS_{innerhalb}} \\ F_{AxB} &= \frac{MS_{AxB}}{MS_{innerhalb}} \end{aligned} \quad (6.17)$$

Diese F-Werte werden herangezogen, um einen Signifikanzwert  $p$  zu errechnen und damit die Hypothese mit dem vorher definierten Signifikanzniveau  $1 - \alpha$  zu testen. In dieser Arbeit wird ein Signifikanzniveau von 95% verwendet.

Um nicht nur eine Aussage zur Signifikanz, sondern auch zur Stärke der jeweiligen Effekte treffen zu können, wird die Stärke der Effekte  $\eta^2$  als Anteil der jeweiligen Varianz an der Gesamtvarianz berechnet (vgl. Formel 6.18).

$$\eta^2 = \frac{QS}{QS_{tot}} \quad (6.18)$$

Dieser Wert ist nicht zu verwechseln mit dem partiellen  $\eta_p^2$ , bei dem der Nenner aus der Summe der Quadratsumme des jeweils zu untersuchenden Effekts und der Gesamtquadratsumme besteht. Ein  $\eta_p^2$  Wert drückt die Variation aus, die im Datensatz stecken würde, wäre der jeweilige Faktor oder die jeweilige Interaktion nicht im Modell (B. H. Cohen, 2002; Rasch et al., 2010).

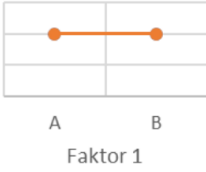
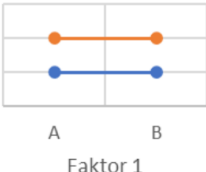
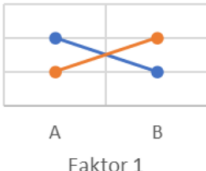
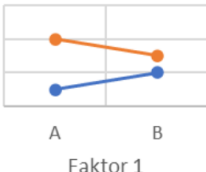
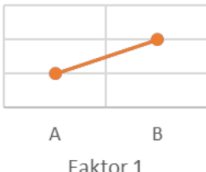
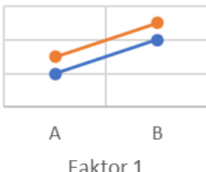
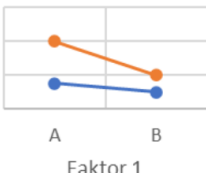
Aus diesem partiellen Wert kann mit Formel 6.19 zusätzlich die Effektstärke nach Cohen ( $f$ ) berechnet werden. Laut J. Cohen (1988) wird bei Effektstärken ab 0.1 von einem schwachen, ab 0.25 von einem mittleren und ab 0.4 von einem starken Effekt gesprochen.

$$f = \sqrt{\frac{\eta_p^2}{1 - \eta_p^2}} \quad (6.19)$$

In dieser Arbeit werden sowohl zwei zwei-Wege- ( $2 \times 2$  und  $2 \times 3$ ) als auch eine drei-Wege-ANOVA ( $2 \times 2 \times 2$ ) durchgeführt, um die Effekte der unterschiedlichen Datensätze und Schalter voneinander getrennt analysieren zu können. Zusätzlich werden sogenannte Interaktionsplots oder Profillinien angefertigt, um die unterschiedlichen Einflüsse und Interaktionen auch graphisch zu veranschaulichen. Je nach Anzahl der Faktorausprägungen können eine unterschiedliche Anzahl an Effekten und Interaktionseffekten auftreten. Für eine ANOVA mit zwei Faktoren mit jeweils zwei Ausprägungen ( $2 \times 2$  ANOVA) existieren beispielsweise acht Möglichkeiten (vgl. Tab. 13). Bei den unterschiedlichen Ausprägungen der Effekte unterscheidet man auch keine, ordinale, semidisordinale oder disordinale Wechselwirkung (Rasch et al., 2010).

Um die zeitliche und räumliche Autokorrelation der Werte der verschiedenen Datensätze zu minimieren, wurde in dieser Arbeit jeweils der Mittelwert pro Pixel und Zeitpunkt aller in die jeweilige ANOVA eingehenden Modelldaten von den einzelnen Datensätzen abgezogen oder mehrmals eine zufällige Stichprobe aus allen Datenpunkten gezogen.

Tab. 13: Übersicht der möglichen Effekte bei einer 2 x 2 ANOVA  
 (mit X: kein signifikanter Effekt bzw. ✓: signifikanter Effekt nachgewiesen)

Möglichkeit	Faktor 1	Faktor 2	Interaktion	Graph
1	X	X	X	
2	X	✓	X	
3	X	X	✓	
4	X	✓	✓	
5	✓	X	X	
6	✓	✓	X	
7	✓	X	✓	
8	✓	✓	✓	

—●— Faktor 2 A  
—●— Faktor 2 B

### 6.1.3 Verschiedene Indizes zur Feststellung der Güte

Um die Güte einer Modellierung festzustellen, wird häufig der sogenannte RMSE verwendet. Dieser Wert gibt an, wie stark der Modellwert durchschnittlich von den Beobachtungswerten abweicht. Berechnet wird der RMSE mit Formel 6.20. Je größer der RMSE-Wert, desto schlechter stimmen Modell- und Beobachtungsdaten überein.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - x_{i,obs})^2} \quad (6.20)$$

mit:

$n$  = Stichprobenumfang

$x_i$  = i-ter Wert der Modelldaten

$x_{i,obs}$  = i-ter Wert der Validationsdaten

Der Nachteil dieses „einfachen“ RMSEs liegt darin, dass die Größenordnung des Index stark von der betrachteten Variable und deren Einheit abhängt. Um die RMSEs von verschiedenen Klimavariablen besser vergleichen zu können, bieten sich deshalb verschiedene Normierungen an. Dadurch werden die absoluten RMSE-Werte, die jeweils in der Einheit der klimatischen Variable vorliegen, in relative RMSE-Werte (rRMSE) umgewandelt. Dieser rRMSE wird mit Formel 6.21 berechnet. Auch hier gilt, je kleiner der Wert, desto besser das Modellergebnis im Vergleich zu den Validationsdaten (Duan et al., 2013; Despotovic et al., 2016).

$$rRMSE = \frac{RMSE}{\bar{x}_{obs}} = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - x_{i,obs})^2}}{\bar{x}_{obs}} \quad (6.21)$$

mit:

$n$  = Stichprobenumfang

$x_i$  = i-ter Wert der Modelldaten

$x_{i,obs}$  = i-ter Wert der Validationsdaten

$RMSE$  = root mean square error (vgl. Formel 6.20)

$\bar{x}_{obs}$  = Mittelwert der Validationsdaten

Weitere Normierungsmethoden des RMSE sehen die Variationsbreite, welche jedoch wie die vorgestellte Version mit dem Mittelwert stark von Ausreißern beeinflusst wird, oder die IQR der Beobachtungsdaten als Zähler von Formel 6.21 vor. Ebenso kann der RMSE mit der Standardabweichung der Beobachtungsdaten normiert werden. Moriasi et al. (2007) nennen diesen Index RMSE-observations Standard deviation Ratio (RSR). Er lässt sich wie folgt berechnen:

$$RSR = \frac{RMSE}{s} = \frac{\sqrt{\sum_{i=1}^n (x_{i,obs} - x_i)^2}}{\sqrt{\sum_{i=1}^n (x_{i,obs} - \bar{x}_{obs})^2}} \quad (6.22)$$



mit:

$n$  = Stichprobenumfang

$x_i$  = i-ter Wert der Modelldaten

$x_{i,obs}$  = i-ter Wert der Validationsdaten

$\bar{x}_{obs}$  = Mittelwert der Validationsdaten

$s$  = Standardabweichung der Validationsdaten

Die Entwickler des Index geben an, dass RSR-Werte kleiner als 0.5 ein sehr gutes Ergebnis sind. Werte über 0.7 werden von ihnen als ungenügend bezeichnet.

Eine weitere Möglichkeit zur numerischen Beurteilung der Modellgüte im Vergleich zu Beobachtungsdaten stellt die sogenannte Nash-Sutcliffe Effizienz (NSE) dar. Es handelt sich dabei um den von Nash & Sutcliffe (1970) entwickelten Index, der mittels der Varianz ausdrückt, wie gut die Modelldaten zu den Beobachtungsdaten passen:

$$NSE = 1 - \frac{\sum_{i=1}^n (x_{i,obs} - x_i)^2}{\sum_{i=1}^n (x_{i,obs} - \bar{x}_{obs})^2} \quad (6.23)$$

mit:

$n$  = Stichprobenumfang

$x_i$  = i-ter Wert der Modelldaten

$x_{i,obs}$  = i-ter Wert der Validationsdaten

$\bar{x}_{obs}$  = Mittelwert der Validationsdaten

Der NSE kann Werte von  $]-\infty;1]$  annehmen, wobei negative Werte ausdrücken, dass der Mittelwert der Beobachtungsdaten ein besserer Prädiktor ist als die Modelldaten. Je höher ein positiver NSE-Wert, desto kleiner ist der Quotient aus der Summe der quadratischen Abweichungen von Modell und Beobachtung sowie der Varianz der Beobachtungswerte und damit desto besser die Modellperformanz. NSE-Werte kleiner als 0.5 werden im Allgemeinen als ungenügend angesehen (Moriassi et al., 2007). Eine Normalisierung des NSE, sodass der Wertebereich des Index zwischen 0 und 1 liegt, entwickelten Nossent & Bauwens (2012):

$$NNSE = \frac{1}{2 - NSE} = \frac{\sum_{i=1}^n (x_{i,obs} - \bar{x}_{obs})^2}{\sum_{i=1}^n (x_{i,obs} - x_i)^2 + (x_{i,obs} - \bar{x}_{obs})^2} \quad (6.24)$$

mit:

$n$  = Stichprobenumfang

$x_i$  = i-ter Wert der Modelldaten

$x_{i,obs}$  = i-ter Wert der Validationsdaten

$\bar{x}_{obs}$  = Mittelwert der Validationsdaten

Der prozentuale Bias (PBIAS) gibt an, wie stark und welche Richtung tendenziell der Fehler zwischen Modell- und Beobachtungsdaten ist:

$$PBIAS = 100 \cdot \frac{\sum_{i=1}^n (x_{i,obs} - x_i)}{\sum_{i=1}^n x_{i,obs}} \quad (6.25)$$

mit:

$n$  = Stichprobenumfang

$x_i$  = i-ter Wert der Modelldaten

$x_{i,obs}$  = i-ter Wert der Validationsdaten

PBIAS-Werte  $< 0$  signalisieren einen positiven Modellbias und positive Werte weisen auf eine Unterschätzung hin. Je höher der Betrag von PBIAS, desto schlechter ist das Modell. Durch die Multiplikation mit dem Faktor 100 kann PBIAS in der Einheit Prozent angegeben werden. Die Größe wird vor allem bei der hydrologischen Modellierung eingesetzt und in der Literatur existieren unterschiedliche Grenzwerte zur Einschätzung der Modellperformanz, die jedoch abhängig von der zu beurteilenden Variable sind (Moriasi et al., 2007; Golmohammadi et al., 2014). Manchmal wird PBIAS auch mit umgedrehten Minuenden und Subtrahenden bei der Subtraktion im Zähler definiert (z. B. Wagena et al., 2018). Dies müsste bei der Interpretation der Vorzeichen der PBIAS-Werte berücksichtigt werden.

## 6.1.4 Zusammenhänge von Variablen

### Korrelationskoeffizient nach Pearson

Um den Zusammenhang von zwei Variablen zu untersuchen, wird häufig zunächst der sogenannte Korrelationskoeffizient nach Pearson ( $r$ ) berechnet (vgl. Formel 4.4b). Er drückt als Quotient der Kovarianz zweier Variablen und dem Produkt der jeweiligen Standardabweichung der Variablen, den linearen Zusammenhang der beiden Variablen aus. Der Koeffizient ist somit einheitslos und kann Werte zwischen -1 und 1 annehmen. Je größer der Betrag, desto stärker ist der Zusammenhang. Wird der Wert 1 erreicht, liegt ein perfekter Zusammenhang vor. Positive Werte weisen darauf hin, dass der Wert einer Variable steigt, wenn gleichzeitig die andere Variable wächst. Umgekehrt bedeutet ein negativer Korrelationskoeffizient, dass sich die Variablen konträr verhalten.

Zudem ist bei der Analyse zu beachten, dass sich die Werte des Korrelationskoeffizienten nicht linear verhalten. Das heißt, dass beispielsweise ein Wert von  $r = 0.6$  keine doppelt so hohen Zusammenhang wie ein Wert von  $r = 0.3$  hat. Zudem sind Korrelationsanstiege um einen Wert im betragsmäßig unteren Wertebereich nicht so stark zu bewerten, als Korrelationsanstiege um den gleichen Wert im oberen Wertebereich. Für eine Mittelwertbildung verschiedener Korrelationskoeffizienten muss daher auf eine Fisher-z-Transformation des Korrelationskoeffizienten mit folgender inversen hyperbolischen Tangensfunktion entsprechender Formel 6.26 zurückgegriffen

werden (Wilks, 2006; Bortz & Schuster, 2010).

$$r_z = \frac{1}{2} \cdot \ln\left(\frac{1+r}{1-r}\right) \quad (6.26)$$

Das Zurücktransformieren nach dem Mitteln passiert analog mit der hyperbolischen Tangensfunktionsformel:

$$r = \frac{e^{2 \cdot r_z} - 1}{e^{2 \cdot r_z} + 1} \quad (6.27)$$

Wie der Mittelwert und die Standardabweichung wird auch die Korrelation stark von Extremwerten und Ausreißern innerhalb einer Stichprobe beeinflusst (Moriassi et al., 2007). Um die Autokorrelation der Werte, die beispielsweise durch in beiden Vergleichsvariablen enthaltene Jahrgänge auftritt, zu reduzieren, kann der Korrelationskoeffizient mit den jeweiligen Anomaliewerten der Variablen berechnet werden. Es wird dann auch manchmal vom sogenannten Anomaly Correlation Coefficient (ACC) gesprochen (Jolliffe & Stephenson, 2012).

### Einfache lineare Regressionsanalyse

Eine weitere häufig verwendete Methode ist die einfache lineare Regression. Diese kann nicht nur zur Trendanalyse verwendet werden, sondern drückt durch ihre Koeffizienten die Ähnlichkeit zweier Stichproben aus. Durch diese statistische Methode wird eine lineare Gerade  $\hat{y} = b \cdot x + a$  mit Steigung  $b$  und der Konstanten  $a$  (y-Achsenabschnitt) so an die Stichprobenpunkte (definiert durch die jeweiligen x- und y-Werte) angepasst, dass die quadrierte Summe der vertikalen Abstände zwischen den Punkten und der aufgestellten Gerade minimiert ist. Die Werte der Stichprobe, welche die x-Werte liefert, werden Prädiktoren genannt, da diese benutzt werden können, um Näherungswerte für die abhängigen y-Werte (Kriterium) zu berechnen. Liegt die Steigung nahe bei der Zahl 1 und der Achsenabschnitt nahe bei Null, dann ähneln sich die Stichproben stark. Die Formeln 6.28a und 6.28b liefern die entsprechenden Berechnungsformeln für die beiden Koeffizienten der Gerade. Die Übereinstimmung der beiden Stichproben wird durch den Determinationskoeffizient ( $r^2$ ) ausgedrückt, welcher als Anteil der modellierten (erklärte Varianz) an der in der Stichprobe gemessenen Varianz (Gesamtvarianz) berechnet wird (Formel 6.28c) und auch als Bestimmtheitsmaß bezeichnet wird. Er kann Werte von 0 bis 1 annehmen, wobei das Maximum nur bei Gleichheit der beiden Varianzen und damit bei identischen Stichproben erreicht wird. Außerdem entspricht der Koeffizient  $r^2$  der quadrierten Korrelation (Formel 4.4b) der beiden Stichproben (Wilks, 2006; Bortz & Schuster, 2010).

$$b = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (6.28a)$$

$$a = \bar{y} - b \cdot \bar{x} \quad (6.28b)$$

$$r^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (6.28c)$$

mit:

$b$  = Steigung der linearen Trendgeraden

$a$  = y-Achsenabschnitt der linearen Trendgeraden

$n$  = Stichprobenumfang

$x_i$  = i-ter Wert der unabhängigen Stichprobe

$y_i$  = i-ter Wert der abhängigen Stichprobe

$\bar{x}$  = Mittelwert aller  $x_i$

$\bar{y}$  = Mittelwert aller  $y_i$

$r^2$  = Determinationskoeffizient, Bestimmtheitsmaß

$\hat{y}_i = b \cdot x_i + a$  = modellierter y-Wert für Index  $i$

### 6.1.5 Weitere Gütemaße

Zusätzlich zu den oben genannten Indizes bietet die sogenannte Konfusionsmatrix und die daraus ableitbaren Gütekriterien eine weitere Möglichkeit verschiedene Modellversionen und Beobachtungsdaten miteinander zu vergleichen. Dieses sehr häufig bei der Validation von Klassifikationsalgorithmen eingesetzte Verfahren eignet sich aber auch für klimatische Variablen. Grundsätzlich werden bei dem Verfahren zwei Datensätze miteinander verglichen. Nachfolgend sind die verschiedenen Gütemaße mit ihren jeweiligen Formeln für zwei und mehrklassige Konfusionsmatrizen dargestellt.

Die Produzenten-Genauigkeit wird im Englischen Producer Accuracy (PA) oder auch „recall“ oder True Positive Rate (TPR), „hit rate“ oder „sensitivity“ genannt. Sie gibt das Verhältnis der positiv richtigen Vorhersagen im Vergleich zu allen positiven Werten an:

$$PA = \frac{TP}{TP + FN} = \frac{TP}{P}$$

$$PA_j = \frac{x_{i,i}}{\sum_{j=1}^n x_{i,j}} \quad (6.29)$$

mit:

$n$  = Anzahl der Klassen

$j$  = Spaltenindex = entsprechende Klasse in den Beobachtungsdaten

$i$  = Zeilenindex = entsprechende Klasse in den Modelldaten

$x_{i,j}$  = Anzahl der Werte in Klasse  $i$  bei den Modelldaten und gleichzeitig in Klasse  $j$  bei den Beobachtungsdaten

Als Nutzer-Genauigkeit, User Accuracy (UA) oder „Consumer’s Accuracy“, „precision“ oder Positive Predicted Rate (PPR) hingegen wird das Verhältnis der positiv richtigen Werte im Vergleich zur Gesamtzahl der positiv vorhergesagten Werte bezeichnet:

$$\begin{aligned}
 UA &= \frac{TP}{TP + FP} \\
 UA_i &= \frac{x_{i,i}}{\sum_{j=1}^n x_{i,j}}
 \end{aligned}
 \tag{6.30}$$

Die Gesamtgenauigkeit, die auch Overall Accuracy (OA) oder nur „accuracy“ genannt wird, drückt das Verhältnis aller richtig klassifizierten Pixel zur Gesamtanzahl aller Pixel aus:

$$\begin{aligned}
 OA &= \frac{TP + TN}{TP + FP + FN + TN} \\
 OA &= \frac{\sum_{k=1}^n x_{k,k}}{\sum_{i=1}^n \sum_{j=1}^n x_{i,j}} = \frac{\sum_{k=1}^n x_{k,k}}{n}
 \end{aligned}
 \tag{6.31}$$

Als zusätzliche Kennwerte gibt es die durchschnittliche Genauigkeit Average Accuracy (AA) und die mittlere Genauigkeit Mean Accuracy (MA). Wobei die durchschnittliche Genauigkeit dem Mittelwert aller PA entspricht, während die mittlere Genauigkeit der Mittelwert von OA und AA (vgl. Formel 6.32) ist.

$$MA = \frac{OA + AA}{2}
 \tag{6.32}$$

Die jeweils zugehörigen Fehler OE (Omission Error, False Negative Rate) und CE (Commission Error, False Discovery Rate) ergeben sich aus 1-PA bzw. 1-UA. Außerdem gibt es noch die weniger häufig angegebenen und nur bei binären Konfusionsmatrizen möglichen Werte: true negative rate (TNR=TN/FP+TN=TN/N), false positive rate/false alarm rate/fallout (FPR=1-TNR=FP/FP+TN =FP/N), negative predictive value (NPV=TN/FN+TN) und false omission rate (FOR=FN/FN+TN =1-NPV). Abbildung 43 verdeutlicht den Zusammenhang der verschiedenen Indizes.

Der häufig fälschlicherweise als Level der Übereinstimmung deklarierte Kappa-Koeffizient ( $\kappa$ ), repräsentiert das Übereinstimmungslevel korrigiert von der zufällig richtigen Zuweisung.

$$\kappa = \frac{OA - c}{1 - c} = \frac{OA - \left( \sum_{k=1}^n \left( \sum_{i=1}^n x_{i,k} \cdot \sum_{j=1}^n x_{k,j} \right) / n^2 \right)}{1 - \left( \sum_{k=1}^n \left( \sum_{i=1}^n x_{i,k} \cdot \sum_{j=1}^n x_{k,j} \right) / n^2 \right)}
 \tag{6.33}$$

Ein Kappa-Wert kleiner als 0.4 gilt als schlechte Übereinstimmung, während Werte über 0.8 eine ausgezeichnete Performanz ausdrücken. Ein Kappa von -1 würde die perfekte Umkehrung bedeuten. Trotzdem ist die Interpretation des Wertes schwierig, da es ähnlich wie beim Korrelationskoeffizienten keine einheitliche Richtlinien zur Bewertung gibt und verschiedene Quellen unterschiedliche Grenzwerte für Kappa definieren. Es gibt außerdem verschiedene Abwandlungen des Kappa-Koeffizienten (vgl. Pontius & Millones, 2011) und Formeln zum Schätzen der Varianz sowie einen Signifikanztest für Kappa (vgl. Smits et al., 1999).

		Validationsdaten		Zeilensumme		
		P	N			
Modelldaten	P	TP	FP	TP+FP	UA=PPV=TP/(TP+FP)	CE=FDR=FP/(TP+FP)
	N	FN	TN	FN+TN	FOR=FN/(FN+TN)	NPV=TN/(FN+TN)
Spaltensumme		P=TP+FN	N=FP+TN	n=TP+FN+FP+TN		
		PA=TPR=TP/P	FPR=FP/N		OA=(TP+TN)/n	
		OE=FNR=FN/P	TNR=TN/N		ERR=1-OA=(FP+FN)/n	
					Kappa Koeffizient=(OA-c)/(1-c)	

Abb. 43: Schematischer Aufbau einer Vierfeldertafel und die jeweiligen Indizes  
 P: Anzahl der Werte unterhalb des Mittelwerts / N: Anzahl der Werte oberhalb des Mittelwerts  
 (Erklärung der einzelnen Indizes und deren jeweilige Berechnungsformel siehe Text.)

In der Literatur sind die Bedeutung, der Vorteil gegenüber anderen Indizes und die Aussage des Kappa-Werts umstritten. Laut Pontius & Millones (2011) oder Olofsson et al. (2014) liefert der Kappa-Koeffizient keine zusätzlichen und hilfreichen Informationen, da er unter anderem sehr stark mit der OA korreliert. Andere Forscher hingegen argumentieren, dass der Kappa-Index im Vergleich zur OA bessere Aussagen zur Genauigkeit einer Methode liefern kann, weil hierbei die zufällige Übereinstimmung eliminiert wird. Der Kappa-Koeffizient gilt daher als ein eher konservatives Maß zur Einschätzung der Ähnlichkeit (P. F. Watson & Petrie, 2010; Jolliffe & Stephenson, 2012; McHugh, 2012).

Da die in dieser Arbeit zu analysierenden Variablen entweder metrisch skalierte klimatologische Werte oder nominal skalierte klimatologische Indexdaten (z. B. Sommertag, Frosttag, Regentag, ...) sind, muss die Konfusionsmatrix zunächst auf eine sogenannte Vierfeldermatrix reduziert werden. In Abbildung 43 ist eine solche Matrix mit den jeweiligen Berechnungen der Indizes schematisch dargestellt. Die metrischen Werte werden, wie z. B. in Ring et al. (2018) beschrieben, jeweils in die Klasse größer bzw. kleiner gleich dem jeweiligen Datensatzmittelwert eingeteilt.

Insgesamt muss angemerkt werden, dass bei allen Konfusions-Indizes die „Distanz“ zwischen den Datensätzen bzw. die „Größe des Fehlers“ nicht berücksichtigt wird. Je nach Wahl der Klassengrenzen können unterschiedliche Gütemaße erreicht werden. Dennoch erlauben sie eine gute und schnelle Einschätzung der Ähnlichkeit von verschiedenen Stichproben bzw. im Fall der Verwendung von Modell- und Beobachtungsdaten der Qualität der Vorhersage.

## 6.2 Übersicht der Modellexperimente

In dieser Arbeit wurden zunächst die Landoberflächenkarten, die von REMO als Startbedingungen benötigt werden, verändert. Sowohl der Datensatz der Topographie (GTOPO) als auch der Bodendatensatz (FAO) wurde durch neue Datensätze ausgetauscht. Um die aus diesen Veränderungen der konstanten Oberflächendaten resultierenden Auswirkungen auf Modellierungsergebnisse zu testen, ergibt sich der unvollständige Experimentplan, der in Abbildung 44 dargestellt ist. Er wurde so aufgebaut, dass trotz der Lücken durch den paarweisen Vergleich die Effekte der jeweiligen Änderung nachvollzogen werden können. Durch den in Kapitel 4.4.1 beschriebenen Vergleich der Topographiedaten und der daraus folgenden Vorauswahl sowie dem Vergleich der verschiedenen Bodendatensätze (vgl. Kapitel 4.4.2) lassen sich zwei Topographien, vier Bodendatensätze und zwei Ansätze bei der Berechnung der Bodenparameter (vgl. Kapitel 5) im Plan auf die drei Unterscheidungskategorien aufteilen. Dadurch ergeben sich für diesen Experimentplan neun Läufe, die mit dem bisherigen REMO-Modell (id01), welcher als Validationslauf für alle folgenden Änderungen dient, verglichen werden können. Auch die neun Modellläufe werden mit id01 als Präfix bezeichnet, um diese von den anderen Experimentplanläufen abzugrenzen. Die in der Abbildung jeweils ausgewiesene Suffixe (drei Buchstaben) definieren den verwendeten Topographie- und Bodendatensatz sowie die LSOIL-Einstellung der jeweiligen Modellversion. Die Analyse wurde für das EUR-44- und für das GER-11-Gebiet durchgeführt, sodass auch die verbesserte räumliche Auflösung der Datensätze in der Analyse berücksichtigt werden kann.

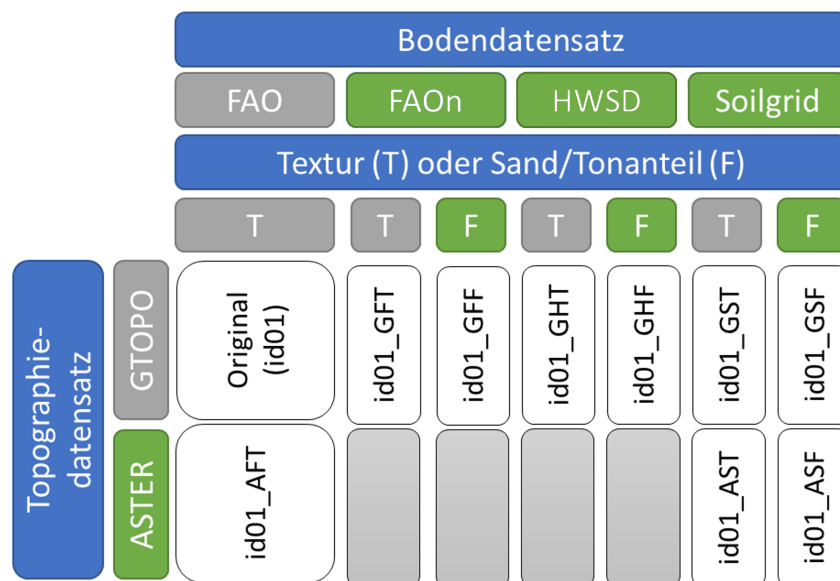


Abb. 44: Experimentplan 1: Übersicht der durchgeführten Läufe zum Vergleich verschiedener Topographie- und Bodendatensätze  
Erklärung der Abkürzungen und Unterschiede zwischen den Modellläufen siehe Text.  
(dunkelblau: Kategorie / grau: bisherige Standardeinstellung / grün: neue Einstellung)

Im zweiten Teil der Arbeit wurden neue Parametrisierungen sowie Erweiterungen für die Bodenvariablen und ein neues numerisches Schema bei der Bodentemperaturberechnung getestet (vgl. Kapitel 5). Der in Abbildung 45 dargestellte zweite Experimentplan zeigt die zur Analyse durchgeführten Modellläufe. Die vier Buchstaben lassen sich jeweils aus der Einstellung des

dazugehörigen Schalters in der Namelist ableiten. Der erste Buchstabe steht für die bereits im ersten Versuchsplan eingebundene Einstellung, ob die Textur (T) oder der Sand- und Tonanteil (F) für die Bodenparametrisierung verwendet werden soll. Der zweite Buchstabe drückt aus, ob das Bucket- (F) oder das neue 5Layer-Schema (T) für die Berechnung der hydrologischen Prozesse im Boden verwendet wird. Da das 5Layer-Schema in einer weiteren Dissertation ausführlich analysiert und noch weiter verbessert wird, ist bei diesem zweiten Experimentplan dieser Schalter immer auf F gestellt. Beim dritten Schalter handelt es sich um die Entscheidung, ob (wie bisher) mit in die Tiefe konstanten Bodenparametern (F) oder mit den neu eingebauten fünfschichtigen Bodenparametern (T) im Modell gerechnet wird. Hierbei ist zu beachten, dass die Kombinationen TFFF und TFTF sowie TFFT und TF TT identische Modellergebnisse liefern, da für die Textur (erster Schalter) wie bisher nur die Bodenparameter der Oberfläche vorliegen und somit nur in die Tiefe konstante Texturwerte benutzt werden können. Der letzte Schalter steht für das verwendete numerische Schema zur Berechnung der Bodentemperaturen (F für das alte Schema, T für das Crank-Nicolson Schema). Für diesen Experimentplan ergeben sich entsprechend acht Läufe, von denen sechs zu numerisch unterschiedlichen Modellergebnissen führen. Der Modelllauf TFFF, der wie bei dem ersten Experimentplan als Validierungslauf dient, stimmt allerdings nicht mit dem Originallauf aus dem ersten Plan überein. Weil für den ersten und dritten Schalter ein Bodendatensatz benötigt wird, bei dem beide Einstellungen möglich sind, und wegen des in Kapitel 5.2 erläuterten Problems mit dem funktionierenden neuen Bodentemperaturschema (vierter Schalter), das zum Vergleichen der Testläufe gleich sein soll, muss auf die verbesserte Modellversion (vgl. Kapitel 5.2) zurückgegriffen werden. Daher wird dieser Experimentplan mit dem Modellcode id03 (für genauere Unterschiede zwischen den Modellversionen vgl. Abb. 46) durchgeführt. Auch hier wurde sowohl die gröbere Auflösung mit dem EUR-44- als auch die feinere Auflösung mit dem GER-11-Gebiet getestet.



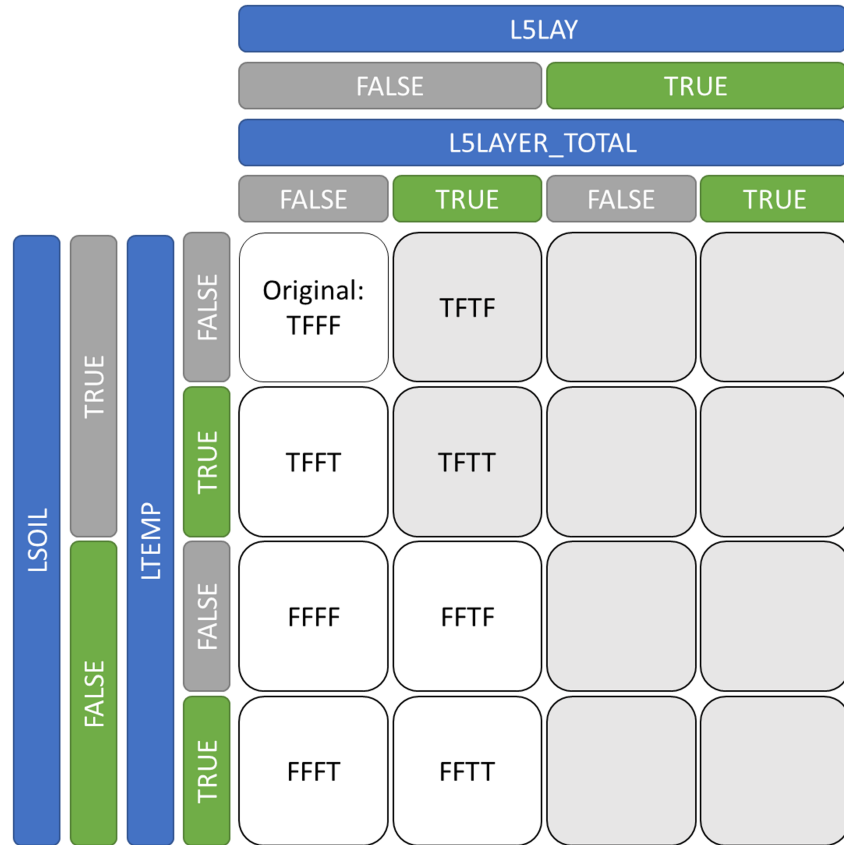


Abb. 45: Experimentplan 2: Übersicht der durchgeführten Läufe zum Vergleich der verschiedenen Modellschalter  
 Erklärung der Abkürzungen und Unterschiede zwischen den Modellläufen siehe Text.  
 (dunkelblau: Schaltereinstellung im Jobskript / grau: bisherige Standardeinstellung / grün: neue Einstellung)

Im nächsten Teil der Arbeit sind noch weitere Änderungen und neue Parametrisierungen, wie beispielsweise die verschiedenen PTFs und andere kleine Änderungen (vgl. Kapitel 5) sukzessiv und zusätzlich zu bereits in den vorherigen Experimentplänen entwickelten Änderungen ins Modell eingebaut worden. Ein paarweiser Vergleich zur vorherigen Modellversion kann deshalb jeweils den Effekt der Neuerung zeigen (vgl. Abb. 46).

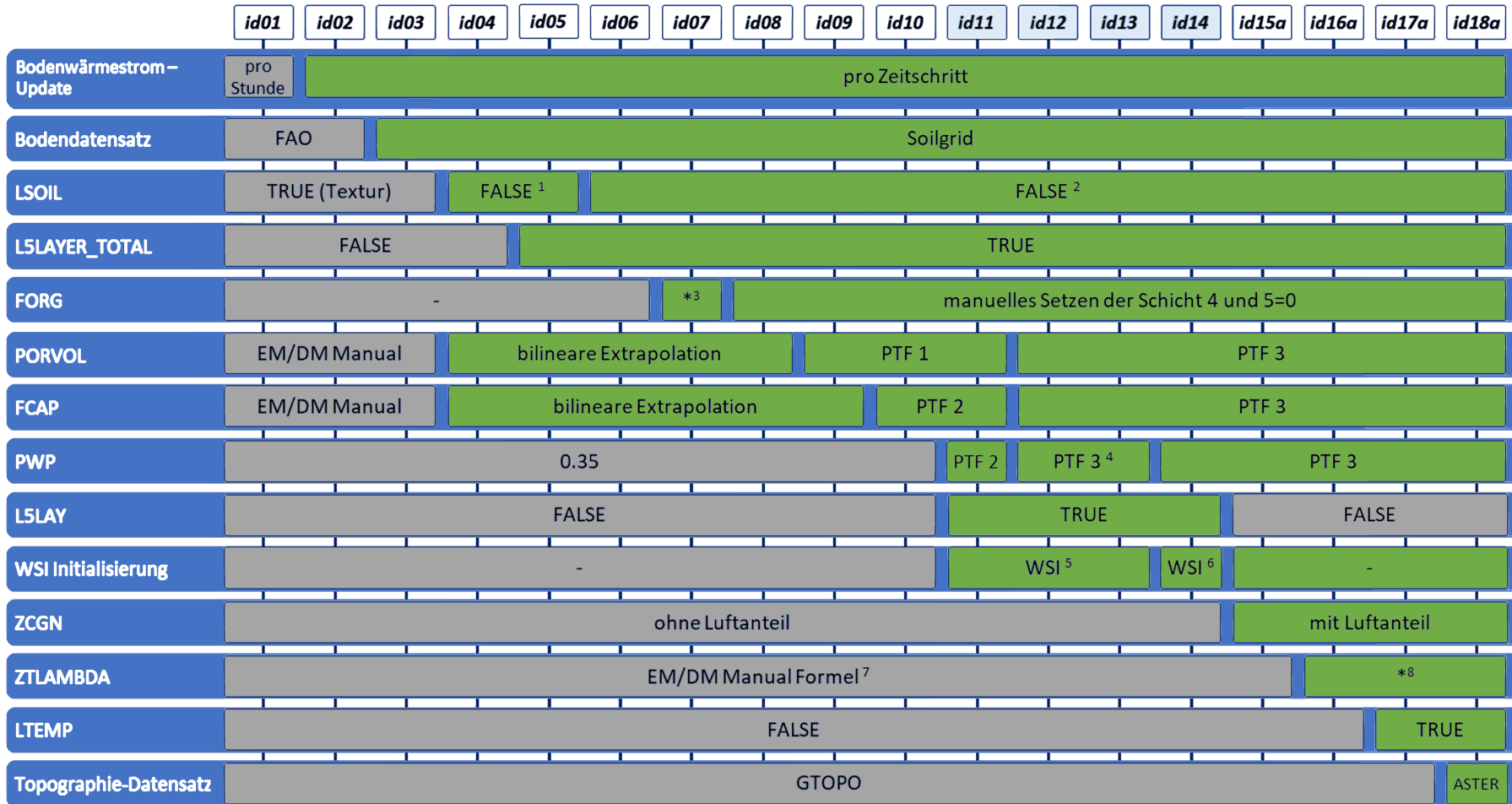


Abb. 46: Experimentplan 3: Übersicht der durchgeführten Läufe zum Vergleich der verschiedenen Parametrisierungen und kleiner Änderungen (dunkelblau: Kategorie / grau: bisherige Standardeinstellung / grün: neue Einstellung)

<sup>1</sup>: bilineare Extrapolation mit Sand- / Tonanteil (eigene PTF); <sup>2</sup>: Sand- /Tonanteil für de Vries (1963) und Farouki (1981b); <sup>3</sup>: organischer Anteil für alle Schichten (nach vertikaler Interpolation); PTF 1: Lawrence & Slater (2008); PTF 2: Adhikary et al. (2008); PTF 3: Saxton & Rawls (2006); <sup>4</sup>: in *surf* PWP mit PTF 3, in *vdiff* PWP=0.35; <sup>5</sup>: WSI wird mit konstanten Wassergehalt initialisiert; <sup>6</sup>: WSI wird linear von WS auf alle fünf Schichten verteilt; <sup>7</sup>: Formel 3.3; <sup>8</sup>: Lawrence & Slater (2008)

Aufgrund der Vielzahl der unterschiedlichen Einstellungen und der damit verbundenen Rechenzeit und Speicherkapazität wurden alle hier aufgelisteten Läufe zunächst nur für das Jahr 2000 als Kaltstart für EUR-44 gerechnet. Für einzelne Modellstufen wurde zudem ein Lauf für GER-11 realisiert, um auch die Auswirkung bei höheren Auflösungen zu testen. Anhand der in Abbildung 46 dargestellten Kategorien (Bodenwärmestrom-Update (vgl. Kapitel 5.2), Bodendatensatz, LSOIL, L5LAYER\_TOTAL, FORG, PORVOL, FCAP, PWP, L5LAY, WSI-Initialisierung, ZCGN, ZTLAMBDA, LTEMP und Topographie-Datensatz) lassen sich die Änderungen von Lauf id02 bis id18a einteilen. Ein Lauf wird jeweils mit den untereinander stehenden Einstellungen für die jeweilige Kategorie betrieben. Beispielsweise lässt sich aus Abbildung 46 ablesen, dass durch den Wechsel von Lauf id03 auf id04, LSOIL vom bisherigen Wert TRUE auf FALSE gesetzt wird und sich dadurch die Berechnungsmethode von PORVOL und FCAP ändert (vgl. Kapitel 5.1). Gleichzeitig ist zu erkennen, dass Modellversion id04 wie Version id03 Soilgrid und GTOPO als Eingangsdaten, keinen organischen Anteil (FORG) und keine Änderungen an den thermalen Parametern (ZCGN und ZTLAMBDA) verwendet. Die Reihenfolge der Veränderungen wurde so angepasst, dass alle gewählten Kombinationen lauffähig sind. Zum Beispiel werden das Wärmestrom-Update und der neue Bodendatensatz sehr früh eingeführt, da ohne diese Änderung andere Abwandlungen wie das Verwenden der Schalter „LSOIL“ oder „LTEMP“ oder das Einführen von PTFs nicht möglich wäre. Die Läufe id11 bis id14 sind in der Abbildung farblich hervorgehoben, weil diese mit dem 5Layer-Schema angetrieben wurden. Weil dieses Bodenwasserschema in einer eigenen Dissertation näher analysiert wird, wird in dieser Arbeit nur vereinzelt ein Vergleich auf id11 erfolgen, um die vorherigen Auswirkungen auf diesen Lauf zu zeigen. Ab Version id15 wurden die Läufe mit dem Zusatz a versehen, um diesen „Rückschritt“, der durch das Verwenden des ursprünglichen Bucket-Schemas statt des 5Layer-Schemas zur Berechnung des Bodenwasserhaushaltes entsteht, ganz klar anzudeuten. Wie sich die jeweiligen Einstellungen auswirken, wird in den nachfolgenden Kapiteln erläutert.

Im letzten Schritt wurden die entscheidenden Meilensteine bei der Entwicklung von der bisherigen REMO- zur neuen REMOLAND-Version, anhand der vorausgegangenen Analysen der einzelnen Änderungen als Warmstart und für einen längeren Zeitraum durchgeführt:

- EUR-44 id01 (Referenzlauf)
- GER-11 id01 (Referenzlauf)
- GER-11 id06: Lauf mit neuem Bodendatensatz, FFTF, ohne organischen Anteil, mit neuer Parametrisierung für RGCGN und TLAMBDA
- GER-11 id15a: Lauf mit neuem Bodendatensatz, FFTF, mit organischem Anteil, mit neuer Parametrisierung für RGCGN und TLAMBDA, mit PTF für PORVOL, FCAP und PWP
- GER-11 id18a: Lauf mit neuem Bodendatensatz, FFTT, mit organischem Anteil, mit neuer Parametrisierung für RGCGN und TLAMBDA, mit PTF für PORVOL, FCAP und PWP, mit neuer Parametrisierung von ZTLAMBDA, mit neuem Topographiedatensatz

Um die Auswirkungen der höheren Auflösung und die Modelländerungen für eine größere Stichprobe untersuchen zu können, wurden diese fünf Läufe als Warmstart von 2000-2018 durchgeführt. Beide Läufe, welche die bisherige REMO-Version (id01) verwenden, können damit

als Referenz sowohl für die weiteren Läufe als auch für die Kaltstarts im Experimentplan 1 dienen. Als erster Zwischenschritt wurde id06 ausgewählt. Diese Version beinhaltet den neuen aufbereiteten Bodendatensatz SoilGrids, verwendet bereits den Sand- und Tonanteil statt die Textur, enthält aber noch keinen organischen Anteil. Zudem ist der Schalter „L5LAYER\_TOTAL“ aktiv und die trockenen Werte für Wärmekapazität und -leitfähigkeit werden mit den Ansätzen nach de Vries (1963) und Johansen (1975) berechnet. Der zweite Zwischenschritt wird nach dem Einfügen des organischen Anteils und der Einführung der Pedotransferfunktion von Saxton & Rawls (2006) für FCAP, PWP und PORVOL sowie der Anpassung des ZCGN-Werts um den Luftanteil gesetzt, was Modelllauf id15a entspricht. Beim finalen Lauf (id18a) mit allen Änderungen wird der neue Topographiedatensatz ASTER und das neue Bodentemperaturschema (LTEMP=T) verwendet. Zudem wird die Wärmeleitfähigkeit nach Lawrence & Slater (2008) berechnet.

### **6.3 Auswirkungen auf die Eingangsdaten**

Die größten räumlichen Auswirkungen auf die in REMO verwendeten Eingangsdaten haben die Änderungen des ersten Experimentplans (vgl. Abb. 44). Der bereits in Kapitel 4.4 für die topographischen Variablen FIB, VAROR und AZ0 sowie für die neuen Bodenparameter SAND1, CLAY1 und FORG1 für alle CORDEX-Regionen ausgeführte Vergleich der Datensätze wird nachfolgend detaillierter für die beiden Untersuchungsgebiete EUR-44 und GER-11 sowie für die in den Modellläufen dieser Arbeit verwendeten Datensätze beschrieben.

#### **6.3.1 Topographie-Daten**

In Abbildung 47 ist die Höhe des bisherigen Topographiedatensatzes GTOPO und die Differenz des neuen Topographiedatensatzes ASTER zu GTOPO in 0.44° dargestellt. Die in Kapitel 4.4.1 näher erläuterte positive Topographieanomalie in Grönland von ASTER ist auf der Differenzkarte zu GTOPO gut zu erkennen. Zusätzlich fallen die teilweise um mehr als 200 m tieferen Küstenpixel auf. Die höchsten negativen Differenzwerte finden sich an der stark zerklüfteten norwegischen Küste sowie den spanischen und griechisch Inseln, da sich hier der große Unterschied der räumlichen Auflösung von 30 m im Vergleich zu 1 km am stärksten auf den auf 50 km aggregierten Mittelwert pro Pixel auswirkt. Insgesamt sind Differenzen von zumeist  $\pm 60$  m im Untersuchungsgebiet festzustellen. Es ist zu erkennen, dass ASTER bei höher liegenden gebirgigen Regionen (außer in Afrika und Norwegen) höhere Werte, ansonsten aber eher etwas niedrigere Werte aufweist.

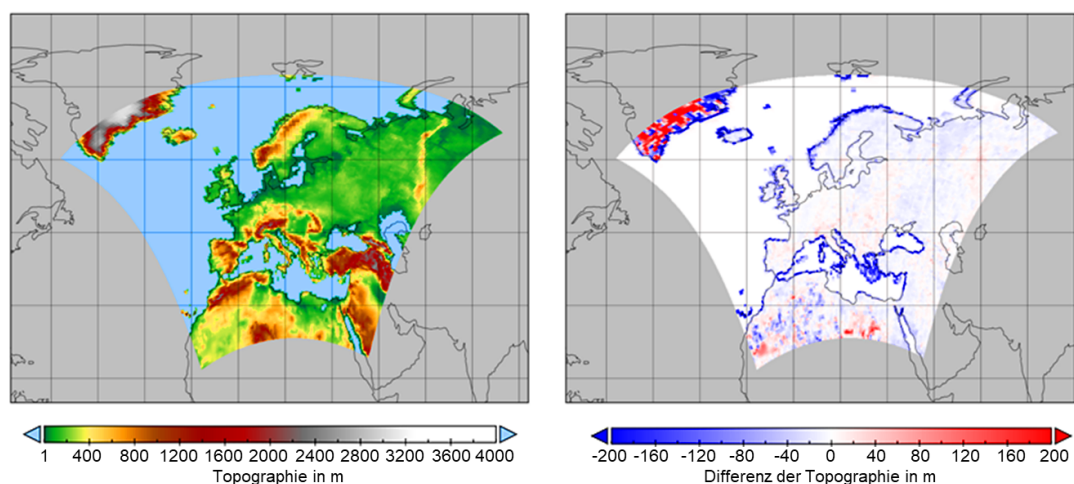


Abb. 47: Topographie-Daten im Vergleich für EUR-44  
links: GTOPO, rechts: ASTER-GTOPO

Diese Tendenzen bestätigen sich auch im GER-11-Gebiet (vgl. Abb. 48). Höher liegende Regionen, wie der westliche Teil der Karpaten, die nördlichen Apenninen oder der Schwarzwald, werden mit ASTER als noch höher ausgewiesen. Allerdings sieht man in den sehr gebirgigen Alpen durch die verbesserte Auflösung des GER-11-Gebiets im Vergleich zum EUR-44-Gebiet auch vermehrt Pixel, die niedriger liegen als mit GTOPO. Ein weiterer sehr markanter flächig zusammenhängender Unterschied der beiden Datensätze liegt an der Westküste Kroatiens. Zudem sind vereinzelt an der Südküste von Frankreich nahe Marseille und Nizza sowie bei Genua größere Unterschiede zu erkennen. Hier werden negative Differenzen von 150–200 m erreicht (vgl. Abb. 48A). Die „fehlerhafte“ Landseemaske vor allem zwischen Dänemark und Schweden wird durch den unteren Grenzwert von 1 m (bei FIB und VAROR) in der Farbgebung der Abbildung sowie der nur grob aufgelösten eingezeichneten Landmaske hervorgerufen und ist kein Fehler in den topographischen Daten. Bei der Varianz der Topographie (vgl. Abb. 48B) liegen die größten Unterschiede in den Bereichen der Alpen und Mittelgebirgen, was durch die großen Differenzen der absoluten Höhe in den beiden Datensätzen ausgelöst wird. Das gleiche Differenzmuster ergibt sich für die Rauigkeitslänge (vgl. Abb. 48C), die aus den Varianzdaten abgeleitet wird.

Da die Varianz der Orographie in REMO zur Korrektur des Beta-Koeffizienten beim verbesserten Arno-Schema (vgl. Formel 3.4) verwendet wird, wirken sich die höheren Werte von ASTER auf die Menge des Oberflächenabflusses und damit einhergehend auf die Infiltration aus. Weil die in REMO genutzte Formel 3.4 jedoch durch eine Obergrenze beschränkt wird, sind die größten Unterschiede zwischen Abflusswerten der Läufe im mittleren Wertebereich der Absolutwerte der Varianz mit den höchsten Varianzdifferenzen zu finden. Durch das veränderte Abflussverhalten hat die Änderung der Topographie damit auch einen indirekt Einfluss auf den Bodenwasserhaushalt, welcher wiederum Einfluss auf die Bodentemperaturen hat. Dass sich die Rauigkeitslängen stark unterscheiden, macht sich später im Modell ebenfalls bemerkbar, denn die Rauigkeitslänge ist ein entscheidender Parameter für die Landoberflächen-Atmosphären Interaktion (Nelli et al., 2020). Sie beeinflusst die Austauschkoefizienten und damit den turbulenten Austausch von Impuls, Energie und Feuchte zwischen der unteren Atmosphäre und der Erdoberfläche (Reijmer et al.,

2004). Daher wirkt sich die Rauigkeitslänge auch besonders auf die Windgeschwindigkeiten, aber auch auf die Wärmeflüsse aus. Höhere Werte der Rauigkeitslänge stehen für einen erhöhten turbulenten Austausch.

Zusammenfassend geht aus dem Vergleich zwischen dem bisherigen GTOPO und dem neuen ASTER Topographiedatensatz hervor, dass es trotz der Interpolation auf 50 km bzw. 12 km Pixelgröße sehr deutliche Unterschiede gibt. Der Betrag der Differenz fällt zudem regional sehr unterschiedlich aus. Besonders groß sind die Unterschiede in gebirgigen Regionen und an Küsten im Untersuchungsgebiet.

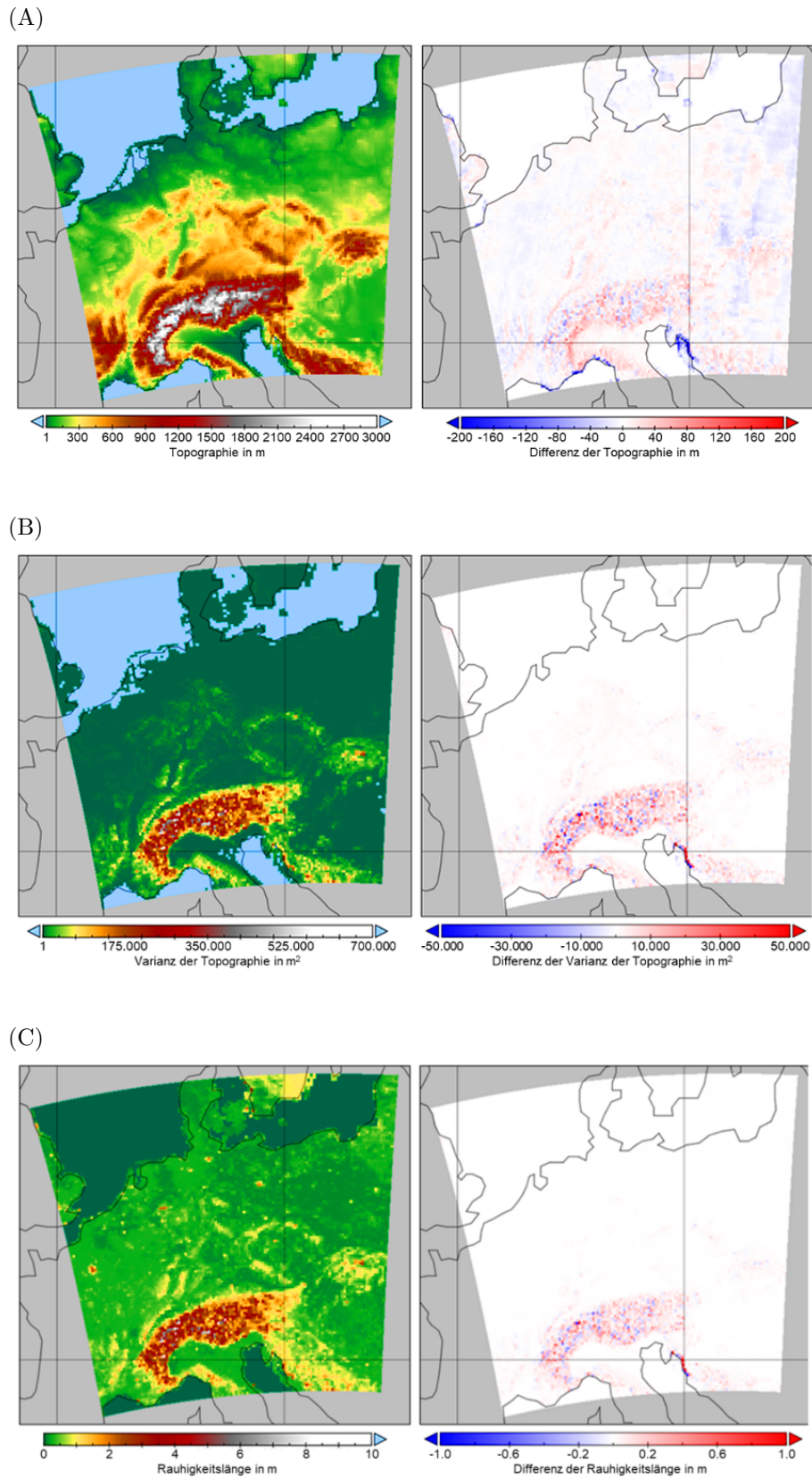


Abb. 48: Topographie-Daten im Vergleich für GER-11

(A) FIB

(B) VAROR

(C) AZ0

links: GTOPO, rechts: ASTER-GTOPO

### 6.3.2 Bodendaten

Wie in Kapitel 4.3.2 bereits gezeigt wurde, differenzieren die alten FAO-Bodendaten nur zwischen fünf Texturklassen und der Klasse Torf, was den Vergleich mit den jeweils 12 Texturklassen der anderen Bodendatensätze erschwert. Die Reklassifikation der 12 Texturklassen auf die fünf bisherigen Texturklassen mittels einer Konfusionsmatrix, die aus der Überlagerung der beiden globalen Texturkarten erstellt wurde, konnte kein eindeutiges Ergebnis liefern. Daher wird in diesem Kapitel zunächst nur der Vergleich der drei neuen Datensätze (FAOn, Soilgrid und HWSD) durchgeführt. Die Differenzbildung zwischen dem bisherigen FAO-Datensatz und den neuen Bodendatensätzen wird erst durch die abgeleiteten Parameter möglich, was im folgenden Kapitel 6.4 analysiert wird. Es werden jeweils nur die Eingangswerte der ersten Schicht verglichen, da sich die Werte der anderen Schichten in ihrer räumlichen Verteilung und im Unterschied zwischen den Datensätzen ähnlich verhalten.

Zunächst fällt bei der Betrachtung von Abbildung 49 auf, dass - wie bereits beim weltweiten Vergleich von FAOn und Soilgrid festgestellt wurde (vgl. Abb. 26) - FAOn im nördlichen Afrika deutlich weniger Sandanteil ausweist als Soilgrid. In Grönland hingegen werden beim FAOn wesentlich höhere Sandanteile angenommen als bei Soilgrid und HWSD. Der Sandanteil beim HWSD unterliegt im Gegensatz zu FAOn und Soilgrid größeren Schwankungen auf kleinen Gebieten. Trotz der Originalauflösung von 1 km im Vergleich zu Soilgrid (ebenfalls 1 km) bzw. FAOn (10 km) erscheint die auf ca. 50 km Auflösung interpolierte Karte deutlich pixeliger. Des Weiteren ist im HWSD anders als bei den anderen beiden Datensätzen eine klare, unnatürliche Kante zwischen Schweden und Norwegen zu erkennen, die auch in den Originaldaten vorkommt und höchstwahrscheinlich ein Artefakt eines Datenbankzusammenschlusses ist. In den genannten Gebieten fallen die Unterschiede des Sandanteils mit bis zu  $\pm 60\%$  Differenz zwischen den Datensätzen sehr groß aus. Weitere zusammenhängende Flächen, die bei allen drei Datensätzen unterschiedliche Sandanteile haben, sind der Norden Deutschlands oder beispielsweise die Bereiche Südukraine, Südwestrussland, Westkasachstan. Im erstgenannten Gebiet werden sowohl beim Soilgrid als auch beim HWSD um ca. 30% höhere Sandanteile erreicht. Diese bilden beim Soilgrid ein gut zusammenhängendes Gebiet, während es beim HWSD vermehrt durch Pixel mit niedrigeren Anteilen unterbrochen ist. Das zweite Gebiet sticht beim FAOn Datensatz mit höheren Sandanteilen heraus, während es sich bei Soilgrid und HWSD nicht deutlich von den umliegenden Sandanteilen abhebt. Dementsprechend ergeben sich negative Differenzwerte von bis zu 35%. Insgesamt liegt Soilgrid beim Sandanteil tendenziell etwas höher als FAOn und HWSD.

Diese Effekte sind noch klarer in Abbildung 50 zu erkennen. In dieser Abbildung sind die Sandanteile für das GER-11-Untersuchungsgebiet dargestellt. Besonders bei FAOn, aber auch bei Soilgrid sind flächendeckende und zusammenhängende Muster zu erkennen, die auf die Kartengrundlage der jeweiligen Datensätze hinweisen, welche als Basis für die räumliche Interpolation der Eigenschaften verwendet wird. Diese Muster finden sich in allen Bodenkarten, jedoch unterschiedlich stark ausgeprägt je nach Parameter und verwendeter Auflösung.



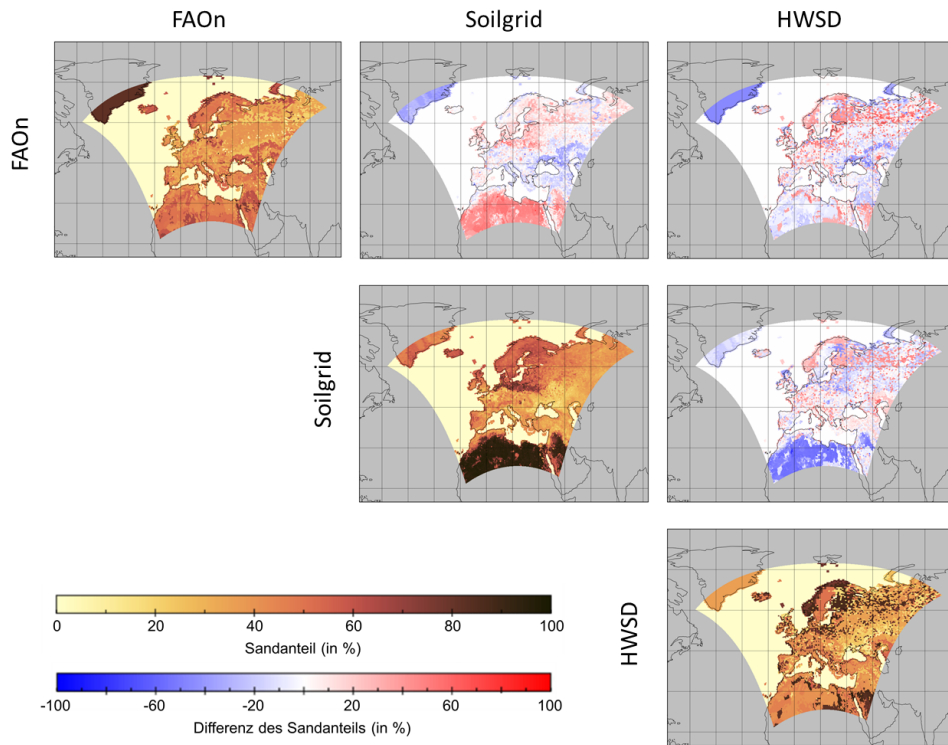


Abb. 49: Sandanteil der neuen Bodendatensätze im Vergleich für EUR-44

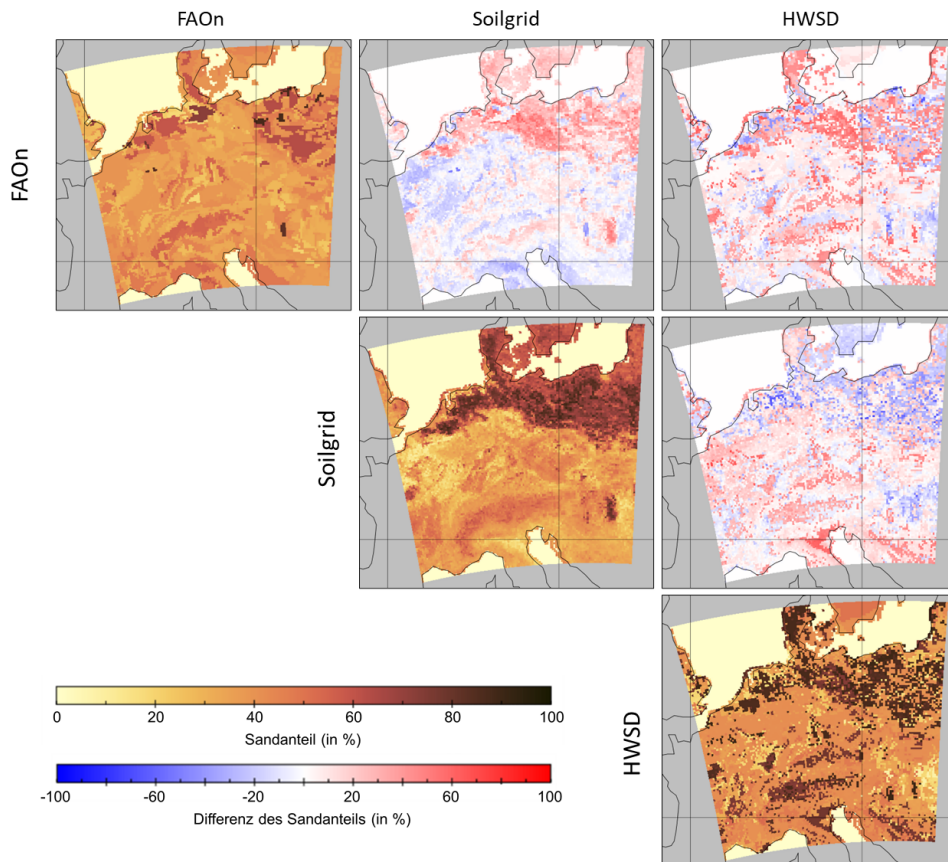


Abb. 50: Sandanteil der neuen Bodendatensätze im Vergleich für GER-11

Vergleicht man den Sandanteil der drei Datensätze direkt miteinander, fällt beispielsweise die Fläche der Po-Ebene im Norden Italiens auf, bei der beim Soilgrid mit ca. 15% Sandanteil relativ kleine, beim FAOn (ca. 50%) mittlere und beim HWSD mit ca. 75% (im östlichen Teil) hohe Sandanteile ausgewiesen werden. Ein weiteres auffälliges Gebiet befindet sich im südöstlichen Bereich des Ausschnitts (Ungarn). In allen drei Karten sind Pixel mit einem Anteil von über 75% Sand gekennzeichnet und diese treten auch deutlich von den geringeren Anteilen der umliegenden Pixel hervor. Allerdings ist anhand der Differenzabbildungen zu erkennen, dass diese kleine Region beim FAOn weiter westlich liegt bzw. dass beim Soilgrid und HWSD diese entsprechende Region nicht so hohe Anteile ausweist, aber eine zweite weiter östliche liegende Fläche die hohen Sandanteile erreicht. Es gibt aber auch Regionen im Untersuchungsgebiet, die in allen drei Datensätzen ähnliche Sandanteilswerte erreichen.

Dem Sandanteil genau entgegengesetzt, ergeben sich bei den Tonanteilen der drei Datensätze (vgl. Abb. 51) bei EUR-44 viele Regionen mit erhöhten Tonanteilen bei FAOn und HWSD, während Soilgrid keine Werte über 40% besitzt. Auch im afrikanischen Bereich des Untersuchungsgebiets sind die Anteile von Soilgrid fast flächendeckend um 20-30% kleiner als bei HWSD und FAOn. Die beim Sandanteil des HWSD stark hervorgetretene Grenze zwischen Norwegen und Schweden ist beim Tonanteil kaum sichtbar. Die stark erhöhten Tonanteilswerte des FAOn am nordöstlichen Rand der Karte treten weder beim Soilgrid noch beim HWSD auf, wobei beim HWSD ebenfalls in diesem Bereich höhere Werte erreicht werden, die aber nicht mit den Pixeln des FAOn identisch sind. Die Beträge der Differenzen zwischen den drei Datensätzen sind für den Tonanteil insgesamt mit  $\pm 20\%$  etwas kleiner als beim Sandanteil.

Bei der Betrachtung des GER-11 Ausschnitts (vgl. Abb. 52) zeigt sich wie beim EUR-44, dass sich der Tonanteil im Vergleich zum Sandanteil fast überall konträr verhält. Hohe Sandanteile entsprechen niedrigen Tonanteilen, woraus sich schließen lässt, dass die dritte Komponente (Schluffanteil) im Untersuchungsgebiet relativ gleichverteilt ist. Es fällt auf, dass nur beim HWSD vermehrt Pixel mit über 50% Tonanteil auftreten. In den gleichen Regionen weisen die anderen beiden Datensätze (z. B. nordöstliches Frankreich oder im Grenzgebiet zwischen Ungarn und Rumänien) nur Werte von 25-35% auf.

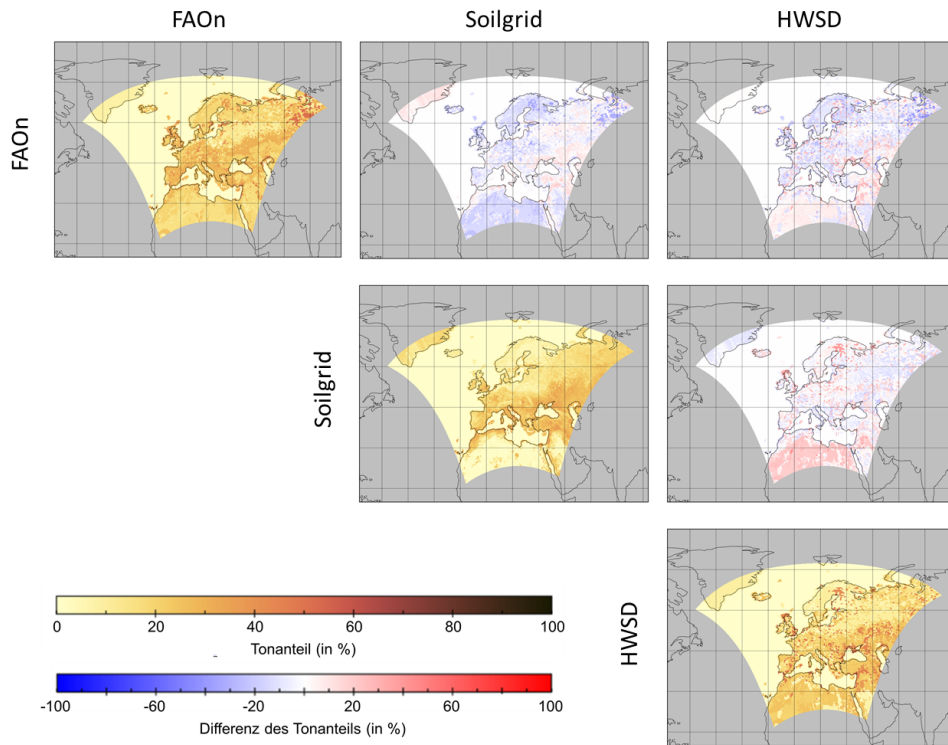


Abb. 51: Tonanteil der neuen Bodendatensätze im Vergleich für EUR-44

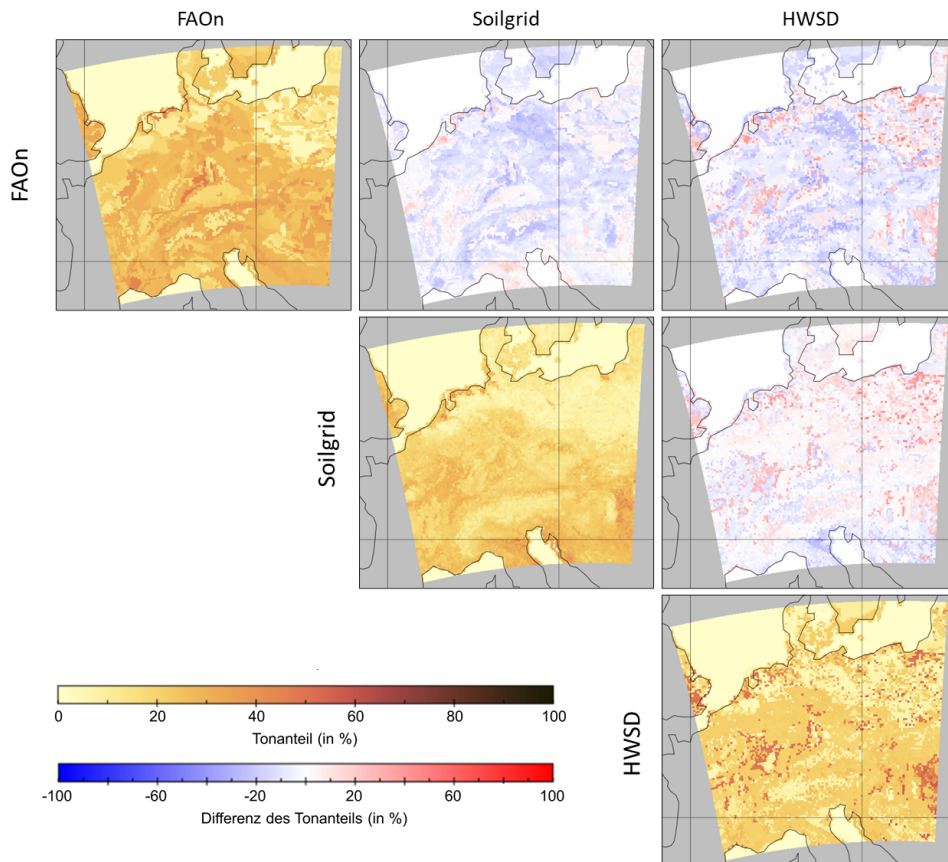


Abb. 52: Tonanteil der neuen Bodendatensätze im Vergleich für GER-11

Beim organischen Anteil der obersten Bodenschicht (FORG1), welcher in Abbildung 53 für das Untersuchungsgebiet EUR-44 dargestellt ist, erkennt man bei FAOn, dass nur an der Nordküste von Litauen und Lettland sehr hohe Werte von über 0.8 erreicht werden. Im nordöstlichen Bereich treten beim FAOn im gleichen Bereich, wo zuvor die erhöhten Tonanteile erkennbar waren, organische Werte von ca. 0.6 auf. Ähnlich dem Tonanteil werden beim HWSD ebenfalls in diesem Gebiet, aber nicht an den exakt gleichen Pixeln, höhere organische Anteile ausgewiesen. Beim SoilGrids-Datensatz hingegen werden viel mehr Pixel mit organischen Anteilen von über 0.3 ausgewiesen, sodass eine geschlossene Fläche mit sehr hohen Werten entsteht.

Im GER-11-Vergleich des organischen Anteils (Abb. 54) fallen die Unterschiede zwischen den Karten der drei Datenquellen noch stärker auf. Der HWSD kennzeichnet nur vereinzelt hohe organische Anteile an den Küsten im nördlichen Bereich des Untersuchungsgebietes (Anteile größer 0.7). Im restlichen Gebiet schwankt der Anteil zwischen 0–0.15. Anders als FAOn und HWSD wird bei Soilgrid ein stark erhöhter Anteil an organischem Material vor allen in den Alpen, aber auch in anderen Mittelgebirgen dargestellt. Insgesamt liegt der organische Anteil im Soilgrid (außer in kleinflächigen Ausnahmen) viel höher als bei den anderen beiden Datensätzen. Bei Soilgrid fällt zudem auf, dass für Tschechien, dem nordöstlichen Teil Österreichs sowie Frankreich sehr niedrige Werte ausgegeben werden. Da auch beim FAOn in diesen Bereichen im Vergleich zu umliegenden Regionen niedrigere organische Anteile vorliegen, ist davon auszugehen, dass es sich nicht um Datenfehler aufgrund von unvollständig angepassten nationalen Datenbanken handelt, wie es die realitätsnahe Abbildung der Ländergrenzen vermuten ließe.

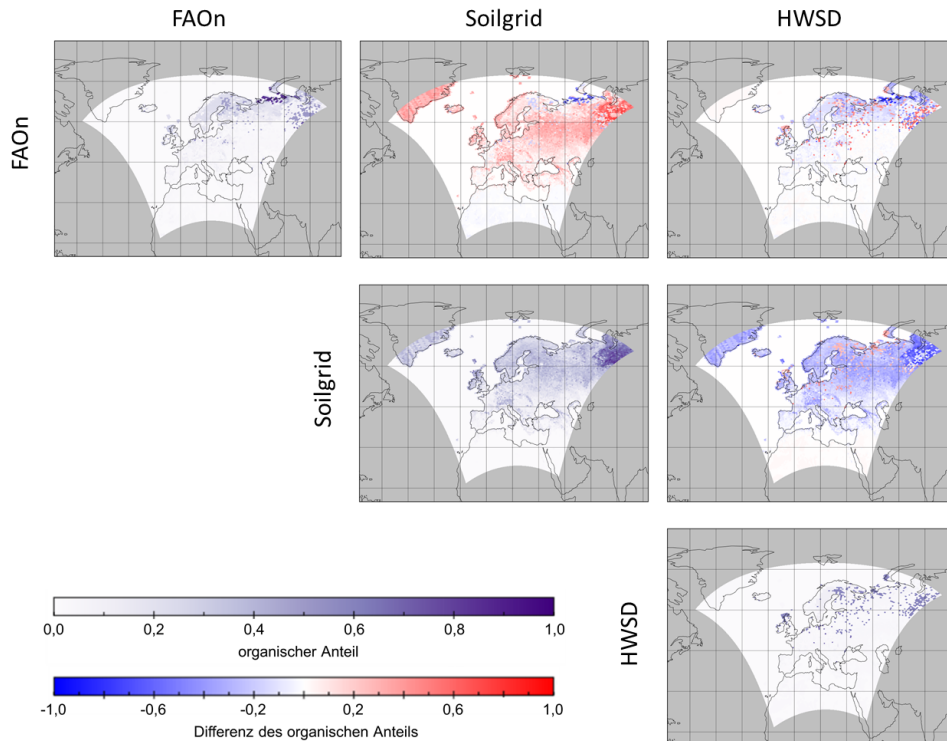


Abb. 53: Organischer Anteil der neuen Bodendatensätze im Vergleich für EUR-44

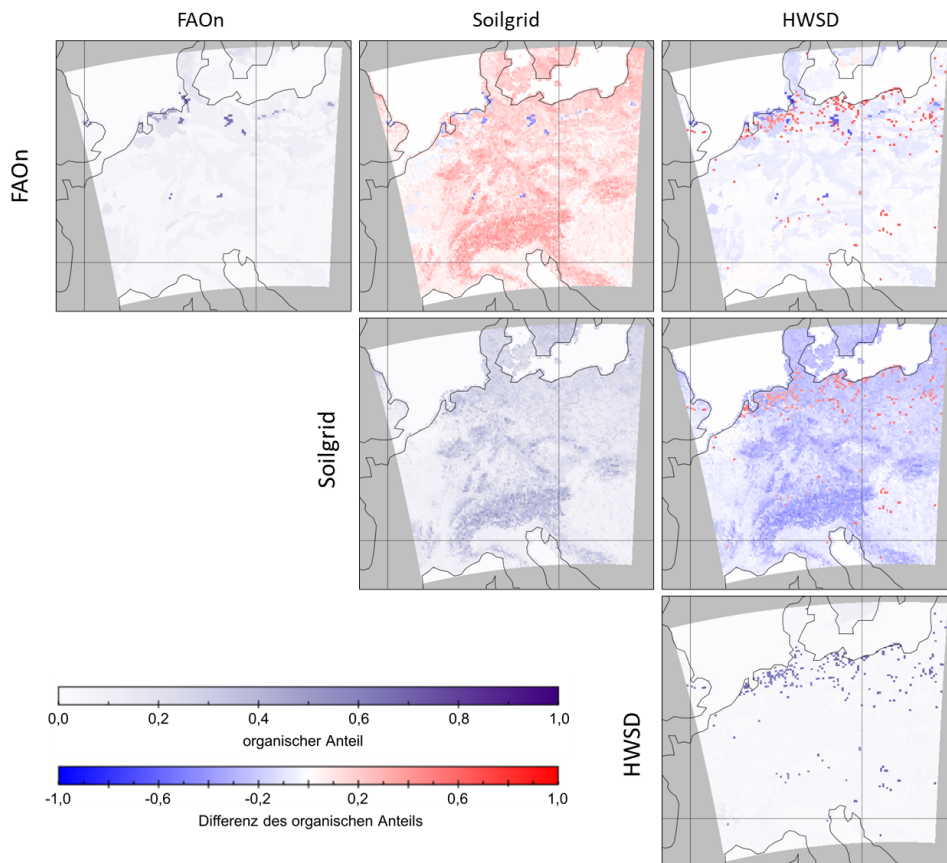


Abb. 54: Organischer Anteil der neuen Bodendatensätze im Vergleich für GER-11

Zusammenfassend lässt sich beim Vergleich aller drei Variablen (Sand, Ton und organischer Anteil) und aller drei neuen Bodendatensätze für beide Untersuchungsgebiete festhalten, dass zwar nur wenige komplett gegensätzliche Aussagen in den Daten stecken und die grobe Verteilung der Werte bei allen übereinstimmt, die absoluten Zahlen jedoch stark variieren können. Des Weiteren sticht vor allem der HWSD-Datensatz mit teilweise für die hohe Datensatzauflösung untypischen „pixeligen Flecken“ hervor, die vermutlich auf die schlechte Datenlage in jenen Regionen zurückzuführen ist, sodass hier nur gröber aufgelöste Originaldaten zur Bildung der HWSD-Karte zur Verfügung stehen (vgl. Kapitel 4.3.2). Besonders groß ist der Unterschied zwischen den Daten beim organischen Anteil. Dennoch stellen hier alle drei Datensätze sowohl wegen der vielfach besseren Auflösung als auch wegen der numerischen Abstufung von organischem Material eine Verbesserung im Vergleich zur bisherigen Ausweisung der Klasse Torf dar.

## 6.4 Auswirkungen auf die abgeleiteten Parameter

Bei den nachfolgend dargestellten abgeleiteten Parametern handelt es sich um die Bodenparameter, die ausgehend von dem jeweiligen Bodendatensatz und der gewählten Parametrisierung für die Bodentemperatur- und die Bodenhydrologieberechnung in REMO benutzt werden. Diese Untersuchung der abgeleiteten Parameter statt der direkten Eingangsdaten ist notwendig, weil sich zum einen die Anzahl der Texturklassen ändert, was einen direkten Vergleich erschwert, und weil zum anderen die Werte von mehreren sich überlagernden Eingangsfaktoren abhängen (vgl. Kapitel 5.1).

Tabelle 14 zeigt jeweils die Änderungen der drei hydrologischen Bodenvariablen (PORVOL, FCAP und PWP) sowie der drei thermischen Bodenvariablen (TLAMBDA, DLAMBDA und RGCGN) für die einzelnen Modellläufe des Experimentplans 3 (vgl. Abb. 46). Zusätzlich sind die resultierenden thermischen Variablen ZTLAMBDA und ZCGN in der Tabelle dargestellt, weil sich diese durch eine Veränderung der Berechnungsformel direkt oder durch die Veränderung der in die Formel eingehenden Bodenparameter indirekt ebenfalls ändern.

Die Änderungen von Modelllauf id03, id04 und id05 hängen mit dem Austausch des Datensatzes zusammen und sind hauptsächlich räumlicher Art. In den Modellläufen id06 bis id08 werden die thermischen Eingangsparameter direkt (id06) oder indirekt durch das Berücksichtigen des organischen Anteils (id07 und id08) modifiziert. Die Veränderungen der thermischen Variablen bei Version id11 werden durch die Verwendung des 5Layer-Schemas ausgelöst, weil hier statt eines mittleren Wassergehalts über alle Bodenschichten der jeweilige Bodenwassergehalt jeder einzelnen Schicht zur Berechnung verwendet wird.

Diese unterschiedlichen Versionen der sechs Eingangsparameter werden nachfolgend mit den Bodenvalidierungsdaten (vgl. Kapitel 4.5.3) für die oberste Bodenschicht verglichen. Die hydrologischen Parameter werden mit dem HiHydroSoil-Datensatz validiert und die thermischen Variablen sind in den Dai-Daten enthalten. Für PORVOL wird auf Daten beider Datensätze zurückgegriffen, da das Porenvolumen auch für die thermischen Variablen eine große Rolle spielt und, wie in der Tabelle zu erkennen ist, alle thermischen Bodenparameter beeinflusst.

Für eine bessere Übersicht und in Hinblick auf die Modellläufe der nachfolgenden Analysen werden nur jeweils drei Modellläufe dargestellt:

- Die Karten der bisherigen Version (id01).
- Die Werte der Version mit dem neuen Bodendatensatz Soilgrid und den beiden neuen Schaltern LSOIL und L5LAYER\_TOTAL, welche Änderungen in allen Parametern hervorruft (id05).
- Die Karten der Version, bei der alle thermischen (außer DLAMBDA) und hydrologischen Eingangsparameter neu parametrisiert wurden (id12). Für DLAMBDA wird dagegen Lauf id16 statt id12 verwendet, da bei Lauf id12 keine Änderung dieser Variable stattfindet, aber bei Lauf id16 eine neue Berechnungsformel angewendet wird.

Tab. 14: Übersicht der Bodenparameteränderungen

Eine angezeigte Änderung bezieht sich jeweils zur vorherigen ID (genauere Erklärung der Modellversionen vgl. Abb. 46):

x= Änderung der Eingangsdaten oder Berechnungsmethode

x\*= Änderung wegen L5LAY=TRUE

(x)= indirekte Änderung, weil sich ein für die Berechnung benutzter Parameter ändert

ID	PORVOL	FCAP	PWP	TLAMBDA	DLAMBDA	ZTLAMBDA (abhängig von TLAMBDA und DLAMBDA)	RGCGN	ZCGN (abhängig von RGCGN)	Änderungen
id01									
id02									
id03	x	x	(x)	x	x	(x)	x	(x)	Soilgrid statt FAO
id04	x	x	(x)	x	x	(x)	x	(x)	Sand/Ton- statt Texturwert Wert für jede Bodenschicht
id05	x	x	(x)	x	x	(x)	x	(x)	statt einem Wert für den Gesamtboden
id06				x		(x)	x	(x)	de Vries (1963) und Farouki (1981b)
id07				x		(x)	x	(x)	statt Majewski (1991)
id08				x		(x)	x	(x)	aufbereiteter organischer Anteil in allen Schichten
id09	x			(x)		(x)	(x)	(x)	organischer Anteil in Bodenschicht 4 und 5 = 0
id10		x	(x)						PORVOL PTF nach Lawrence & Slater (2008)
id11			x			x*		x*	FCAP PFT nach Adhikary et al. (2008)
id12	x	x	x	(x)		(x)	(x)	(x)	neue Parametrisierung für PWP alle hydrol. Parameter nach Saxton & Rawls (2006)
id13									
id14									
id15								x	ZCGN-Berechnung mit Luft und RGCGN nach de Vries (1963)
id16					x	x			neue Formel für DLAMBDA und TLAMBDA nach Farouki (1981b) bzw. Lawrence & Slater (2008)
id17									
id18									

Exemplarisch für die thermischen Variablen ist in Abbildung 55 die Amplitude der Wärmeleitfähigkeit (DLAMBDA) für die drei ausgewählten Modellläufe und die Validierungsdaten sowie die jeweilige Differenz dargestellt. Weil in der Abbildung das GER-11-Gebiet dargestellt ist, fallen besonders die aufgrund der groben Ausgangsaufösung des bisherigen FAO-Datensatzes entstehenden Kacheln der jeweiligen Texturklassen bei Abbildung 55B auf. Diese Kacheln sind bei allen Variablen zu finden. Die Übergänge zwischen den einzelnen Klassen und die Differenz zu den Validierungsdaten fallen je nach Parameter unterschiedlich aus. Bei DLAMBDA

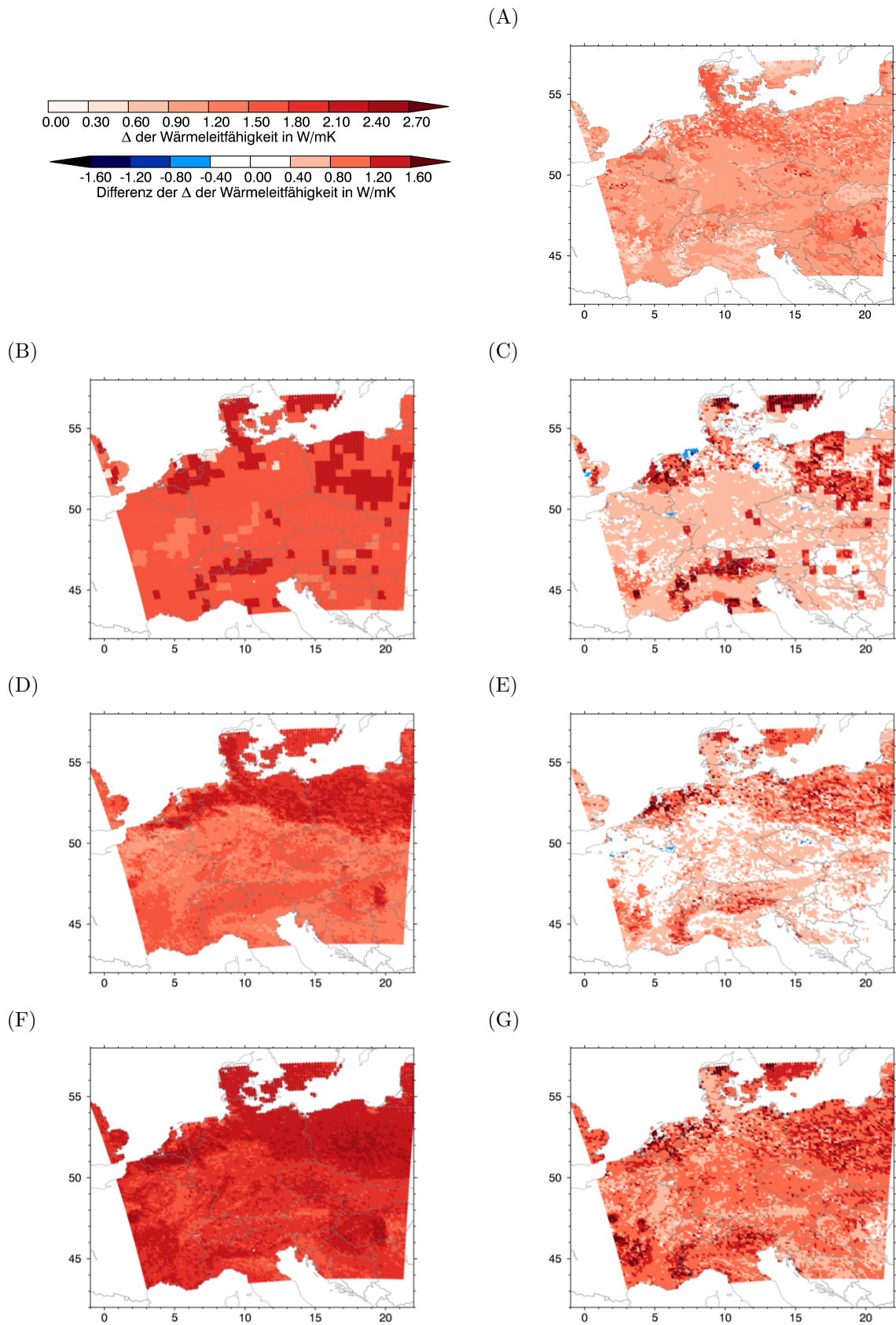


Abb. 55: Räumlicher Vergleich der Amplitude der Wärmeleitfähigkeit (DLAMBDA) von Validationsdaten und unterschiedlichen Modellläufen (Erklärung vgl. Abb. 46) in GER-11  
 (A) Validationsdaten (Dai-Daten)  
 (B), (D), (F) jeweilige Amplitude der Wärmeleitfähigkeit von id01, id05 und id16  
 (C), (E), (G) jeweilige Differenz von (B), (D) und (F) zu (A)



ist dieser Unterschied besonders stark ausgeprägt, vor allem für die Bereiche in den südlichen Alpen, Niederlande, Norddeutschland sowie in größeren Teilen Polens, wo jeweils die Texturklasse „lehmiger Sand“ ausgewiesen wird. Die mit Abstand niedrigsten Werte werden für die Flächen der Texturklasse „Torf“ erreicht, die großflächig nur in England, nordwestlich von Berlin und in Nordostdeutschland auftritt. Insgesamt wird DLAMBDA bei id01 generell überschätzt. Nur in Teilen der Torfflächen treten negative Differenzen zum Validationsdatensatz auf. Da mit Modelllauf id05 DLAMBDA fasst überall abnimmt, minimiert sich die Differenz zu den Validationsdaten. Das sowohl für Lauf id05 als auch für Lauf id16 gezeigte Muster der Absolutwerte stimmt größtenteils mit dem in Abbildung 50 dargestellten Sandanteil von Soilgrid überein, da der Sandanteil der dominante Faktor in der Berechnungsformel (vgl. Tab. 12) ist. Wegen des Anstiegs der Amplitude der Wärmeleitfähigkeit bei Modellversion id16 wird der Unterschied zwischen Validations- und Modelldaten hier größer. Die Werte werden im Modell sehr deutlich überschätzt.

Für RGCGN und TLAMBDA (nicht dargestellt) nehmen die absoluten Werte hingegen von Version id01 über id05 zu id12 ab, sodass die Differenz zum Validierungsdatensatz kleiner wird und höhere Differenzen bei Lauf id12 nur in einem kleinen Gebiet in Zentralfrankreich am südwestlichen Randbereich und vereinzelt im nördlichen Bereich des Untersuchungsgebietes vorkommen. Hierzu ist anzumerken, dass die Dai-Daten in diesen Bereichen teilweise Fehlwerte ausweisen.

Bei den hydrologischen Variablen PORVOL und FCAP, die in Abbildung 56 und 57 dargestellt sind, fallen sowohl bei den absoluten Werten als auch bei der Differenz zu den Validierungsdaten wieder die bereits angesprochenen Torfflächen auf. Beim Vergleich des Porenvolumens (Abb. 56), bei dem zwei Validierungskarten zur Verfügung stehen, zeigt sich außerdem, dass die Dai-Daten (vgl. Abb. 56A) im zentralfranzösischen Gebiet, in dem auch die hohen DLAMBDA-Differenzen auftreten, und vereinzelt in Norddeutschland und Polen sehr hohe Porenvolumenwerte von über  $0.8 \frac{\text{m}^3}{\text{m}^3}$  erreicht werden. Die gleichen Flächen sind bei den auf die hydrologischen Parameter spezialisierten Datensatz HiHydroSoil (vgl. Abb. 56B) jedoch nicht mit höheren Werten ausgewiesen, obwohl beide Datensätze größtenteils auf Soilgrid basieren. Außerdem fällt auf, dass bei HiHydroSoil die Alpen höhere Porenvolumenwerte erreichen als die ansonsten relativ homogene Umgebung. Bei den Dai-Daten hingegen unterscheidet sich diese Region nicht in diesem Maß von der Umgebung. Aufgrund der Differenzkarten des nicht dargestellten Vergleichs auf dem EUR-44-Gebiet ist zu schließen, dass in den Dai-Daten der organische Anteil in die Berechnung des Porenvolumens einfließt, während beim HiHydroSoil dieser nur indirekt über die Trockenroh-dichte (vgl. Formel 4.16) berücksichtigt wird. Wegen des großen Wertebereichs von PORVOL und FCAP und der deshalb gewählten Klassenbreite ergeben sich insgesamt für PORVOL kaum regionale Unterschiede in den Karten. Bei Lauf id05 resultiert die gewählte Einteilung sogar darin, dass im gesamten Untersuchungsgebiet nur eine Klasse ausgewiesen wird. Weil bei Version id12 die Porenvolumenwerte insgesamt reduziert werden, wechselt ungefähr die Hälfte der Fläche des Untersuchungsgebiets in eine Klasse mit geringerem Wert, wodurch die jeweiligen Differenzen von einer leichten Überschätzung der Validationswerte in die neutralen Differenzklassen wechseln. Die Porenvolumenwerte in den Alpen, die bereits bei Modelllauf id01 und auch bei id05 im Vergleich zu den Validationswerten unterschätzt werden, werden jedoch durch die Verminderung

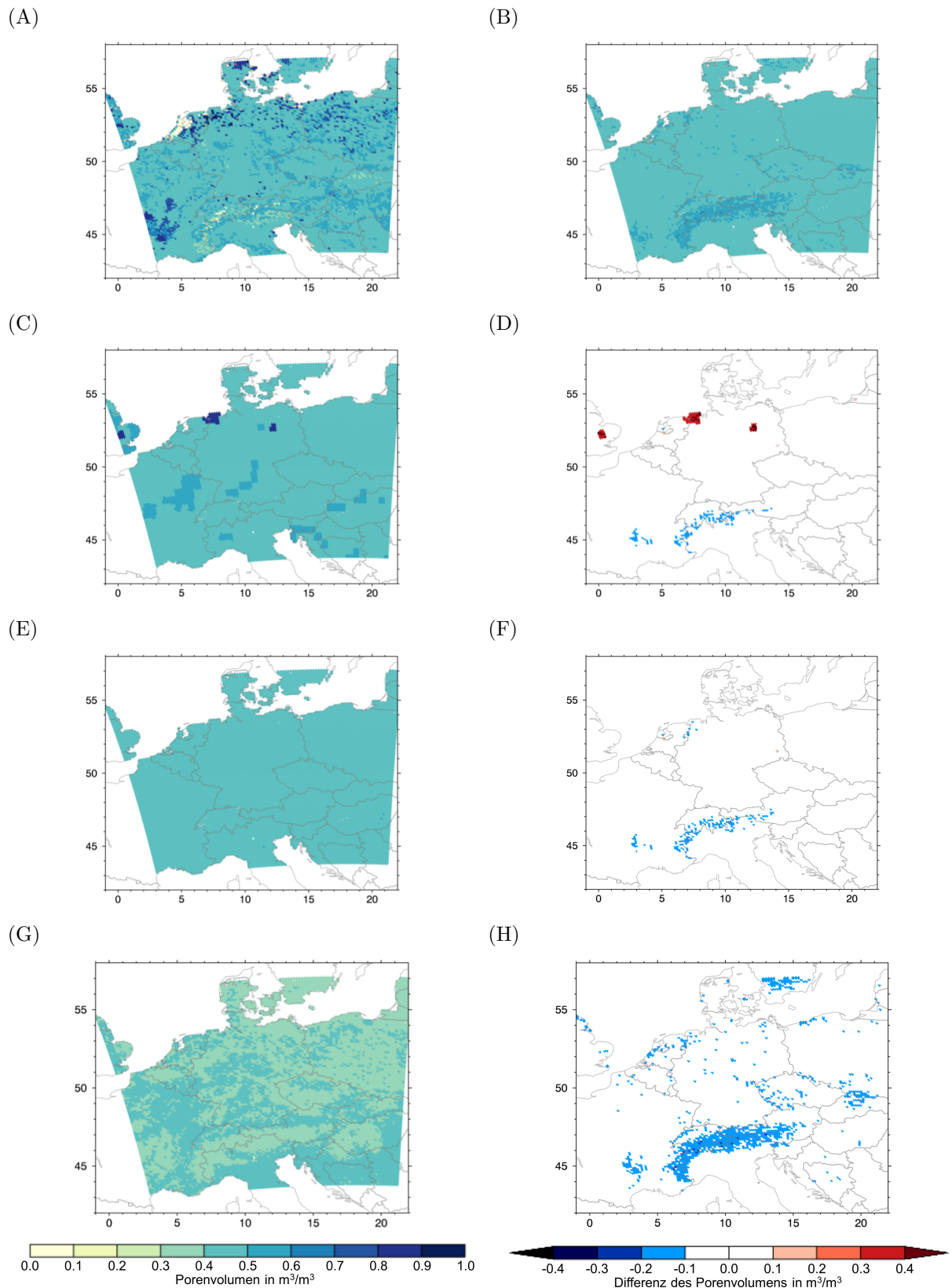


Abb. 56: Räumlicher Vergleich des Porenvolumens (PORVOL) von Validationsdaten und unterschiedlichen Modellläufen (Erklärung vgl. Abb. 46) in GER-11  
 (A) Validationsdaten (Dai-Daten); (B) Validationsdaten (HiHydroSoil)  
 (C), (E), (G) jeweiliges Porenvolumen von id01, id05 und id12  
 (D), (F), (H) jeweilige Differenz von (C), (E) und (G) zu (B)

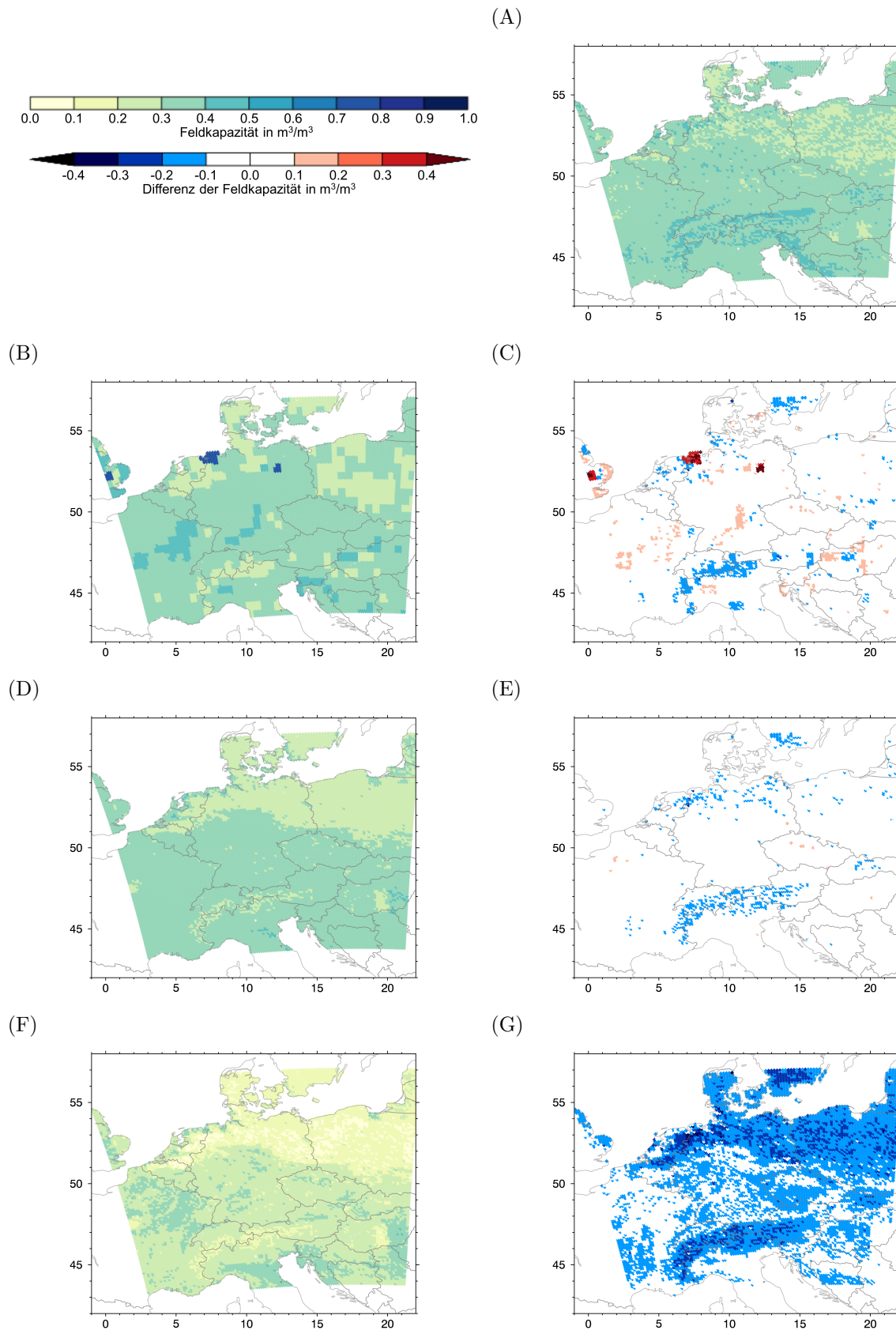


Abb. 57: Räumlicher Vergleich der Feldkapazität (FCAP) von Validationsdaten und unterschiedlichen Modellläufen (Erklärung vgl. Abb. 46) in GER-11  
 (A) Validationsdaten (HiHydroSoil)  
 (B), (D), (F) jeweilige Feldkapazität von id01, id05 und id12  
 (C), (E), (G) jeweilige Differenz von (B), (D) und (F) zu (A)

der tatsächlichen Porenvolumenwerte noch stärker unterschätzt. Auch die Feldkapazitätswerte (vgl. Abb. 57) nehmen durch die Parameteränderungen von Modellversion id01 bis zu id12 ab. Jedoch sind bei Version id05 insgesamt hauptsächlich zwei Klassen in der Karte zu finden. Bei Lauf id12 lassen sich mehr Strukturen anhand der Absolutwerte erkennen. Der Unterschied von id05 zu id12 ist einerseits durch die unterschiedlich großen negativen Koeffizienten des Sandanteils in der jeweils verwendeten Formel (vgl. Kapitel 5.1.1) und andererseits durch das Berücksichtigen des organischen Anteils in der für Lauf id12 verwendeten Berechnungsformel (vgl. Formel 5.5) zu erklären. Weil ein höherer Sandanteil zu einer geringeren Feldkapazität führt und wie bereits bei DLAMBDA der Sandanteil der dominierende Faktor bei der Berechnung des Bodenparameters ist, nimmt die Feldkapazität ab und das räumliche Muster von Lauf id12 ähnelt dem des Sandanteils von Soilgrid (vgl. Abb. 50). Die starke Abnahme aller FCAP-Werte von Modelllauf id05 auf id12, die besonders deutlich in Norddeutschland und im Alpenraum zu erkennen ist, führt jedoch beim Vergleich mit den Beobachtungsdaten dazu, dass die gute Übereinstimmung von Version id05 mit den Validationswerten durch eine starke Unterschätzung der Validationswerte von Version id12 ersetzt wird.

Da die Einfärbung der Werte bei beiden hydrologischen Vergleichsabbildungen identisch gewählt wurde, ist daraus abzuleiten, dass zum einen der Porenvolumenwert im gesamten Untersuchungsgebiet nicht so stark variiert wie die Feldkapazität und dass zum anderen die Feldkapazität durch die vorgenommenen Parametrisierungsänderungen v. a. für Lauf id12 wesentlich stärker betroffen ist, obwohl FCAP bei der Berechnung von PORVOL verwendet wird (vgl. Formel 5.6). Wie die Feldkapazität nimmt der PWP-Wert (nicht dargestellt) in einem ähnlichen Muster ab, was aber zu einer besseren Übereinstimmung mit den Validationsdaten führt.

Zusammenfassend lässt sich festhalten, dass zwar die Feldkapazität durch Version id12 im Vergleich zu den Beobachtungsdaten stärker unterschätzt wird als durch Modellversion id05 und durch andere nicht gezeigte Parametrisierungen von anderen Modellversionen, PORVOL und PWP jedoch durch diese Version kleinere oder ähnliche Differenzwerte erreichen als bei Modelllauf id01 und id05. Ferner muss beachtet werden, dass sowohl der Feldkapazitätswert als auch der permanente Welkepunkt je nach Datensatz als unterschiedliche Wassergehalte auf der WRC definiert sind. Der dargestellte Validationsdatensatz HiHydroSoil verwendet beispielsweise für die Feldkapazität pF2 (vgl. Kapitel 4.5.3), was einem Wassergehalt von ca. 100 hPa entspricht. Die ab Lauf id12 verwendeten Formeln von Saxton & Rawls (2006) geben die Feldkapazität jedoch mit 33 kPa an. Aufgrund von mangelnden Messwerten oder anderen Beobachtungsdaten lässt sich daher keine eindeutige Beurteilung der neuen Parameterwerte durchführen und die größeren negativen Differenzen von id12 sind daher nicht sofort mit einer Verschlechterung der Eingangsdaten ins Modell gleich zu setzen.

Ähnliche Beobachtungen können auch beim EUR-44-Gebiet gemacht werden (nicht dargestellt). Die Tatsache, dass die Texturklassen „Sand“ und „lehmiger Sand“ so hohe DLAMBDA-Werte im Vergleich zu den restlichen Klassen in Version id01 ausweisen und vor allem, dass die Feldkapazität mit der Parametrisierung von Version id12 sehr stark reduziert wird, intensiviert sich für das größere Untersuchungsgebiet, weil ein differenziertes Bodenspektrum durch das größere Gebiet

abgebildet wird. Besonders große Unterschiede zu den jeweiligen Validierungsdaten sind vor allem in Skandinavien, im nordöstlichen und auch im afrikanischen Bereich des Untersuchungsgebiets zu erkennen. Der bei GER-11 angesprochene Kacheleffekt tritt bei diesem Testgebiet in den Absolutwerten von Lauf id01 kaum auf, wenngleich der direkte Vergleich der Werte von Modelllauf id01 und id05 die teilweise räumlich abrupten Werteübergänge trotzdem deutlich macht.

## 6.5 Auswirkungen auf klimatische Variablen

Neben den Effekten auf die abgeleiteten Parameter (Kapitel 6.4) wirken sich die Änderungen der Eingangsdaten und der Parametrisierung sowohl räumlich als auch zeitlich auf die anderen Variablen des Klimamodells aus. In den folgenden Unterkapiteln werden die verschiedenen Modellversionen untereinander und mit den in Kapitel 4.5 vorgestellten Validationsdaten auf unterschiedlichen räumlichen (EUR-44 und GER-11) und zeitlichen Skalen verglichen. Der Relaxationsbereich des Modells (acht Modellpixel am Rand, vgl. Kapitel 3.2.1) wird in allen Analysen nicht verwendet.

### 6.5.1 Vergleich zwischen den verschiedenen Modellversionen

Aufgrund der Vielzahl an Änderungen, die in dieser Arbeit getestet wurden (vgl. Abb. 44, 45 und 46), und der begrenzten Rechen- und Speicherkapazität wurden alle Änderungen einzeln nur für das Jahr 2000 als Kaltstart analysiert. Die sich daraus ergebenden Hinweise und Erkenntnisse über die Entwicklung und den Effekt der Änderungen wurden anschließend genutzt, um eine Aggregation der Änderungen anzuwenden und damit längere Läufe (2000-2018) als Warmstart durchzuführen. Die Warmstartläufe werden in Kapitel 6.5.2.2 vorgestellt.

#### 6.5.1.1 Ergebnisse der Hypothesentests

Um den Variablenraum des Modells zu reduzieren, auf dem die beschriebenen Änderungen wirken, und zu untersuchen, ob sich verschiedene Effekte überlagern, sind verschiedene statistische Methoden durchgeführt worden, deren Ergebnisse nachfolgend kurz beschrieben werden.

#### Signifikanztestergebnisse der Modelllaufdifferenzen

Mittels der in Kapitel 6.1 vorgestellten Testmethoden werden zunächst die signifikanten Änderungen verschiedener Modellversionen untersucht. Je nachdem, ob nach einem Lillifors-Test eine Normalverteilung für die Daten vorliegt, werden t-Test oder Wilcoxon-Test zur Überprüfung verwendet. Als Datengrundlage dienen die Tages- und Monatsanomalien der jeweiligen Läufe des Jahres 2000, welches für alle Modellläufe gerechnet worden ist. Bei den Tagesanomalien wird das jeweilige Monatsmittel und bei den Monatsanomalien das jeweilige Jahresmittel abgezogen. Es

werden jeweils alle 115 Variablen, die standardmäßig in die stündlichen Ergebnisdateien (e-files) von REMO (vgl. Anhang B) geschrieben werden, paarweise für zwei Modellläufe miteinander verglichen. Trotz des kleineren Stichprobenumfangs der monatlichen Werte führen die Untersuchung dieser Daten zu annähernd identischen Ergebnissen wie jene mit den Tagesanomalien. Die weiteren Untersuchungen in diesem Zusammenhang fokussieren sich auf die täglichen Werte. Die Analyse wird sowohl auf räumlicher Basis (Anomaliebildung mit dem räumlichen Mittel) als auch auf zeitlicher Basis (Anomaliebildung mit zeitlichen Mittel) durchgeführt. Die Ergebnisse der räumlichen und zeitlichen Ansätze unterscheiden sich deutlich. Insgesamt ergeben sich auf räumlicher Basis mehr signifikante Änderungen für die Variablen als auf zeitlicher Basis.

In Tabelle 15 ist exemplarisch die Zusammenfassung aller Änderungen im GER-11-Testgebiet für die räumliche Differenzanalyse für Experimentplan 1 dargestellt. Es fällt auf, dass alle Modelllaufvergleiche hohe Werte an signifikant unterschiedlichen Variablen erreichen. Die kleinste Anzahl mit 67 von 115 Variablen haben die beiden Läufen id01\_GST und id01\_AST. Der einzige Unterschied beider Modellläufe ist der veränderte Topographiedatensatz. Die anderen beiden Modellpaare, die sich ebenfalls nur durch die Topographie unterscheiden, sind id01 und id01\_AFT sowie id01\_GSF und id01\_ASF. Bei diesen unterscheiden sich 74 bzw. 77 Variablen signifikant. Die meisten signifikant unterschiedlichen Variablen erreichen die Modellläufe GFT und AST sowie GST und ASF mit jeweils 92 Variablen. Die Analyse zeigt, dass trotz der Verwendung der jeweiligen räumlichen und zeitlichen Anomalien, was eine Beeinflussung des Testergebnisses durch auftretende konstante räumliche und zeitliche Differenzen zwischen den Modellläufen ausgleichen sollte, sehr viele von den 115 Testvariablen bereits durch den Austausch der Datensätze signifikante Änderungen erfahren.

Tab. 15: Anzahl der Variablen pro Modelllauf des Experimentplan 1 mit signifikanten Unterschieden für GER-11

Als Datengrundlage dienen tägliche Anomaliedaten, die auf räumlicher Basis erstellt wurden, indem das räumliche Mittel vom jeweiligen Pixelwert abgezogen wird. Angegeben ist jeweils die Anzahl an Variablen, die aus den 115 Testvariablen signifikante Unterschiede (für  $\alpha = 5\%$ ) zwischen zwei Modellläufen aufweisen. Die Bedeutung der Modelllaufkürzel ist Abbildung 44 zu entnehmen.

	id01_GFT	id01_GFF	id01_GHT	id01_GHF	id01_GST	id01_GSF	id01_AFT	id01_AST	id01_ASF
id01	80	81	82	85	82	88	74	89	86
id01_GFT		80	82	84	88	88	84	92	84
id01_GFF			75	85	88	89	87	89	90
id01_GHT				83	90	90	86	90	89
id01_GHF					79	77	88	88	79
id01_GST						80	90	67	92
id01_GSF							90	84	77
id01_AFT								91	84
id01_AST									81

Insgesamt ist es daher einfacher, zunächst die Variablen ohne signifikante Änderung auszusortieren als die Variablen mit signifikanter Änderung auszuwählen, um die Auswirkungen der Datensatzänderungen zu analysieren. In Anhang J ist die zur Tabelle 15 gehörende Tabelle abgebildet, bei der die paarweisen Ergebnisse der Analyse pro Variable aufgelistet sind. Aus dieser kann abgeleitet werden, dass nur die jeweiligen Variablen der Eisfraktion (u- und v-Windänderung, sensibler Oberflächenenergiefluss, Oberflächentemperatur, Rauheitslänge, Oberflächenalbedo, spezifische Feuchte, Oberflächenevaporation sowie Meereseisbedeckung und -dicke) keine signifikanten Veränderungen vorweisen, da sie im EUR-44-Gebiet kaum und im hier dargestellten Ergebnis von GER-11 gar nicht vorkommen. Auch die Oberflächentemperatur und -albedo der Landoberflächenklasse „Wasser“ werden jeweils durch die neuen Datensätze nicht geändert. Zudem gibt es keine Auswirkungen bei den konstanten Werten wie Oberflächenhintergrundalbedo, Vegetationsanteil und LAI.

Genauere paarweise Untersuchungen zeigen den komplexen Zusammenhang der einzelnen Variablen. Kleine Modifikationen an einzelnen Eingangsvariablen wirken sich auf Variablen aus, die auf den ersten Blick nicht nachvollziehbar sind. Beispielsweise hat Modelllauf id01\_AFT laut der Analyse mit 74 von 115 Variablen die geringste Anzahl an signifikant unterschiedlichen Variablen zu Lauf id01. Trotz der Tatsache, dass keine Veränderungen an den Bodenparametern von Modelllauf id01 zu id01\_AFT vorgenommen wurden, da hier nur die Topographie ausgetauscht wurde, ergeben sich laut Differenzanalyse unter anderem signifikante Änderungen im Eisanteil der unteren drei Bodenschichten und Änderungen in den Bodentemperaturen der obersten zwei Bodenschichten. Dies könnte an den ebenfalls ausgewiesenen Veränderungen im großskaligen Niederschlagssignal (APRL) und den daraus folgenden Änderungen der Drainage (DRAIN) und des Bodenwassergehalts (WS) liegen. Während bei allen anderen Modelllaufkombinationen sowohl sensible als auch latente Wärmeflüsse durch die Modifikationen beeinflusst werden, fällt bei diesen beiden Modellläufen auf, dass alle sensiblen Wärmeflussvariablen durch jene nicht betroffen sind. Ebenfalls auffällig ist, dass jeweils für die minimale und maximale Oberflächentemperatur sowie für die minimale 2-m-Temperatur keine signifikanten Änderungen, für die mittlere und maximale 2-m-Temperatur jedoch signifikante Unterschiede durch das Austauschen der Topographie vorliegen.

Auch für die anderen Experimentpläne ergeben sich in fast allen Variablen signifikante Änderungen. Die geringsten Werte zwischen zwei Modellläufen ergeben immer noch in ca. 50% der Variablen signifikante Änderungen. Dies liegt einerseits an den relativ großen Stichprobenumfängen durch die Verwendung von täglichen Daten bzw. durch die Anzahl der Gitterboxen und andererseits an den komplexen Wechselwirkungen zwischen den Variablen trotz der Verwendung von Anomalien. Zudem ist das gewählte Signifikanzniveau von 95% relativ liberal, sodass bereits kleinere Veränderungen zum Ergebnis „signifikanter Unterschied“ führen. Für eine kritischere Entscheidung könnte für die Tests ein Wert von 99% gewählt werden. Dies ist ein Nachteil dieser Art von Analyse, da durch die Verwendung eines festen Signifikanzniveaus nur ein binäres Ergebnis entsteht. Ferner lässt sich erkennen, dass es eine starke Abhängigkeit des Testergebnisses von der Art der Bildung der Anomalien für Zeit und Raum gibt, was außerdem durch die unterschiedlichen Ergebnisse für die beiden Untersuchungsgebiete deutlich wird. Es

findet außerdem keine Berücksichtigung der Stärke des Effektes statt. Aus diesen Erkenntnissen folgt, dass weitere Analysen speziell zur Minimierung des Variablenraums und zur Detektion der zusammenhängenden Variablen notwendig sind, um genauer auf die Auswirkungen der jeweiligen Daten- bzw. Parametrisierungsänderung eingehen zu können.

### Ergebnisse der ANOVA

Neben dem Test der Mittelwerte wurde auch eine Zwei-Wege-ANOVA (vgl. Kapitel 6.1.2) für verschiedene Modellläufe durchgeführt, um Effekte von beispielsweise verschiedenen Datensätzen räumlich zu detektieren und Interaktionseffekte bei den folgenden Analysen berücksichtigen zu können. Dabei wurden die Effekte zum einen anhand der Topographie- und der Bodendaten mit den Ausprägungen GTOPO und ASTER (Faktor 1) bzw. FAO und Soilgrid (Faktor 2) und zum anderen anhand der neuen Bodendaten (FAOn, Soilgrid und HWSO) als Faktor 1 und die LSOIL-Schalter (Textur bzw. Sand/Tonanteil) als Faktor 2 für alle zeitlich nicht konstanten Ergebnisvariablen (vgl. Anhang B) getestet. Die Datengrundlage hierfür liefern die täglichen Mittel- bzw. Summenwerte der verschiedenen Läufe von Experimentplan 1 (vgl. Abb. 44). Um eine Übersicht zu bekommen wurden zunächst alle Pixel des Untersuchungsgebiets zu jedem Zeitpunkt als Gruppe behandelt, woraus durch die ANOVA für jede Variable eine Aussage zu den Effekten der beiden Faktoren sowie zur Interaktion und dem Residuum gemacht werden kann. Im zweiten Schritt wurde für jeden Pixel eine ANOVA durchgeführt. Die Einträge der Gruppen bestehen damit nur noch aus den Tageswerten eines jeden Pixels. Der pixelspezifische Jahresgang wurde vor der jeweiligen Analyse durch die Subtraktion des Mittelwerts entfernt.

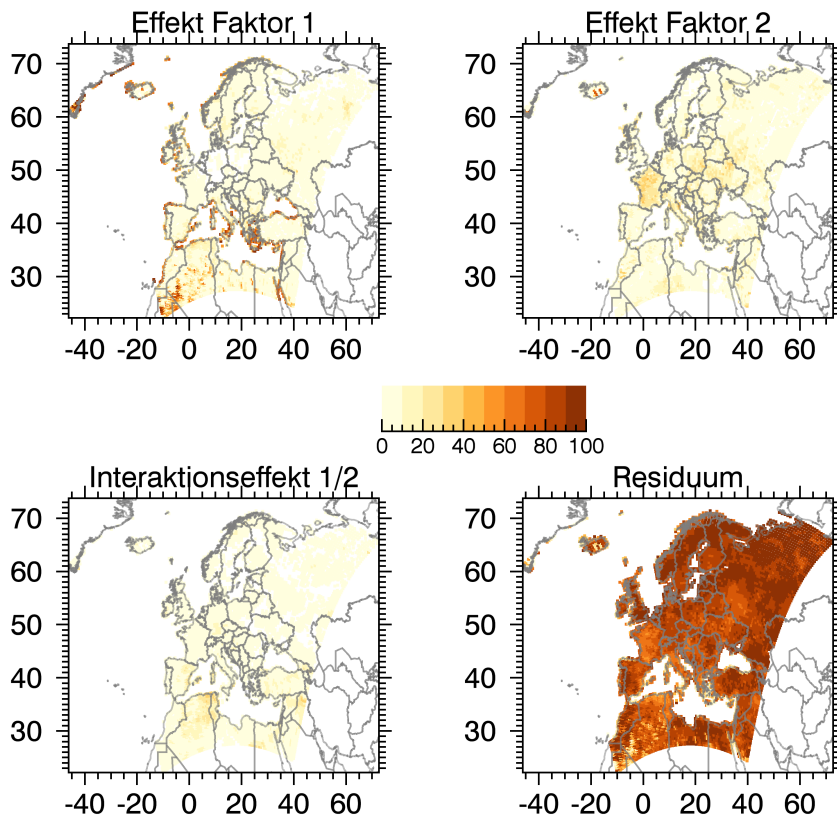
Exemplarisch werden die Karten der Ergebnisse von der 2 x 2 ANOVA mit der Topographie- als Faktor 1 und dem Bodendatensatz als Faktor 2 für EUR-44 und GER-11 erläutert. Wie zu erwarten ist, erreichen die direkt von den Datensatzänderungen betroffenen Variablen flächendeckend hohe Effektwerte, während das Residuum für diese Variablen klein ist. Für Faktor 1 sind das beispielsweise die Rauheitslänge und deren Einzelwerte für die verschiedenen Landfraktionen. Dagegen sind in kleinerem Umfang die 2 m-Temperaturen, die Oberflächen- und Bodentemperaturen für Faktor 2 die Variablen mit den höchsten Varianzwerten. Die Interaktion aller Variablen ist dagegen sehr klein oder nicht signifikant.

Ein Beispiel, bei dem in beiden Hauptfaktoren höhere Werte erreicht werden, ist die Bodentemperatur der mittleren Bodenschicht (TD5), welche in Abbildung 58 dargestellt ist. In der Abbildung werden nur signifikante ( $\alpha=5\%$ ) Pixel gezeigt, sodass die Summe der Varianzkarten für jeden Pixel nicht immer genau 100% ergibt. In der Analyse des EUR-44-Gebiets (Abb. 58A) ist zu erkennen, dass bei Faktor 1 (Topographie) vor allem an den Küsten höhere Varianzwerte erreicht werden, was mit den Regionen der höchsten negativen Topographiedifferenzen (vgl. Abb. 47) übereinstimmt. Bei GER-11 (Abb. 58B) dagegen sind auch Pixel mit höheren positiven Topographiedifferenzen (vgl. Abb. 48A) mit höheren Varianzwerten ausgewiesen. Es fällt jedoch auf, dass die Muster nicht übereinstimmen. Besonders deutlich wird das anhand der Alpenregion. Zwar erreichen die Gebirgepixel, die die Po-Ebene in Italien umschließen und die positive Differenzwerte beim Topographievergleich hatten, Determinationskoeffizienten von über 50%, die höher



gelegenen Regionen innerhalb Deutschlands, welche ein variables Vorzeichen aber ähnlich hohe Beträge bei den Differenzen der Topographiedaten haben, haben dagegen in der Varianzkarte Werte unter 30% und sind nicht von der Umgebung abzugrenzen. Bei den Varianzwerten des zweiten Faktors (Bodendatensatz, FAO und Soilgrid) fällt beim EUR-44 besonders Frankreich mit flächendeckenden Werten von 20–40% auf. Auch beim GER-11-Gebiet werden dort die höchsten Werte ausgewiesen, welche jedoch deutlich höher liegen. Die der groben Auflösung der bisherigen FAO-Bodendaten geschuldete Kachelstruktur, die bereits bei den Differenzkarten der Bodenparameter (z. B. Abb. 55B und 55C) nachgewiesen wurde, ist bei den Ergebnissen der ANOVA nur im südöstlichen Teil des GER-11-Gebiets (Österreich, Ungarn, Slowenien, Kroatien) bei Faktor 2 und konträr gelegen beim Residuum mit höheren erklärten Varianzen zu erkennen. Es fällt zudem auf, dass bei der höheren Modellauflösung insgesamt höhere Werte bei den beiden Hauptfaktoren jedoch niedrigere Interaktionseffekte und Residuen berechnet werden. Sehr ähnliche Muster finden sich auch bei den anderen Bodentemperaturvariablen.

(A)



(B)

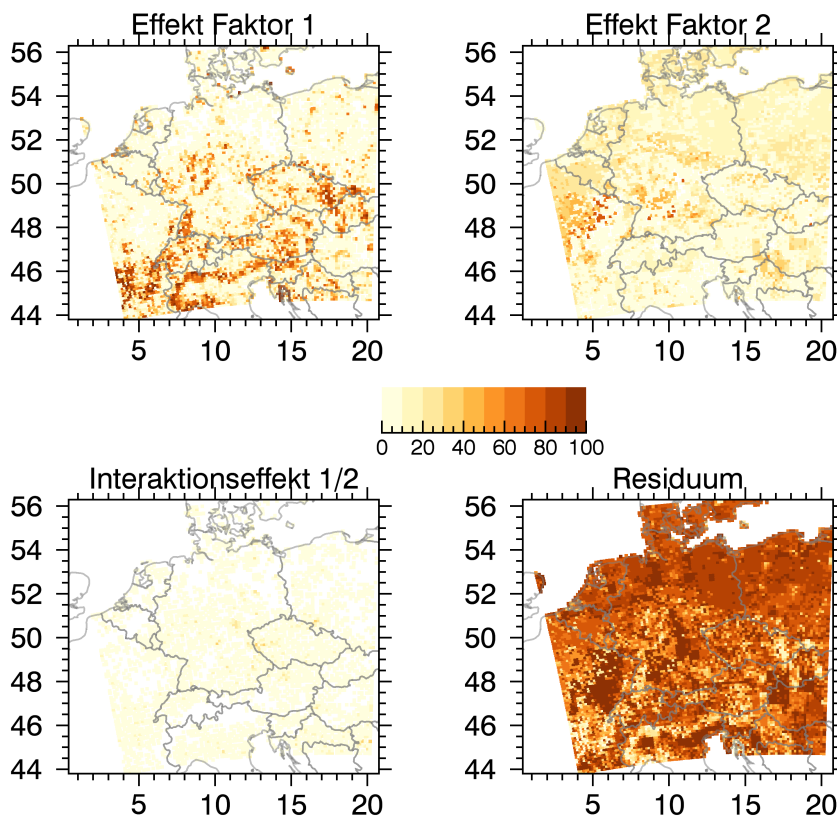


Abb. 58: Ergebnis der pixelweisen 2 x 2 ANOVA mit den Topographiedaten (GTOPO und ASTER) als Faktor 1 und den Bodendaten (FAO und Soilgrid) als Faktor 2 (dargestellt sind nur die Pixel mit signifikanten Werten ( $\alpha=5\%$ ))

(A) EUR-44

(B) GER-11

In der zweiten ANOVA, bei der die drei neuen Bodendatensätze und der LSOIL-Schalter als Faktoren zählen, sind vereinzelt beim Wassergehalt des Bodens und in den erwähnten (Boden-)Temperaturen auch Varianzen von bis zu 40% beim Interaktionseffekt zu erkennen. In allen anderen Variablen liegt nahezu die gesamte erklärte Varianz im Residuumteil, was der intraannuellen Variabilität der einzelnen Variablen entspricht.

Bei der ebenfalls getesteten 2 x 2 x 2 ANOVA, die auf der Grundlage der Läufe des Experimentplans 2 (vgl. Abb. 45) zur Analyse der Interaktionseffekte der drei Schalter (LSOIL, L5LAYER\_TOTAL und LTEMP) verwendet wurde, können ebenfalls in den Bodenparametern und den Bodentemperaturen leichte Effekte in allen drei Einzelfaktoren und der Interaktion von LSOIL und L5LAYER\_TOTAL nachgewiesen werden. In allen anderen Interaktionen liegt die erklärte Varianz nahe Null. Auch in der Kartendarstellung, die bei dieser dreifaktoriellen Varianzanalyse acht Karten umfasst, werden flächendeckend in den allermeisten Variablen beim Residuum die höchsten Werte erreicht. Zudem kann die Variabilität, die nicht im Residuum enthalten ist, auf mehr Faktoren und Interaktionen verteilt werden, sodass die Werte in den jeweiligen Karten kleiner ausfallen. Ausnahmen hiervon bilden wie bei den anderen ANOVA-Analysen nur die direkt durch die Schalter beeinflussten Variablen (Bodenparameter und -temperaturen).

Aus den Ergebnissen der ANOVA kann abgeleitet werden, dass die Interaktionseffekte der verschiedenen Datensatzänderungen in allen Variablen kaum eine Rolle spielt. Zudem liegt trotz der jeweiligen Anomaliebildung zur Elimination des Jahresgangs der überwiegende Teil der Varianz bei fast allen Variablen im Residuum. Variablen mit größeren Effekten sind bei allen verschiedenen Kombinationen die Rauigkeitslänge bei der Topographie bzw. der Wassergehalt des Bodens, der Eisanteil und die Bodentemperaturen der verschiedenen Schichten. Je nach ANOVA und Modellgebiet erreichen auch die Oberflächen- und die 2 m-Temperaturen höhere Varianzwerte. Außerdem kann aufgrund des geringen Einflusses der Interaktionen in den nachfolgenden Analysen direkt auf die jeweiligen Haupteffekte eingegangen werden.

### 6.5.1.2 Veränderungen in ausgewählten Variablen

Die vorausgegangenen Analysen zeigen, dass die Modelländerungen viele verschiedene Variablen in unterschiedlichen Stärken beeinflussen. Aufgrund der komplexen Zusammenhänge, der Interpretierbarkeit der Ergebnisse und auch wegen des Umfangs dieser Arbeit wird daher nachfolgend exemplarisch nur auf wenige Variablen eingegangen.

#### Veränderungen in den 10 m-Windvariablen

Das Verwenden der ASTER-Topographiedaten statt der bisherigen GTOPO ist eine Änderungen in Experimentplan 1 (Abb. 44). Zusätzlich wird ASTER auch im Modelllauf id18a verwendet. Es ist bekannt, dass die Topographie neben dem Einfluss auf Temperaturen auch einen starken Zusammenhang mit den Windvariablen hat (z. B. Walter et al., 2006; Pryor et al., 2012), weshalb zunächst die Auswirkungen der neuen Topographie auf die Windkomponenten analysiert wird.

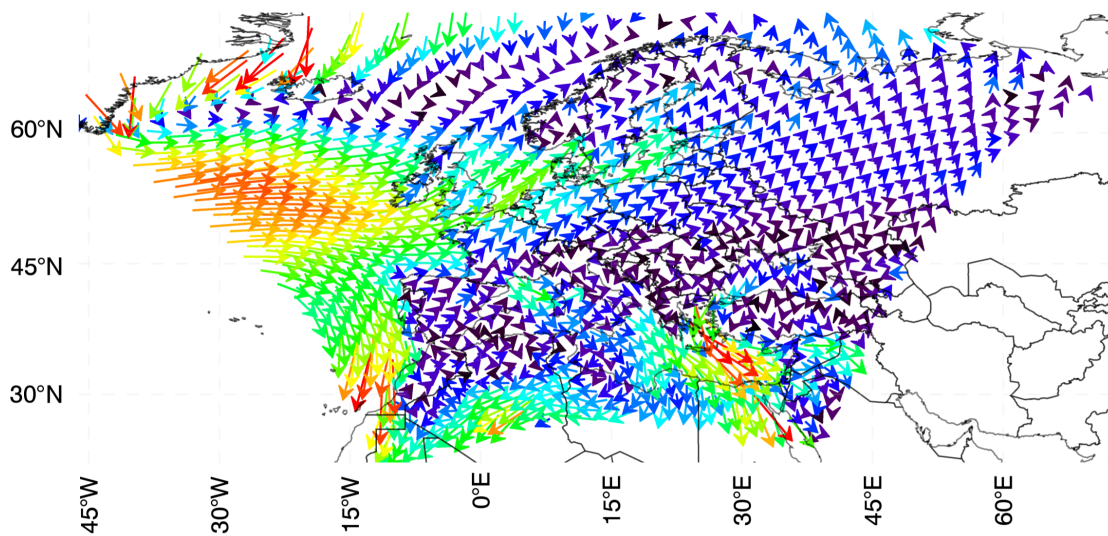
### Bisherige Modellversion im Vergleich zu ERA5

Um den Einfluss der Topographie-Datensatzänderung besser untersuchen und die Effektstärke einordnen zu können, zeigt Abbildung 59 zunächst die mittlere Windgeschwindigkeit und Windrichtung der ERA5-Daten für beide Untersuchungsgebiete. Während die Richtung der eingezeichneten Pfeile anzeigt, woher der Wind durchschnittlich kommt, kennzeichnen Länge und Farbe der Pfeile die durchschnittliche Windgeschwindigkeit. Es fällt sowohl bei EUR-44 (Abb. 59A) als auch bei GER-11 (Abb. 59B) auf, dass die Windgeschwindigkeit über dem Meer deutlich höher ist als über Land. Infolgedessen werden in Küstenbereichen ebenfalls höhere Windgeschwindigkeiten gemessen als weiter im Landesinneren. Außerdem liegen die Windgeschwindigkeiten in topographisch erhöhten Gebieten, wie z. B. den Alpen, im dargestellten zeitlichen Mittel im unteren Wertebereich. Besonders hohe Windgeschwindigkeiten von durchschnittlich über  $4 \frac{\text{m}}{\text{s}}$  werden nur über dem Wasser östlich von Grönland, südlich von Island, (nord-)östlich der Kanaren und einem schmalen Bereich, der ausgehend von Athen über das Mittelmeer ins Rote Meer geht, erreicht. Die mittlere Windrichtung wechselt bei EUR-44 je nach Ausschnitt. Im nordwestlichen Bereich über  $60^\circ\text{N}$  liegen Nordwinde vor, welche im südwestlichen Bereich (bis  $45^\circ\text{N}$ ) auf Westwind drehen. Noch weiter südlich treten dann wieder eher nordwestliche Winde auf. Im mittleren und nordöstlichen Teil des Untersuchungsgebiets liegen hauptsächlich südwestliche Windströme vor. Für den südöstlichen Teilbereich werden dagegen eher nördliche Winde angezeigt. Insgesamt fällt zudem auf, dass besonders in höher gelegenen Gebieten „Turbulenzen“ auftreten. Die Windrichtung von benachbarten Pixeln unterscheidet sich deutlich und dreht sich teilweise sogar um  $180^\circ$ . Dies ist aufgrund der Darstellung besonders gut beim GER-11 im Alpenbereich zu erkennen. Für beide Auflösungen stimmen für das GER-11-Gebiet sowohl Windrichtung als auch Windgeschwindigkeit überein. Der Großteil von Deutschland unterliegt Strömungen mit mittleren Windgeschwindigkeiten zwischen  $0.5\text{--}2 \frac{\text{m}}{\text{s}}$ , die aus Südwesten kommen.

Zusätzlich sind in Abbildung 60 die jeweiligen Zeitreihen der täglichen räumlichen Mittel von Windgeschwindigkeit und -richtung für die in Abbildung 59B dargestellten Validationsdaten von ERA5 für GER-11 und die dazugehörigen Modelldaten von Lauf id01 gezeigt. Die Zeitreihe der Windgeschwindigkeit zeigt, dass besonders im Januar, Februar und November 2000 hohe Werte von bis zu  $8 \frac{\text{m}}{\text{s}}$  gemessen wurden. Im Sommer liegt das räumliche Mittel von GER-11, welches auch die erhöhten Windgeschwindigkeiten über Nord- und Ostsee beinhaltet, nur zwischen  $2\text{--}4 \frac{\text{m}}{\text{s}}$ . Bei der mittleren Windrichtungszeitreihe fällt ebenfalls eine zeitliche Abhängigkeit auf. Während die Windrichtung von Januar bis Ende März hauptsächlich zwischen  $200\text{--}300^\circ$  liegt, was einer westlichen Strömung entspricht, werden im April, Mai und September eher (nord-)östliche Windrichtungen angezeigt. Die starken Ausreißer sind der Tatsache geschuldet, dass bei dieser Darstellungsart nicht berücksichtigt wird, dass der y-Wert  $0^\circ$  identisch mit dem y-Wert  $360^\circ$  ist. Trotz dieser Darstellungsart lässt sich ein leichter Zusammenhang zwischen beiden Zeitreihen erkennen. Dieser lässt sich insbesondere zwischen Oktober und Dezember ablesen. Höhere Windgeschwindigkeiten werden von Winden mit westlicher Komponente (höhere Gradzahlen) erreicht. Der Unterschied zwischen den beschriebenen ERA5-Daten und den bisherigen Modellwerten von Lauf id01 ist in beiden Zeitreihen kaum erkennbar. Die Windgeschwindigkeit wird vom Modell immer leicht überschätzt. Bei der Windrichtung lässt sich kein systematischer Bias anhand der

Zeitreihen feststellen.

(A)



(B)

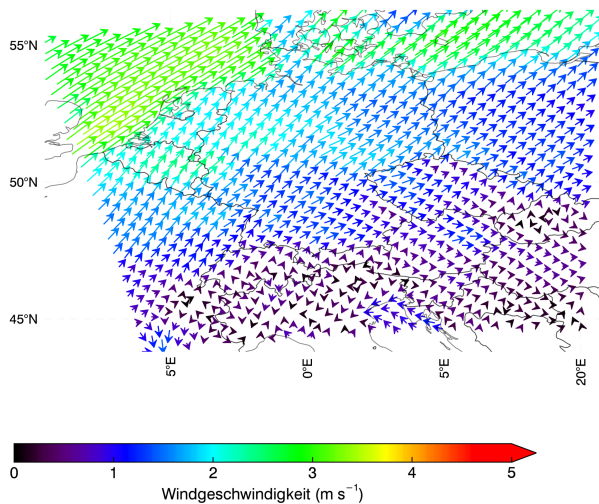


Abb. 59: Mittlere Windgeschwindigkeit und -richtung der ERA5-Daten für beide Untersuchungsgebiete für das Jahr 2000

(Die Windrichtung wird durch die Pfeilrichtung angezeigt. Die Windgeschwindigkeit kann sowohl mit der Farbskala als auch anhand der Länge der Pfeile abgelesen werden. Je länger der Pfeil, desto höher ist die Geschwindigkeit.)

(A) EUR-44 (zeitlicher Mittelwert)

(B) GER-11 (zeitlicher Mittelwert)

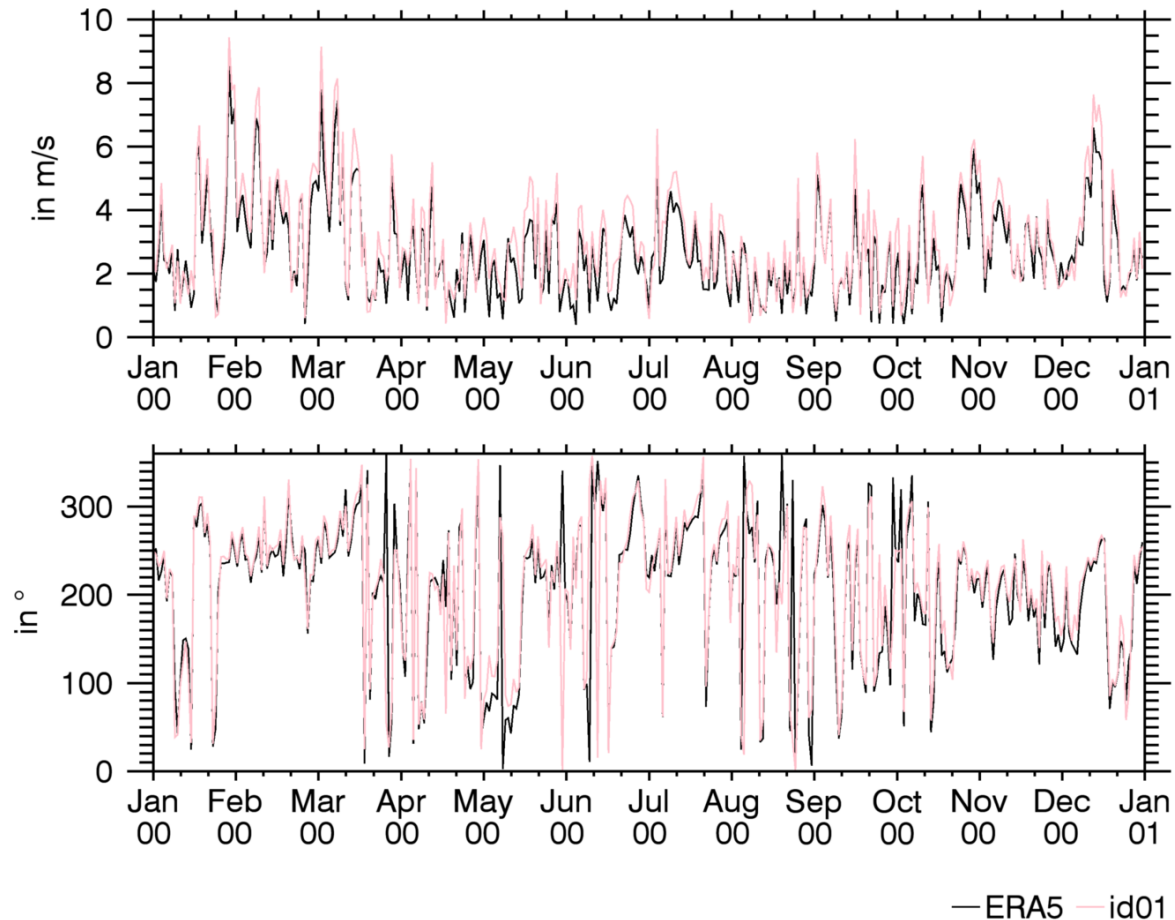


Abb. 60: Zeitreihe des räumlichen Mittelwerts der Windgeschwindigkeit von GER-11 (oben) und der mittleren Windrichtung von GER-11 (unten) für ERA5 und Modelllauf id01 für das Jahr 2000

In Kombination mit diesen ERA5-Validationsdaten, die jeweils als Histogramm in der oberen Teilabbildung von Abbildung 61 dargestellt sind, zeigt diese Abbildung ebenfalls die Differenz zum bisherigen Modelllauf id01. Dabei ist die Winkeldifferenz zwischen Modell- und Validationsdaten abhängig von der Windrichtung der ERA5-Daten aufgelistet und außerdem wird in dieser Abbildung unterschieden, ob die jeweilige Windgeschwindigkeit unterschätzt (obere Punktdiagramm-Teilabbildung) oder überschätzt (untere Punktdiagramm-Teilabbildung) wird. Jeder Punkt im Diagramm repräsentiert einen Monatswert eines einzelnen Pixels im jeweiligen Untersuchungsgebiet. An dieser Art von Abbildung können Zusammenhänge zwischen Differenzen der Richtung bzw. Geschwindigkeit, die abhängig von der Windrichtung sind, durch Häufungen der jeweiligen Punkte gezeigt werden. Außerdem enthält die Abbildung am rechten Rand die jeweiligen Histogramme der positiven und negativen Winkeldifferenz abhängig von der jeweiligen Windgeschwindigkeitsdifferenzklasse, womit die Häufigkeit der Fehler pro Winkeldifferenz und Geschwindigkeitsdifferenz angezeigt wird.

Bei Abbildung 61A, welche die Differenzen zwischen Lauf id01 und ERA5 für EUR-44 zeigt, bestätigt sich durch das obere Histogramm die Aussage von Abbildung 59A, dass im Untersuchungsgebiet westliche Winde am häufigsten vorkommen. Ebenso ist aus der Einfärbung der

Histogrammbalken abzulesen, dass der Großteil der Windgeschwindigkeiten unter  $5 \frac{\text{m}}{\text{s}}$  liegt. Lediglich aus Westsüdwesten und Westen werden höhere Windgeschwindigkeiten häufiger gemessen, sodass diese im Histogramm gut zu erkennen sind. Analysiert man die jeweilige Winkeldifferenz zwischen Modelllauf id01 und den ERA5-Daten, zeigt sich, dass kein großer Zusammenhang vorliegt. Alle Differenzpunkte liegen sehr gleichmäßig über die Winkeldifferenz-Achse verteilt, was auch an der jeweiligen Länge der Histogrammbalken abzuleiten ist. Auch die Windgeschwindigkeitsdifferenzklassen, welche sich über die Einfärbung der Punkte bzw. Balken nachweisen lässt, zeigt auf den ersten Blick keine systematischen Muster. Die Häufung der positiven Differenzpunkte der Windrichtung von  $250^\circ$  muss in Relation zur insgesamten Häufigkeit dieser Windrichtung (vgl. oberes Histogramm) gesehen werden. Dennoch scheint aus dieser Windrichtung eher eine größere Überschätzung der Windgeschwindigkeit vorzuliegen, da einerseits vermehrt Differenzen mit dunkelroter Einfärbung auftreten und andererseits weniger negative Differenzpunkte für diese Windrichtung eingezeichnet sind. Ein ähnliches, aber nicht so deutlich ausgeprägtes Phänomen ist bei der Windrichtung aus  $40^\circ$  zu erahnen. Ein Zusammenhang der Winkeldifferenz mit der Windrichtung lässt sich nicht nachweisen.

Betrachtet man Abbildung 61B, welche die Werte für das GER-11-Gebiet, erkennt man zunächst, dass sich die Windrichtungsverteilung (oberes Histogramm) für dieses Gebiet klar von den EUR-44-Werten unterscheidet. Zwar werden auch für GER-11 am meisten Westwinde nachgewiesen, die Häufung der nördlichen Winde fehlt dagegen für dieses Gebiet komplett. Dafür wird etwas mehr Wind mit südlicher Komponente angezeigt. Dies lässt sich mit Blick auf die mittleren Windrichtungen der ERA5-Karten (Abb. 59) gut erklären. Die absoluten Windgeschwindigkeiten (Einfärbung der Histogrammbalken) beider Untersuchungsgebiete unterscheiden sich ebenfalls deutlich. Vergleicht man exemplarisch die Histogramme für Westwind, stellt man fest, dass im GER-11-Gebiet mehr Winde mit Geschwindigkeiten zwischen  $2\text{--}3 \frac{\text{m}}{\text{s}}$ , aber weniger Winde mit über  $4 \frac{\text{m}}{\text{s}}$  vorkommen. Für die angesprochenen Winde mit einer nördlichen Komponente werden bei GER-11 nur schwache Winde mit Geschwindigkeiten unter  $1 \frac{\text{m}}{\text{s}}$  gezählt. Bei der Betrachtung der Winkeldifferenzhistogramme fällt auf, dass wieder die positiven Differenzen, die eine Überschätzung der Windgeschwindigkeit durch das Modell anzeigen, höhere Häufigkeiten über alle Winkeldifferenzklassen erreichen als die negativen Differenzen. Diese Überschätzung der Windgeschwindigkeit ist bereits bei der Zeitreihendarstellung (vgl. Abb. 59B) angesprochen worden. Anders als bei den korrespondierenden Histogrammen des EUR-44-Gebiets kommen jedoch die Differenzklassen mit höheren Beträgen ( $>1 \frac{\text{m}}{\text{s}}$ ) kaum vor, was darauf hindeutet, dass die Windvariablen für GER-11 insgesamt besser vom Modell abgebildet werden können. Trotzdem ist auch für das GER-11-Gebiet der bereits bei der EUR-44-Analyse aufgetretene Streifen bei der Windrichtung  $250^\circ$  zu sehen. Anders als bei EUR-44 ist bei GER-11 noch deutlicher zu erkennen, dass nicht nur eine Häufung der Punkte, sondern eine tatsächlich erhöhte positive Windgeschwindigkeitsdifferenz vorliegt, die scheinbar an diese gewisse Windrichtung in den ERA5-Daten gekoppelt ist. In der Kartenanalyse ist dieses Muster jedoch aufgrund der hohen Variabilität der Werte pro Monat und Pixel nicht eindeutig lokalisierbar.

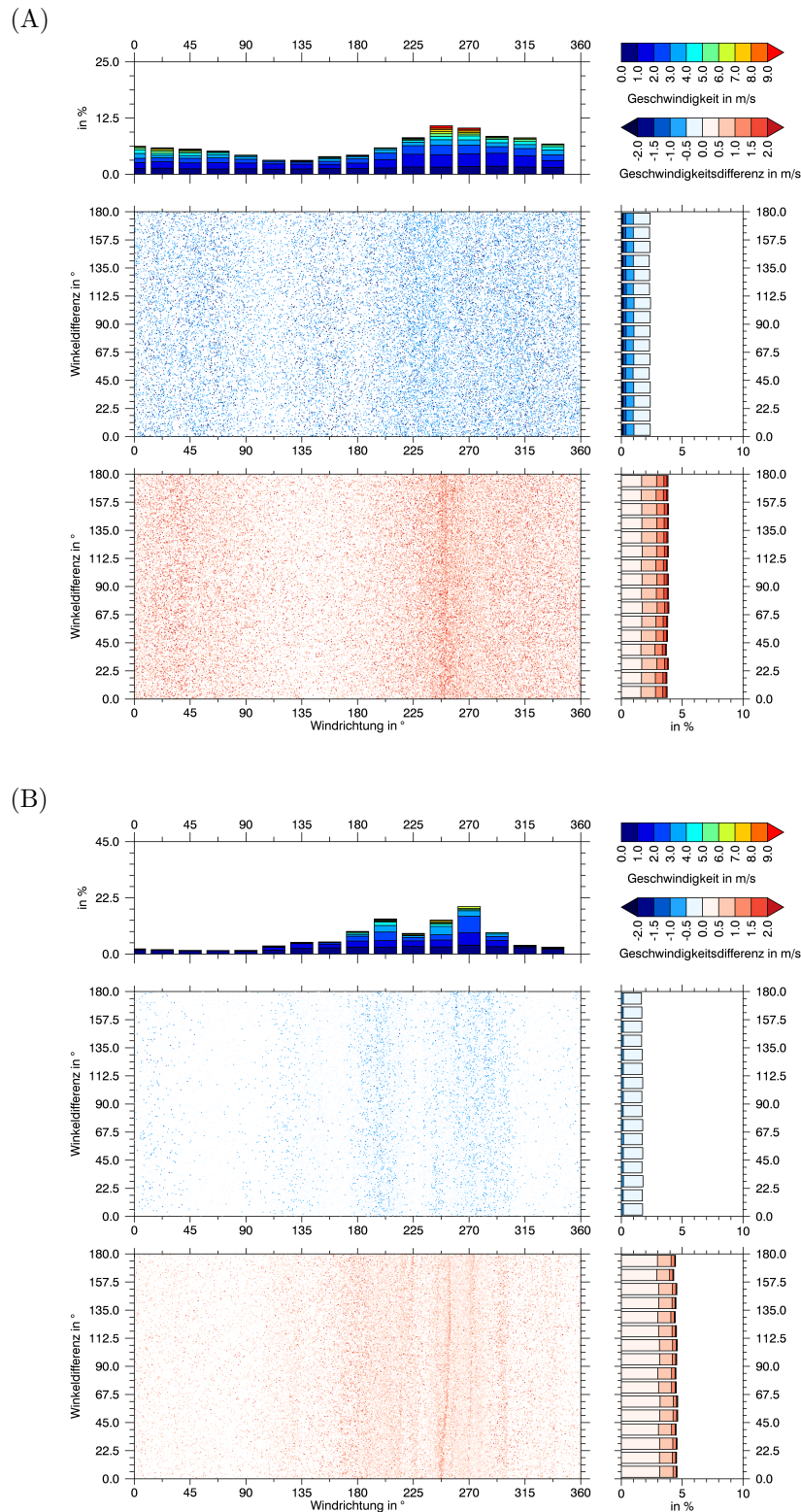


Abb. 61: Unterschied der Windrichtung und -geschwindigkeit von Modelllauf id01 und ERA5 abhängig von der Windrichtung (Datengrundlage: pixelweise Monatsmittelwerte des Jahres 2000)

(A) EUR-44 / (B) GER-11

Histogramm (oben): relative Häufigkeit der Windrichtung von ERA5 aufgeteilt nach der jeweiligen Windgeschwindigkeit / Histogramm (rechts): relative Häufigkeit der Winkeldifferenz zwischen id01 und ERA5 aufgeteilt nach der Windgeschwindigkeitsdifferenz / Punktdiagramm: Winkeldifferenz pro Windrichtung von ERA5 aufgeteilt nach positiven und negativen Windgeschwindigkeitsdifferenzen (Stärke der Differenz siehe Einfärbung)



Um die räumlichen Unterschiede zwischen Validations- und Modelldaten besser nachvollziehen zu können, sind in Abbildung 62 jeweils für EUR-44 und GER-44 die durchschnittlichen Differenzen zwischen Lauf id01 und ERA5 für die Windgeschwindigkeit sowie die mittlere Winkeldifferenz für JJA dargestellt. Der Sommer wurde exemplarisch ausgewählt, weil diese Jahreszeit die Differenz der durchschnittlichen Windgeschwindigkeit von EUR-44 sehr gut repräsentiert, obwohl die Differenz von SON im Vergleich zu den anderen Jahreszeiten deutlich andere Differenzmuster mit teilweise umgedrehten Vorzeichen zeigt. Bei der Winkeldifferenz ist für EUR-44 eine zeitliche Nord-Süd-Verlagerung der größten Differenzen zu erkennen. Die vermehrt größeren Differenzen werden im dargestellten Sommer vor allem im Norden ausgewiesen, während im Winter vor allem im Süden höhere Abweichungen vorliegen. Beim GER-11-Gebiet ist dieser Effekt nicht nachweisbar. Alle Jahreszeiten zeigen ein sehr ähnliches Winkeldifferenzmuster, welches eindeutig von der Topographie abhängig ist. Auch die Windgeschwindigkeitsdifferenz zeigt beim GER-11-Gebiet einen klaren Zusammenhang mit der topographischen Höhe. Beide Differenzmuster unterscheiden sich jedoch, was sich beispielsweise gut an der Po-Ebene und ihrer südlich gelegenen Gebirgskette nachweisen lässt. Während im Gebirge die höchste Überschätzung der Windgeschwindigkeit ( $1.5\text{--}2 \frac{\text{m}}{\text{s}}$ ) gemessen werden, wird bei der Windrichtung die größte Abweichung in der Ebene festgestellt. Im gesamten nördlichen Teil des GER-11-Gebiets, welches als Meer oder Flachland definiert wird, werden bei beiden Variablen geringere Differenzen zwischen Modell- und Validationsdaten ausgewiesen.

#### **Auswirkung der neuen Topographiedaten auf Windvariablen**

Aufgrund der teilweise größeren Höhendifferenzen zwischen dem bisherigen und dem neuen Topographiedatensatz, welche in Kapitel 6.3.1 genauer vorgestellt worden sind, wird nachfolgend der Effekt dieser Änderung auf die Windvariablen (10 m) untersucht. In Abbildung 63A (für EUR-44) und 64A (für GER-11) sind dazu, nach dem gleichen Prinzip wie bei Abbildung 61, die Häufigkeit der Winde pro Windrichtung und -geschwindigkeit sowie die jeweilige Winkel- und Windgeschwindigkeitsdifferenz pro Windrichtung zu sehen. Zu beachten ist, dass in diesem Fall jedoch jeweils die Daten des Modelllaufs id01 statt der ERA5-Daten als Referenzdaten verwendet werden, um Unterschiede zwischen den Modellversionen besser detektieren zu können. Daher ist auch die Differenzlegende mit kleineren Werten aufgetragen. Außerdem werden in dieser Abbildung nur die Daten von Juni bis August statt des ganzen Jahres 2000 benutzt, um genauer auf die Effekte der Topographieänderung eingehen zu können, ohne dass diese durch die jahreszeitlichen Schwankungen der Windvariablen überlagert werden. Zusätzlich zu den einzelnen Pixelwerten sind jeweils die Mittelwerte dieser Differenzen (id01-id01\_AFT für JJA 2000) sowie die tatsächlichen Werte von Lauf id01 in den beiden Abbildungen als Karten dargestellt, wodurch ein Vergleich mit den in Abbildung 62 dargestellten Differenzen zu den Validationsdaten für den gleichen Zeitraum möglich ist.

Aus dem Histogramm der modellierten Windgeschwindigkeit pro Windrichtung in Abbildung 63A lässt sich zunächst ableiten, dass im Sommer 2000 bisher überwiegend Winde aus West, Nordwest und Nordost für EUR-44 modelliert werden, wobei dies mit Blick auf die Karte des Mittelwerts der Windrichtung (Abb. 63D) vor allem eine räumliche Unterteilung ist. Im Norden und

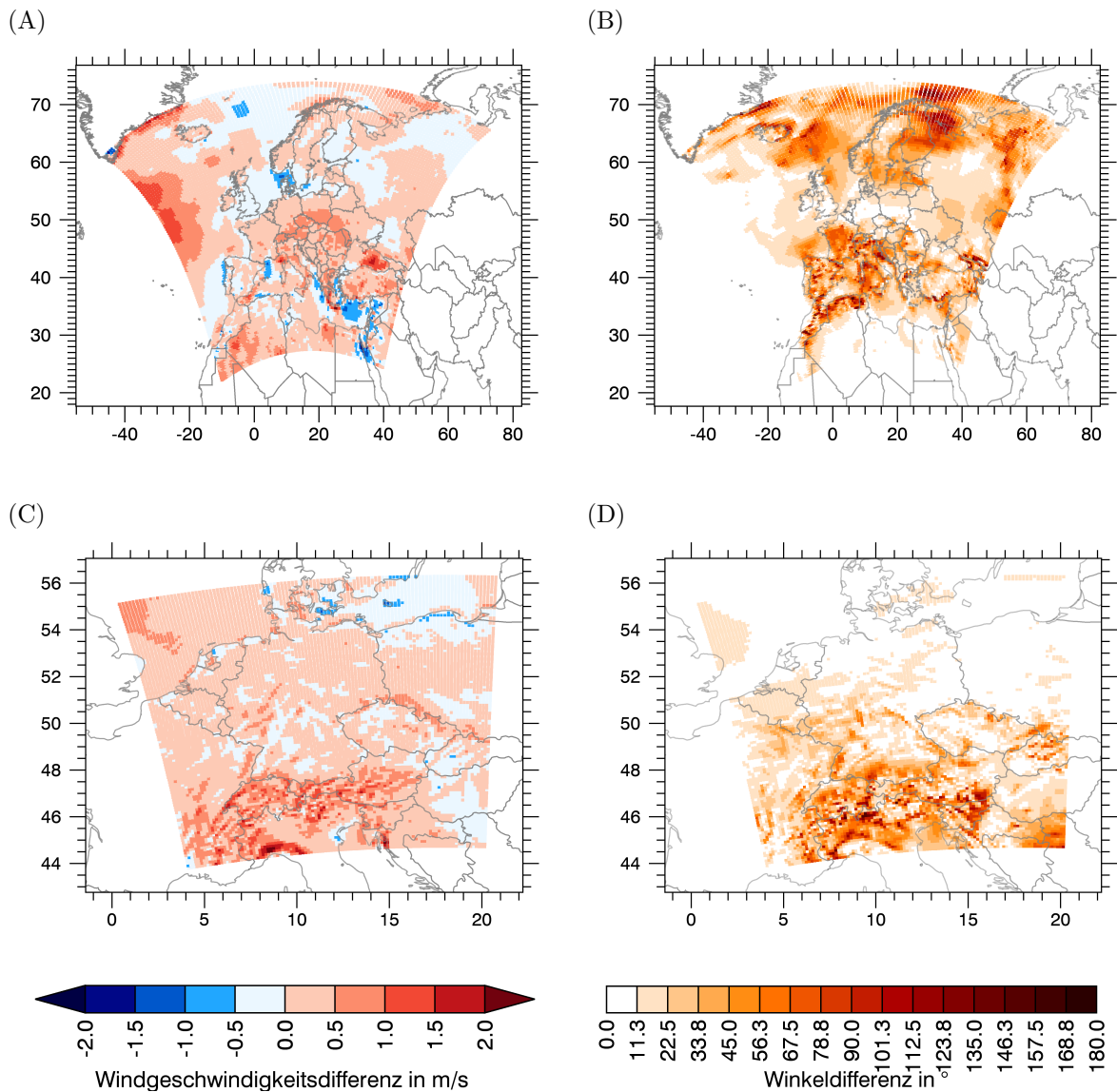


Abb. 62: Differenzkarten für Windgeschwindigkeit und -richtung zwischen Modelllauf id01 und ERA5  
 (Datengrundlage: Monatsmittel JJA 2000)  
 (A) Differenz der Windgeschwindigkeit (EUR-44)  
 (B) Winkeldifferenz (EUR-44)  
 (C) Differenz der Windgeschwindigkeit (GER-11)  
 (D) Winkeldifferenz (GER-11)

Süden des Untersuchungsgebietes werden überwiegend die Winde mit nördlicher Komponente angezeigt, in der Mitte des Gebiets (40–60°N) dagegen die Westwinde. Die durchschnittliche Windgeschwindigkeit (Einfärbung der Histogrammbalken) liegt für die häufiger vorkommenden Windrichtungen relativ gleichverteilt auf diese Richtungen zwischen 0–5  $\frac{m}{s}$ , wobei die Klasse mit 2–3  $\frac{m}{s}$  jeweils am häufigsten vorkommt. Der Vergleich beider Modellläufe ergibt etwas mehr positive Windgeschwindigkeitsdifferenzen, welche durch die Werte über 0.2  $\frac{m}{s}$  ausgelöst wird, da die kleineren Differenzbetragsklassen alle gleichmäßig verteilt sind (Länge der Histogrammbalken). Wegen des kleineren Wertebereichs auf der x-Achse der Histogramme (verglichen mit der Achse

von Abb. 61) sind kleine Unterschiede zwischen den jeweiligen Winkeldifferenzbalken zu erkennen, die jedoch im Promillebereich liegen. Auch an den beiden Punktdiagrammen zeigt sich keine systematische Abhängigkeit zwischen der Windrichtung von Lauf id01 und der resultierenden Winkeldifferenz zu Lauf id01\_AFT, welche aufgrund der konstanten Topographieänderung und der damit verbundenen Strömungsänderung zu erwarten wäre. Lediglich eine kleine Häufung bei einer Windrichtung von knapp  $270^\circ$  speziell bei einer Winkeldifferenz von  $50^\circ$ , die gleichzeitig mit höheren positiven Windgeschwindigkeitsdifferenzen ausgewiesen wird, fällt auf.

Für das EUR-44-Gebiet lässt sich aus Abbildung 63B und 63C erkennen, dass durch die neue Topographie zuvor niedrigere Windgeschwindigkeiten tendenziell höher und zuvor höhere Windgeschwindigkeiten eher niedrigerer modelliert werden. Eine Ausnahme hiervon ist das bereits bei der Analyse der ERA5-Jahresdaten angesprochene Gebiet bei Athen, welches durch hohe Windgeschwindigkeiten aufgefallen ist. Dieses Gebiet erreicht bei Lauf id01 deutlich höhere Geschwindigkeiten als die Umgebung und mit den ASTER-Daten werden noch höhere Windgeschwindigkeiten modelliert. In Verbindung mit Abbildung 62A lässt sich schließen, dass durch ASTER in der Windgeschwindigkeit eine Verbesserung über Deutschland, Tschechien, Russland und vereinzelt Teilen Afrikas erreicht wird. Der Betrag der Differenz zwischen id01\_AFT und id01 ist aber sehr viel geringer als der zwischen id01 und ERA5, sodass alle Änderungen eher als minimal einzuschätzen sind. Ein räumlicher Zusammenhang zwischen den Topographieänderungen (Abb. 47) und der Windgeschwindigkeitsdifferenz lässt sich nicht herstellen.

Die Flächen mit der höchsten Winkeldifferenz zwischen dem bisherigen Lauf id01 und ERA5 (vgl. Abb. 62A) werden beim Vergleich von Lauf id01\_AFT mit id01 (vgl. Abb. 63E) ebenfalls mit den höchsten Unterschieden gekennzeichnet. Aus dieser Analyse kann aufgrund der Eigenschaften der Windrichtung jedoch nicht festgestellt werden, ob durch die neuen ASTER-Daten die Windrichtung besser als zuvor mit den ERA5-Werten übereinstimmt. Dies zeigt aber, dass die Abweichungen der Richtung zwischen den bisherigen Modell- und Validationsdaten mit der Topographie zusammenhängen muss, da durch die Änderung der Höhe an diesen Stellen im Untersuchungsgebiet die größten Effekte auftreten. Dieser räumliche Zusammenhang ist bei der Windrichtungsanalyse für GER-11 (Abb. 62D und Abb. 64) jedoch nicht zu sehen.

Beim Vergleich des Deutschlandausschnittes für beide Auflösungen zeigt sich, dass vor allem die modellierte Windgeschwindigkeit durch die höhere Auflösung für Deutschland zunimmt (Abb. 63B vs. Abb. 64B). Während bei EUR-44 durchschnittlich Geschwindigkeiten von  $0.5\text{--}1.5 \frac{\text{m}}{\text{s}}$  für den Sommer erreicht werden, sind es bei GER-11  $1.0\text{--}2.5 \frac{\text{m}}{\text{s}}$ . Zudem sind bei GER-11 allein aufgrund der Auflösung wesentlich mehr Details zu erkennen, was auch bei der Windrichtung (Abb. 64D) auffällt. Der Unterschied zwischen den beiden Modellversionen hingegen ist für die höhere Auflösung wesentlich geringer. Dies spiegelt sich zum einen bei den Differenzkarten der beiden Windvariablen (Abb. 64C und Abb. 64E) und zum anderen in den dargestellten Differenzhistogrammen von Abbildung 64A wider. An den Histogrammen sieht man, dass sowohl positive als auch negative Geschwindigkeitsdifferenzen bis  $0.05 \frac{\text{m}}{\text{s}}$  mit weitem Abstand am häufigsten vorkommen, was sich grundsätzlich von den Differenzhistogrammen bei EUR-44 unterscheidet, wo die jeweils höchsten Beträge am häufigsten vorkommen. Zudem kann bei den

GER-11-Winkeldifferenzen anders wie bei allen bisherigen Analysen eine exponentielle Abnahme statt einer Gleichverteilung der relativen Häufigkeiten beobachtet werden. Das heißt, für GER-11 unterscheiden sich die beiden Modellversionen nur geringfügig.

Der aus diesen Analysen nicht eindeutig identifizierbare Zusammenhang mit den Topographie-differenzen weist auf den komplexen Zusammenhang hin, wie die Topographie im Klimamodell verwendet wird und die Windvariablen beeinflusst. Eine direkte Analyse der u- und v-Komponente statt der daraus abgeleiteten Windgeschwindigkeit und -richtung wird in Kapitel 6.5.2.2 für andere ausgewählte Modellläufe gemacht, weshalb an dieser Stelle darauf verzichtet wird, die Windänderungen weiterer Modellversionen miteinander zu vergleichen.

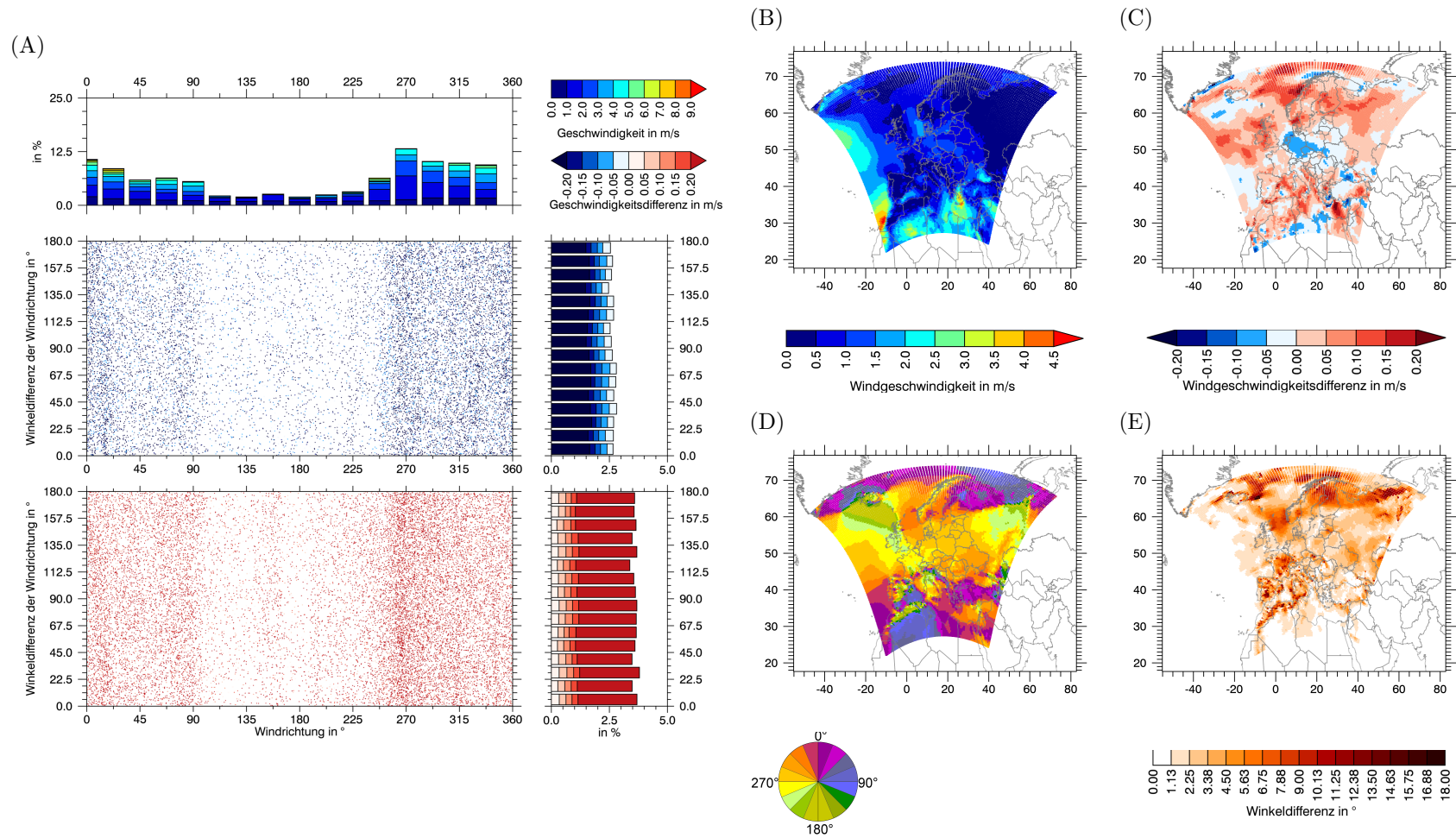


Abb. 63: Unterschied der Windrichtung und -geschwindigkeit von Modelllauf id01\_AFT und id01 für EUR-44

(Datengrundlage: pixelweise Monatsmittelwerte des JJA 2000)

(A) Windrichtungsabhängige Analyse (Erklärung der Teilabbildungen siehe Abb. 61) / (B) Windgeschwindigkeit (id01) / (C) Windgeschwindigkeitsdifferenz (id01\_AFT-id01) / (D) Windrichtung (id01) / (E) Winkeldifferenz (id01\_AFT-id01)

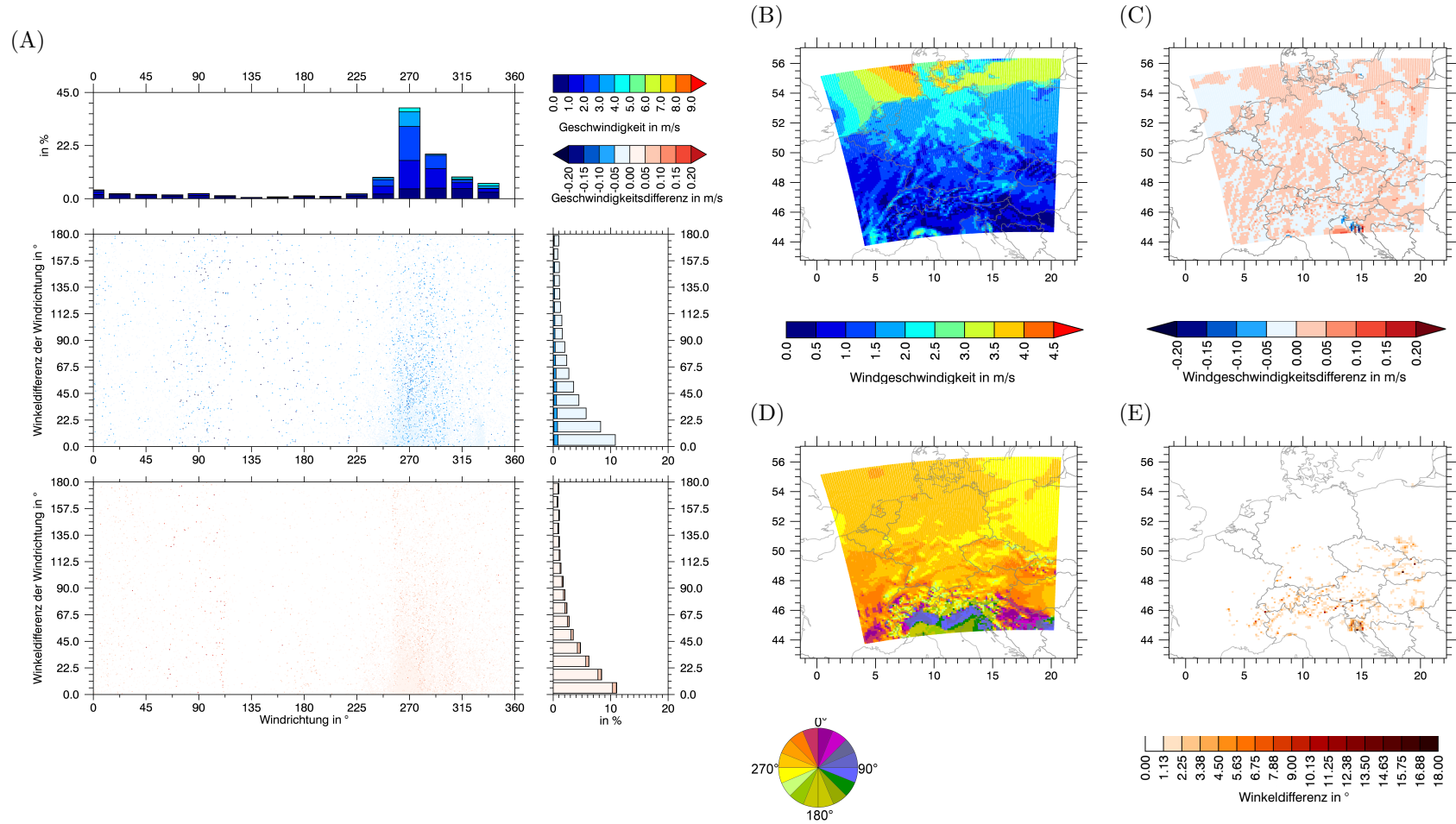


Abb. 64: Unterschied der Windrichtung und -geschwindigkeit von Modelllauf id01\_AFT und id01 für GER-11

(Datengrundlage: pixelweise Monatsmittelwerte des JJA 2000)

(A) Windrichtungsabhängige Analyse (Erklärung der Teilabbildungen siehe Abb. 61) / (B) Windgeschwindigkeit (id01) / (C) Windgeschwindigkeitsdifferenz (id01\_AFT-id01) / (D) Windrichtung (id01) / (E) Winkeldifferenz (id01\_AFT-id01)

## Veränderungen in den verschiedenen Wärmeflüssen und den 2 m-Temperaturen

Wie sich die Datensatz- und Parametrisierungsänderungen auf die Wärmeflüsse auswirken, wird durch einen Vergleich mit den Oberflächenenergiebilanzvariablen von ERA5 festgestellt. Es werden sowohl der sensible und latente Wärmefluss (AHFS und AHFL) als auch die langwellige und kurzwellige Nettostrahlung an der Oberfläche (SRADS und TRADS) dargestellt. Zudem kann aus diesen der Bodenwärmefluss berechnet werden, welcher als „Energieförderer“ einen großen Einfluss auf die Bodentemperaturen hat. Eng mit den Energieflüssen verknüpft sind auch die drei 2 m-Temperaturen (T2M, T2MIN und T2MAX), weshalb die Änderungen in diesen Variablen größtenteils parallel interpretiert werden.

### Auftretende Änderungen durch neue Datensätze (Experimentplan 1)

Exemplarisch ist in Abbildung 65 jeweils der latente Wärmefluss von ERA5 und Modelllauf id01 (Liniendiagramm) sowie deren Differenz und die Differenz zwischen den anderen Modellläufen von Experimentplan 1 und dem bisherigen Lauf id01 (Balkendiagramme) dargestellt. Es handelt sich um die täglichen Flächenmittelwerte aller Landpunkte des EUR-44- (Abb. 65A) und des GER-11-Gebietes (Abb. 65B). Da die Differenzwerte zwischen REMO und ERA5 viel größer sind als die Unterschiede zwischen den REMO-Modellläufen, sind unterschiedliche y-Achsen gewählt worden. Zur Erleichterung des Vergleichs sind in den Modelldifferenzabbildungen die Balken grün eingefärbt, bei denen sich der Betrag der Differenz zwischen dem jeweils neuen Modelllauf und ERA5 im Vergleich zur Differenz zwischen Lauf id01 und ERA5 verringert. Grüne Balken kennzeichnen somit Modellverbesserungen, während rosa Balken für Verschlechterungen stehen. Es ist zu beachten, dass ein negativer Wärmeflusswert einen Energieverlust darstellt. Das heißt beispielsweise, ein negativer latenter Wärmewert steht für Verdunstung.

Die latente Wärme von Modelllauf id01 besitzt im Vergleich zu den ERA5-Daten für das EUR-44-Gebiet einen positiven Bias von  $10\text{--}20 \frac{\text{W}}{\text{m}^2}$ . Es ist zu erkennen, dass der jeweilige Tagesmittelwert von ERA5 immer negativ ist, in den Wintermonaten bei REMO id01 dagegen auch leicht positive Werte vorkommen. Das heißt, die Verdunstung wird von REMO unterschätzt (vgl. Abb. 65A). Für das GER-11-Gebiet tritt dieser Effekt noch stärker zutage und resultiert in noch größeren Differenzwerten zwischen Modell- und Validationsdaten (vgl. Abb. 65B). Bei Betrachtung der Änderungen, die sich durch die neuen Datensätze im EUR-44-Gebiet ergeben, lässt sich feststellen, dass größere Unterschiede in den Differenzmustern vor allem durch die Verwendung von Textur bzw. Sand-/Tonanteil (vgl. LSOIL-Schalter) im Sommer auftreten (sichtbar bei id01\_AST und id01\_ASF im Juni-Juli, bei id01\_GFT und id01\_GFF im Juni-August, bei id01\_GST und id01\_GSF nur vereinzelt im Juni und besonders deutlich bei GHT und GHF im April-Juli). Während bei id01\_AST/ASF und id01\_GHT/GHF die Verwendung des Sand-/Tonanteils (LSOIL=F) eindeutig zu Verbesserungen führt, ist der Effekt bei id01\_GST/GSF und id01\_GFT/GFF nicht so klar. Bei der gleichen Analyse des GER-11 tritt dieser Unterschied nur zwischen id01\_GFF und id01\_GFT auf. Generell führt die Datensatzkombination ASF (von Juni-September) und besonders die Kombination GHF zu geringeren Biaswerten in EUR-44. Auffällig ist zudem, dass die ersten ca. 20 Tage im Dezember in allen neuen Modellläufen wesentlich schlechter abschnei-

den als der bisherige Modelllauf id01. Bei GER-11 sind vor allem bei id01\_AST, id01\_ASF, id01\_GSF und id01\_GHF längere Phasen mit Verbesserungen (v. a. im Sommer) erkennbar, wenngleich der Betrag der Änderungen für das GER-11-Gebiet deutlich geringer ausfällt als beim dargestellten EUR-44. Dennoch scheinen die Datensatzänderungen für das Deutschlandgebiet eher zu systematischen Änderungen im Jahresgang zu führen, während beim gröber aufgelösten Europegebiet eher die intrasaisonale Variabilität verändert wird, wenngleich beachtet werden muss, dass für EUR-44 der Mittelwert über ein viel größeres Gebiet gebildet wurde, welches im bereits für den absoluten latenten Wärmeflusswert viel variabler ist.

Bei der sensiblen Wärme, deren Auswertung nicht gezeigt wird, wird der ERA5-Wert - im Gegensatz zur latenten Wärme - von REMO betragsmäßig überschätzt. Insbesondere im Sommer liegen die Differenzen bei bis zu  $40\text{--}70 \frac{\text{W}}{\text{m}^2}$  je nach Untersuchungsgebiet. Die Datensatzänderungen führen bei allen neuen Modellläufen im EUR-44-Gebiet durchschnittlich zur Minimierung des Bias, während bei der Analyse des GER-11-Gebietes nur in den Wintermonaten leichte Verbesserungen, im Sommer jedoch sowohl höhere als auch niedrigere Differenzen erreicht werden. Damit verhalten sich die jeweiligen Werte für sensible und latente Wärmeflüsse in allen REMO-Läufen, vor allem für die Werte des GER-11-Gebietes, fast konträr zueinander. In dem Zeitraum, in dem sich der Wert von AFHS verbessert, verschlechtert sich AHFL und umgekehrt. Dadurch, dass der jeweilige Betrag der Differenzen jedoch nicht gleich ist, werden vor allem im Sommer vermehrt bessere Modellergebnisse für den Bodenwärmefluss erzielt, der auch zur Bilanzgleichung gehört (vgl. Formel 2.3).

Bei der Analyse der langwelligen und kurzwelligen Nettostrahlung für EUR-44 und GER-11, die ebenfalls nicht graphisch dargestellt sind, lässt sich zunächst festhalten, dass für EUR-44 beide Strahlungsvariablen von Lauf id01 im Vergleich zu ERA5 besonders im Sommer unterschätzt werden (ca.  $20 \frac{\text{W}}{\text{m}^2}$ ). Bei GER-11 ist den Differenzen besonders bei der langwelligen Strahlung kein Jahresgang zu entnehmen und die Differenzen schwanken zwischen  $\pm 10 \frac{\text{W}}{\text{m}^2}$  für die langwellige bzw.  $\pm 20 \frac{\text{W}}{\text{m}^2}$  für die kurzwellige Strahlung. Die Datensatzänderungen führen bei beiden Strahlungen in allen Modellläufen von EUR-44 über den gesamten Zeitraum aber besonders am Anfang des Jahres zu größeren Differenzen. Dagegen treten bei GER-11 bei der langwelligen Strahlung kaum Unterschiede mit mehr als  $0.5 \frac{\text{W}}{\text{m}^2}$  auf, während sich für die kurzwellige Strahlung vor allem von Mai bis September hauptsächlich positive Änderungen von bis zu  $2 \frac{\text{W}}{\text{m}^2}$  in allen neuen Modellläufen ergeben.

Wie aus den oben genannten Beispielen hervorgeht, ist der Vergleich zwischen den unterschiedlichen REMO-Modellläufen und die Bewertung der Änderungen schwierig. Es lassen sich nur vereinzelt charakteristische Merkmale erkennen. Beispielsweise ergeben sich für die Wintermonate klare Verbesserungen der Differenz des latenten Wärmeflusses für die Läufe, die ASTER als Topographie verwenden. Jene Verbesserungen sind für die anderen Läufe nicht zu erkennen. Gleichzeitig wird dagegen der sensible und der Bodenwärmefluss für die ASTER-Läufe in DJF schlechter abgebildet. Der Grund für den starken Einfluss der Topographie auf die Wärmeflüsse ist, dass mit der Topographieänderung auch die Rauigkeitslänge geändert wird, die einen Teil des Energieaustausches zwischen Atmosphäre und Boden steuert (vgl. Kapitel 6.3.1). Dass der



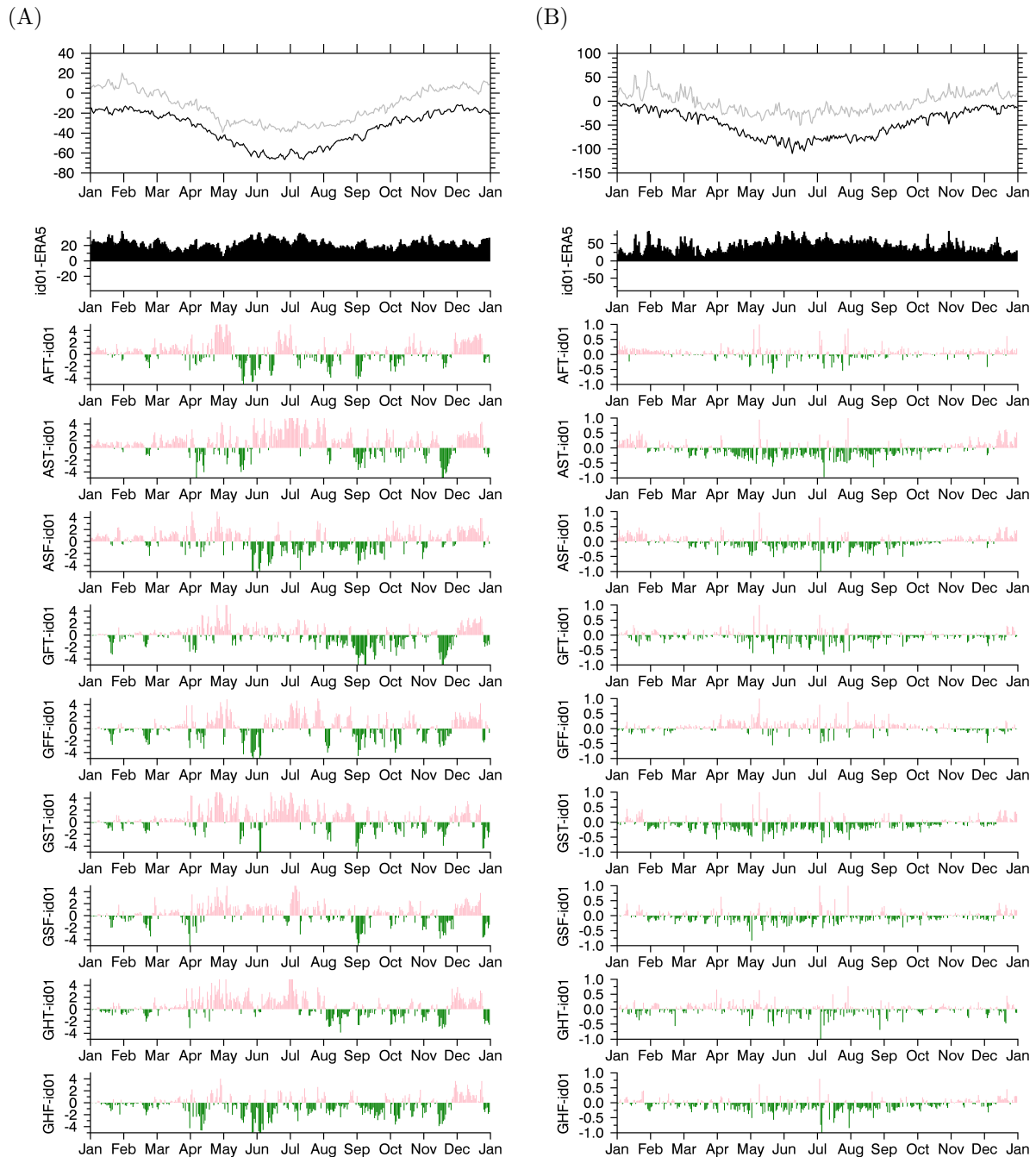


Abb. 65: Vergleich des latenten Wärmeflusses (AHFL) der verschiedenen Modellläufe von Experimentplan 1 (vgl. Abb. 44) mit den ERA5-Validationsdaten für EUR-44 und GER-11 (alle Angaben in  $\frac{W}{m^2}$ )

(A) EUR-44

(B) GER-11

1. Reihe: jeweiliges tägliches Flächenmittel aller Landpunkte der ERA5-Daten (schwarz) und des bisherigen Modelllaufs id01 (grau)

2. Reihe: Differenz zwischen Modelllauf id01 und ERA5

3.-9. Reihe: Differenz zwischen dem jeweiligen angegebenen Modelllauf und Modelllauf id01 (grüne Balken kennzeichnen eine Verbesserung und rosa Balken eine Verschlechterung durch die jeweilige Modellversion im Vergleich zu Modelllauf id01)

Für eine genauere Analyse wird die y-Achse der Differenzen für die Untersuchungsgebiete auf  $\pm 5$  bzw.  $\pm 1 \frac{W}{m^2}$  begrenzt.

Unterschied der Läufe nur im Winter aus den Flächenmitteln erkennbar ist, könnte daran liegen, dass weitere Einflussfaktoren auf die Wärme Flüsse im Winter nicht so ausgeprägt vorkommen und die jeweiligen Absolutwerte im Winter nicht so hoch sind, sodass Änderungen eher auffallen. Zusammenfassend ist festzustellen, dass es zwischen den verschiedenen Modellläufen mit unterschiedlichen Topographien größere Unterschiede gibt als zwischen denen mit unterschiedlichen Bodendaten. Da sich die räumlichen Änderungen des Bodens im Flächenmittel vermutlich ausgleichen, werden nachfolgend die jährlichen Karten der Variablen dargestellt und einige Auffälligkeiten der räumlichen Differenzen analysiert.

Abbildung 66 zeigt die jährlichen Durchschnittstemperaturen T2MIN (Abb. 66A), T2M (Abb. 66C) und T2MAX (Abb. 66F) und die durchschnittlichen Werte von sensiblem (Abb. 66D), latentem (Abb. 66B) und Bodenwärmestrom (Abb. 66G) sowie kurzweilige (Abb. 66H) und langweilige (Abb. 66E) Nettostrahlung für GER-11 aus den ERA5-Daten. Alle drei Temperaturmuster stimmen größtenteils überein. Insgesamt sind im Nord- und Nordostdeutschland jeweils die Temperaturwerte geringer als im Westen und im Süden. Zudem sind besonders die Alpen und Mittelgebirge durch kühlere Temperaturen gekennzeichnet. Außerdem ist in der Abbildung zu erkennen, dass das Bodenwärmeflussmuster relativ gut mit dem der kurzweiligen Strahlung übereinstimmt, wenngleich beachtet werden muss, dass die Legende des Bodenwärmestroms (wie die Legende des sensiblen Wärmestroms) doppelt so hohe Werte hat, wie die anderen drei dargestellten Variablen. Auch hier ist ein deutlicher Nord-Süd-Gradient zu erkennen.

Abbildung 67 ist wie Abbildung 66 aufgebaut, stellt jedoch die Differenz zwischen den jährlichen ERA5-Daten und dem Modellauf id01 dar. Wie an den Zeitreihen der Flächenmittel zu erkennen ist (die nicht dargestellte id01-Auswertung ist ähnlich wie id03 in Abb. 69C), wird T2M bis auf den Alpenbereich bereits relativ gut abgebildet. T2MIN wird dagegen vom bisherigen REMO-Lauf überschätzt. Der Großteil der Fläche ist 0.8–1.6 °C zu warm. Bei T2MAX hingegen liegt eine Unterschätzung um diesen Betrag vor. Bei beiden Variablen nehmen die Fehlerbeträge im Alpenbereich und südöstlichen Bereich zu. Die Differenz der langweiligen Nettostrahlung zeigt eine klare räumliche Abhängigkeit von der vorherrschenden Topographie (vgl. Abb. 48A). Für höher liegende Gebiete werden negative Differenzen berechnet. Aufgrund der negativen Validationswerte bedeutet das, dass der Betrag von id01 für die langweilige Nettostrahlung größer ist als der Betrag von ERA5. Folglich wird entweder weniger Energie eingestrahlt oder mehr Energie ausgestrahlt. Bei den tiefer liegenden Gebieten wird dagegen eine positive Differenz erreicht, was einem betragsmäßig kleineren Wert von id01 entspricht. Bei der kurzweiligen Strahlung sind die Vorzeichen der Differenzen genau umgedreht. Die Mittelgebirge sind zwar nicht so gut wie bei der langweiligen Strahlungsdifferenz zu erkennen, dies liegt aber an der mit auf  $4 \frac{\text{W}}{\text{m}^2}$  genormten relativ großen Klassenbreite. Auffällig ist, dass bei beiden Variablen der Betrag der Differenz an der Küste zunimmt. Beim sensiblen, latenten und beim Bodenwärmestrom tritt dies nicht auf. Die in Abbildung 69A bzw. 69B festgestellte betragsmäßige Unter- bzw. Überschätzung des latenten bzw. sensiblen Wärme flusses ist jeweils im gesamten Untersuchungsgebiet festzustellen.

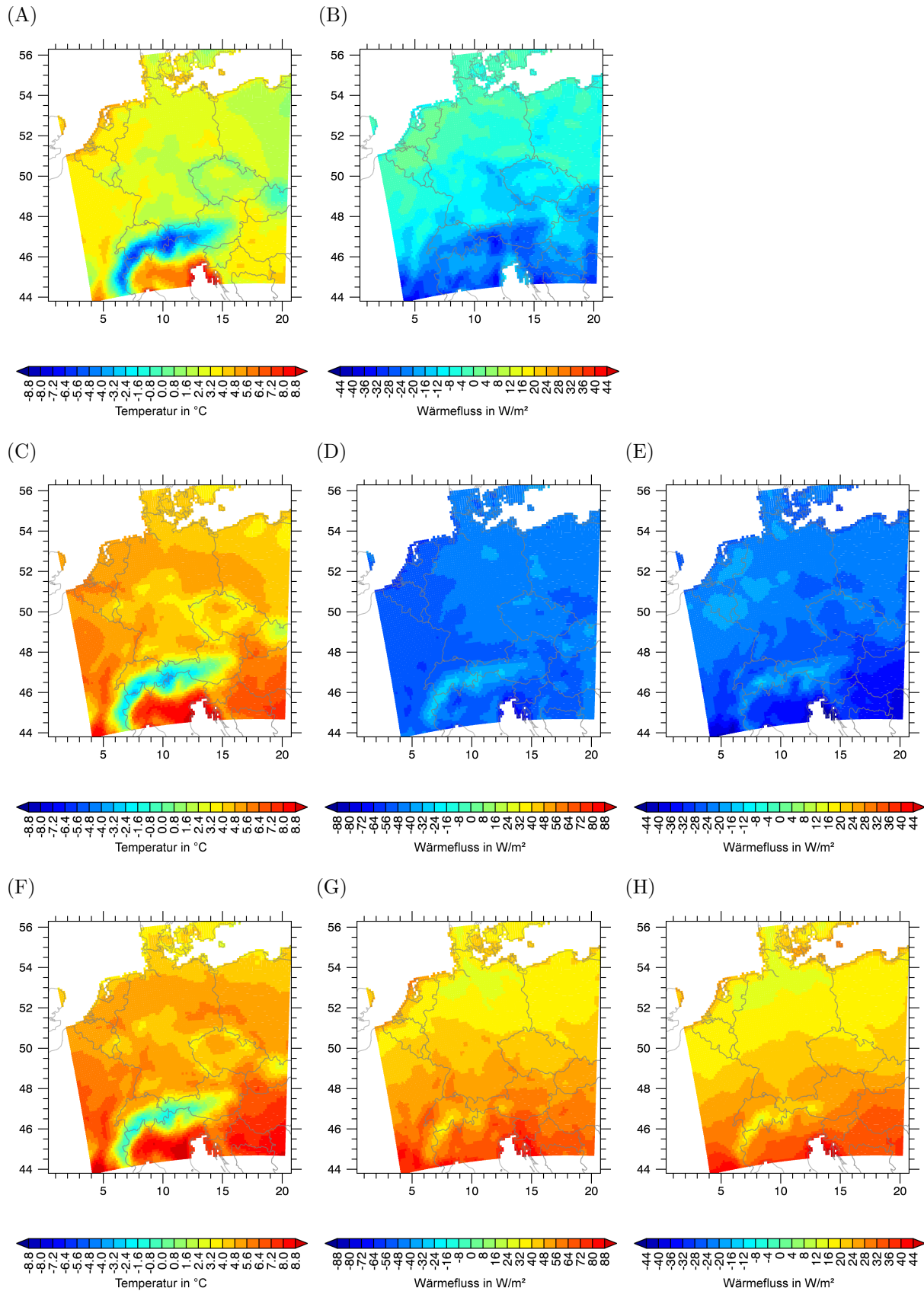


Abb. 66: Mittlere jährliche Temperatur- und Wärmeflussvariablen der ERA5-Daten für GER-11 (2000)  
 (A) Minimumtemperatur / (B) sensible Wärmeflussdichte / (C) Mitteltemperatur /  
 (D) latente Wärmeflussdichte / (E) langwellige Nettostrahlung / (F) Maximumtemperatur / (G) Bodenwärmestromdichte / (H) kurzwellige Nettostrahlung

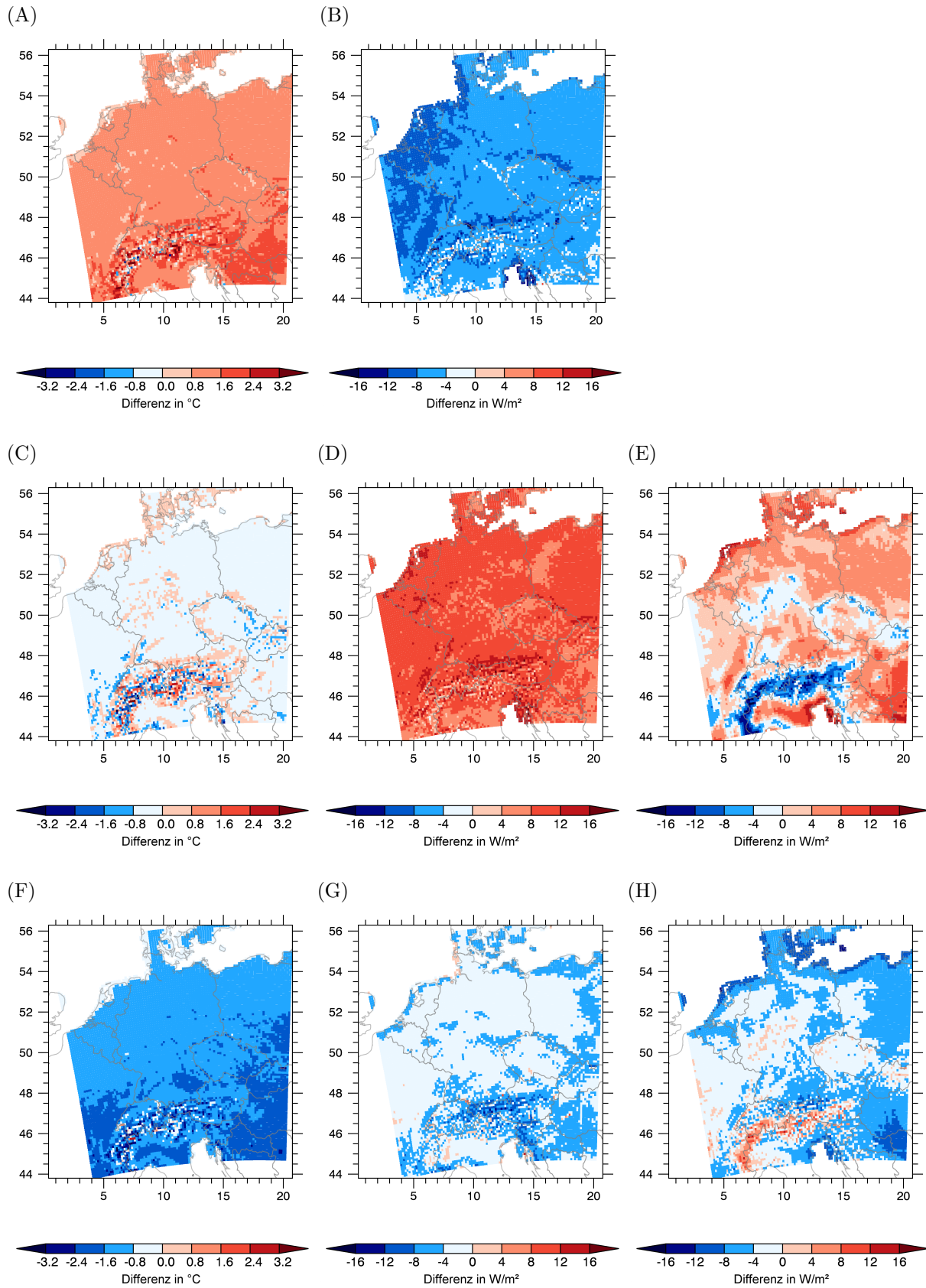


Abb. 67: Differenz der mittleren jährlichen Temperatur- und Wärmeflussvariablen zwischen dem bisherigen Modellauf id01 und den ERA5-Daten für GER-11 (2000)  
 (A) Minimumtemperatur / (B) sensible Wärmeflussdichte / (C) Mitteltemperatur /  
 (D) latente Wärmeflussdichte / (E) langwellige Nettostrahlung / (F) Maximumtemperatur / (G) Bodenwärmestromdichte / (H) kurzwellige Nettostrahlung

Um die durch die neuen Datensätze auftretenden Änderungen analysieren und die Unterschiede besser herausarbeiten zu können, wird - wie bereits bei der Analyse der Zeitreihe der Flächenmittel - aufgrund der wesentlich geringeren Differenz zwischen den unterschiedlichen Modellläufen als zwischen den Modell- und ERA5-Daten auf erstere näher eingegangen. Aus den Flächenmittelanalysen kann außerdem abgeleitet werden, dass sowohl die absoluten Werte als auch die Veränderungen von der Jahreszeit abhängen, sodass der pixelweise Vergleich zwar für das gesamte Testjahr 2000 aber auch für die jeweiligen Jahreszeiten durchgeführt wird.

Exemplarisch sind in Abbildung 68 einige Differenzkarten mit Auffälligkeiten dargestellt. Abbildung 68A zeigt beispielsweise den durchschnittliche Unterschied von T2M zwischen Modelllauf id01 und Lauf id01\_AFT. Dieses Muster steht repräsentativ für alle drei Temperaturen und alle Jahreszeiten, denn für diese Abbildung wurde - anders als in den anderen Analysen - keine Höhenkorrektur der Temperaturdaten vorgenommen. Daher entsprechen die dargestellten Differenzen unter Berücksichtigung der diskreten und kontinuierlichen Farbskalierung dem komplementären Differenzmuster zwischen ASTER und GTOPO in Abbildung 48A. Der Modelllauf id01\_AFT mit überwiegend höheren Topographiewerten im Alpenbereich erzeugt dort niedrigere Temperaturen. Umkehrt werden für niedrigere Gebiete höhere Temperaturen über den gesamten Zeitraum modelliert. Allerdings sind diese durch den neuen Topographiedatensatz ausgelösten Änderungen mit einem Betrag von unter  $0.1\text{ °C}$  bzw. in den meisten Regionen sogar unter  $0.025\text{ °C}$  sehr gering. In Abbildung 68C, welche die Differenz der mittleren Temperatur zwischen Modelllauf id01\_GST und id01 für den Sommer zeigt, ist das charakteristische Muster der veränderten Bodentextur zu erkennen. Wie bereits beim Vergleich der Bodenparameter (vgl. Abb. 55, Abb. 56 und Abb. 57) festgestellt wurde, fallen besonders der westliche Teil der Alpen und der (nord-)westliche Teil von Ungarn auf. Hier werden durchschnittlich ca.  $0.1\text{ °C}$  kühlere Temperaturen modelliert. Die Änderung des Bodendatensatzes löst damit höhere Differenzen zum Originallauf aus als die Änderung der Topographie. Aus diesem Vergleich der räumlichen Muster geht jedoch auch hervor, dass nicht alle Änderungen direkt aus der Texturkartenänderung resultieren. Die großflächige Abkühlung von  $0.025\text{--}0.1\text{ °C}$  um Paris und im Bereich des Oberrheingrabens, die bezogen auf den Betrag den durch die neue Texturklasse ausgelösten Änderungen entspricht, ist in den Parameterdifferenzkarten nicht direkt zu sehen. Bei der Amplitude der Wärmeleitfähigkeit sind in diesem Bereich beide Bodendatensätze in der gleichen Klasse. Bei PORVOL und FCAP hingegen liegt die bisherige Texturklasse jeweils eine Werteklasse höher. Gleichzeitig führt die Abnahme der thermalen Wärmeleitfähigkeitsamplitude in den Westalpen bei relativ gleichbleibenden hydrologischen Parameterwerten zum gleichen Änderungssignal in der dargestellten mittleren Temperatur. Daraus folgt, dass sich die Texturänderung und die sich ergebenden räumlichen Parameteränderungen teilweise verstärken, aber teilweise auch gegenseitig aufheben. Zudem ist dieses Texturklassenmuster zwar in den anderen Jahreszeiten ebenfalls erkennbar, aber bei weitem nicht so deutlich ausgeprägt wie im dargestellten Sommer. Die auch für JJA und diese beiden Modellläufe berechnete Differenz für die sensible Wärme (Abb. 68D) weist das Texturklassenmuster nicht aus. Nur an der markanten Stelle nordwestlich von Berlin, wo beim bisherigen Modelllauf id01 mit FAO die Texturklasse Torf war, wird der Wert der sensiblen Wärme um ca.  $0.75\frac{\text{W}}{\text{m}^2}$  höher, was jedoch aufgrund der negativen Absolutwerte einem betragsmäßig kleineren Wert für Lauf id01\_GST entspricht. Die vorher ebenso als Torf

definierte Fläche an der Küste in Nordostdeutschland, welche bei den Parameteränderungen noch gut in den Differenzkarten zu erkennen ist, ist jedoch weder bei der T2M- noch bei der sensiblen Wärmedifferenzkarte im gleichen Ausmaß bzw. überhaupt von der Umgebung abtrennbar. Abbildung 68B stellt die Differenz des latenten Wärmeflusses zwischen Modelllauf id01 und Lauf id01\_AFT für DJF dar. Aus der Zeitreihenanalyse (Abb. 65B) ging hervor, dass REMO durch die ASTER-Topographie im DJF den latenten Wärmefluss der ERA5-Daten schlechter, in den restlichen Jahreszeiten jedoch besser abbildet. Aus der Karte lässt sich ablesen, dass die durchschnittlich erhöhte Kondensation (vgl. höherer positiver Wert bei Verdunstung) im Winter nicht durch einen flächendeckenden positiven Bias auftritt. An einzelnen Stellen in den Alpen sowie an der bei der Topographiedifferenzanalyse (vgl. Kapitel 6.3.1) bereits auffälligen Küste von Kroatien sind höhere negative Differenzen von  $0.25-1 \frac{\text{W}}{\text{m}^2}$  zu erkennen. Diese Pixel stimmen sehr gut mit den Pixeln der höchsten Topographieänderungen (vgl. Abb. 48A) überein. Auch die nicht dargestellte Differenzkarte der sensiblen Wärme weist für DJF an den gleichen Stellen ähnlich große Differenzwerte aus. Die zusätzlich durchgeführte Analyse der saisonalen räumlichen Korrelationskoeffizienten der Differenz der Rauheitslänge (vgl. Abb. 48C) und der mittleren Differenz der Wärmeflüsse bestätigt, dass in DJF mit ca. 0.15 (für EUR-44) bzw. 0.2 (für GER-11) für die sensible bzw. ca. -0.15 (für beide Untersuchungsgebiete) für die latente Wärme zwar insgesamt ein geringer Zusammenhang vorliegt, dieser Wert aber jeweils deutlich höher als der Korrelationskoeffizient der anderen Saisons und der anderen Variablen ist.

Trotz der Tatsache, dass bei der Zeitreihenanalyse alle Modellläufe ähnliche durchschnittliche Differenzwerte erzielt haben, lässt Abbildung 68E, in der exemplarisch die Differenz der langwelligen Nettostrahlung für JJA dargestellt ist, erkennen, dass hier größere Unterschiede erreicht werden als bei den anderen vorgestellten Differenzkarten. Da aber sowohl positive als auch negative Werte auftreten, ist dies nicht an den analysierten Flächenmitteln ablesbar. Die Differenzmuster des GHT-Laufs korrelieren weitestgehend mit der Bodendifferenzkarte des Sandanteils (vgl. Abb. 50). Die hohen negativen Differenzen in der Po-Ebene und Nordostdeutschland werden vermutlich durch den erhöhten Sandanteil und den daraus folgenden Bodenparameteränderungen ausgelöst. Gegen diese These spricht, dass der nicht dargestellte Vergleich der langwelligen Nettostrahlung von id01 und GST in der Ebene keine großen Änderungen ergibt, obwohl die Sandanteilsdifferenz zwischen Soilgrid und FAO in dieser Region zu ähnlichen Differenzen führt wie HWSD und FAO nur mit umgekehrten Vorzeichen. Außerdem zeigt sich die positive Anomalie der Nettostrahlung zwischen id01\_GHT und id01 im Ruhrgebiet ebenfalls nicht in den Bodenausgangsdaten, sodass die Differenzen in der Nettostrahlung nicht linear aus den geänderten Sandanteilen erschlossen werden können. Setzt man die Änderung der langwelligen Strahlung mit den bisherigen Differenzen zu den ERA5-Daten (vgl. Abb. 67E) in Relation, wird der Bias der langwelligen Nettostrahlung durch den GHT-Lauf in diesen Regionen reduziert. Gleichzeitig sind sehr ähnliche Differenzmuster zwischen id01 und id01\_GHT in unterschiedlicher Ausprägung bei allen Temperaturvariablen zu erkennen. Für die Po-Ebene resultieren die Strahlungsänderungen zum Beispiel in insgesamt höheren Temperaturen ( $0.15-0.2^\circ\text{C}$ ), sodass die T2MIN im Vergleich zu den ERA5-Daten schlechter und die T2MAX besser nachgebildet werden. Die angesprochenen Gebiete fallen auch beim Bodenwärmestrom mit betragsmäßig höheren Differenzen auf. Bei der kurzwelligen Nettostrahlung und bei der sensiblen und latenten Wärmeflussvariable dagegen

ergeben sich andere Muster und das, obwohl bei der Korrelationsanalyse der verschiedenen Variablen jeweils ein mittlerer bis hoher Zusammenhang zwischen allen hier analysierten Variablen berechnet wurde.

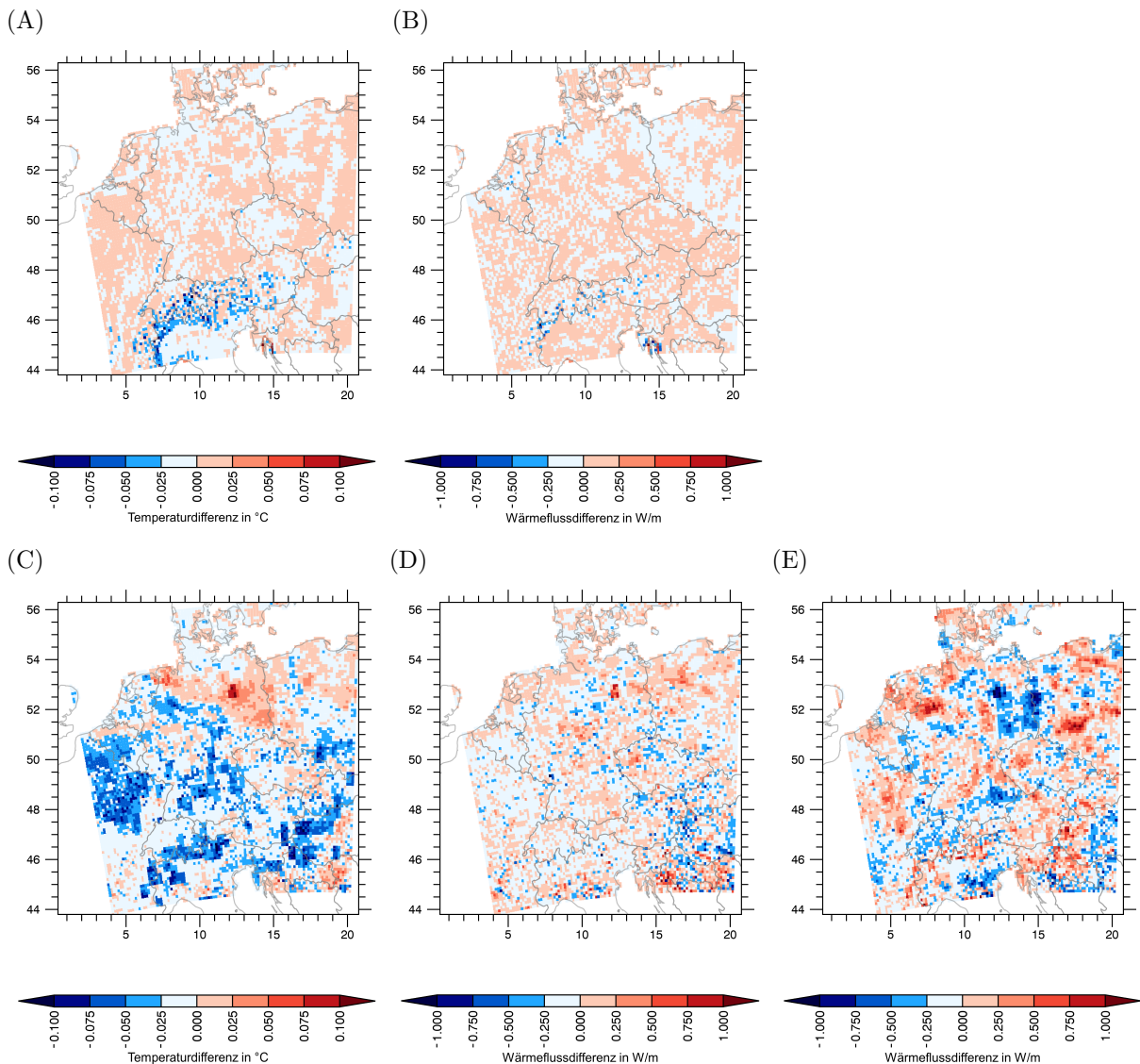


Abb. 68: Differenz verschiedener Modellläufe von Experimentplan 1 (vgl. Abb. 44) und verschiedener Variablen zu Modellauf id01 für GER-11 (2000)

(A) jährliche Differenz: mittlere Temperatur (AFT) / (B) DJF Differenz: latenter Wärmefluss (AFT) / (C) JJA Differenz: mittlere Temperatur (GST) / (D) JJA Differenz: sensibel Wärmeflussdichte (GST) / (E) JJA Differenz: langwellige Nettostrahlung (GHT)

### Auftretende Änderungen durch neue Schalter (Experimentplan 2)

Führt man die gleichen Analysen für diese Variablen mit den Läufen von Experimentplan 2 durch, lässt sich am dargestellten Beispiel der täglichen latenten und sensiblen Wärmeflüsse sowie der mittleren Temperatur für das GER-11-Gebiet (vgl. Abb. 69) erkennen, dass die Beträge der Differenzen der verschiedenen Modellläufe, die durch die Parametrisierungsänderungen ausgelöst

werden, höher sind als die vorher analysierten Änderungen durch die Datensätze (vgl. Abb. 65). Zu beachten ist, dass der Modellreferenzlauf in Abbildung 69 nicht id01, sondern id03 ist. Diese beiden Läufe unterscheiden sich darin, dass bei Modellversion id01 die bisherige FAO-Texturkarte und bei Modellversion id03 bereits die Soilgrid-Texturkarte und damit einhergehend die neuen abgeleiteten Parameter für 12 statt 6 Klassen verwendet werden sowie die neue Strahlungsfrequenz benutzt wird (vgl. Abb. 46). Dieser Wechsel des Bezugmodelllaufs ist notwendig, da die verschiedenen Schalterkombinationen aus Experimentplan 2 mit Modellversion id03 durchgeführt worden sind. In Abbildung 69B ist eine annähernd durchgängige und klare Verbesserung der sensiblen Wärme (AHFS) fast im gesamten Zeitraum in allen neuen Modellläufen zu erkennen. Gleichzeitig wird die latente Wärme (AHFL) (Abb. 69A) etwas schlechter abgebildet. Wie bei der ebenfalls dargestellten mittleren 2 m-Temperatur (Abb. 69C) ist ein Unterschied zwischen den Läufen mit dem bisherigen und dem neuen Bodentemperaturschema am Anfang und am Ende des dargestellten Zeitraums zu sehen. Dadurch dass mit dem neuen Bodentemperaturschema höhere mittlere Wintertemperaturen erreicht werden, verbessert sich der Bias, da in diesem Zeitraum REMO bisher 1–2 °C zu kalt war. Auch in den anderen Jahreszeiten werden durch die Parameteränderungen vermehrt positive Auswirkungen erzielt. Allerdings ist ebenfalls anzumerken, dass die mittlere Temperatur bereits sehr gut vom bisherigen Modell nachgebildet wird und die Beträge der Verbesserung klein sind.

Die berechnete, aber nicht dargestellte Analyse der anderen Variablen ergibt, dass die minimale Temperatur dagegen über den gesamten Zeitraum überschätzt und die maximale Temperatur unterschätzt wird. Da sowohl Minimum- als auch Maximumtemperatur tendenziell durch die Änderungen bis auf die angesprochenen Zeitpunkte im Januar, Februar und Ende Dezember abnehmen, wird die T2MIN besser und T2MAX schlechter. Durch die Erwärmung im Winter fällt weniger Schnee, wodurch die Oberflächenalbedo abnimmt, da Schnee eine sehr hohe Albedo besitzt. Folglich ändert sich das Verhältnis von einfallender und reflektierter Solarstrahlung. Weniger Strahlung wird reflektiert und dementsprechend ändern sich die Energieflüsse im Winter. Bei den Ergebnissen des EUR-44-Gebiets ist dieser Effekt durch das neue Bodentemperaturschema nicht so stark ausgeprägt. Ebenso sind keine signifikanten Änderungen durch die anderen „Schaltereffekte“ in den analysierten Flächenmittel nachzuweisen. Dies liegt vermutlich daran, dass sich diese zum einen eher auf die Wärmeflüsse und die Wärmeverteilung im Boden statt auf die Oberfläche auswirken und zum anderen dass, wie bereits bei der Analyse der Datensätze festgestellt, durch die räumliche Mittlung lokal auftretende Effekte ausgeglichen werden.



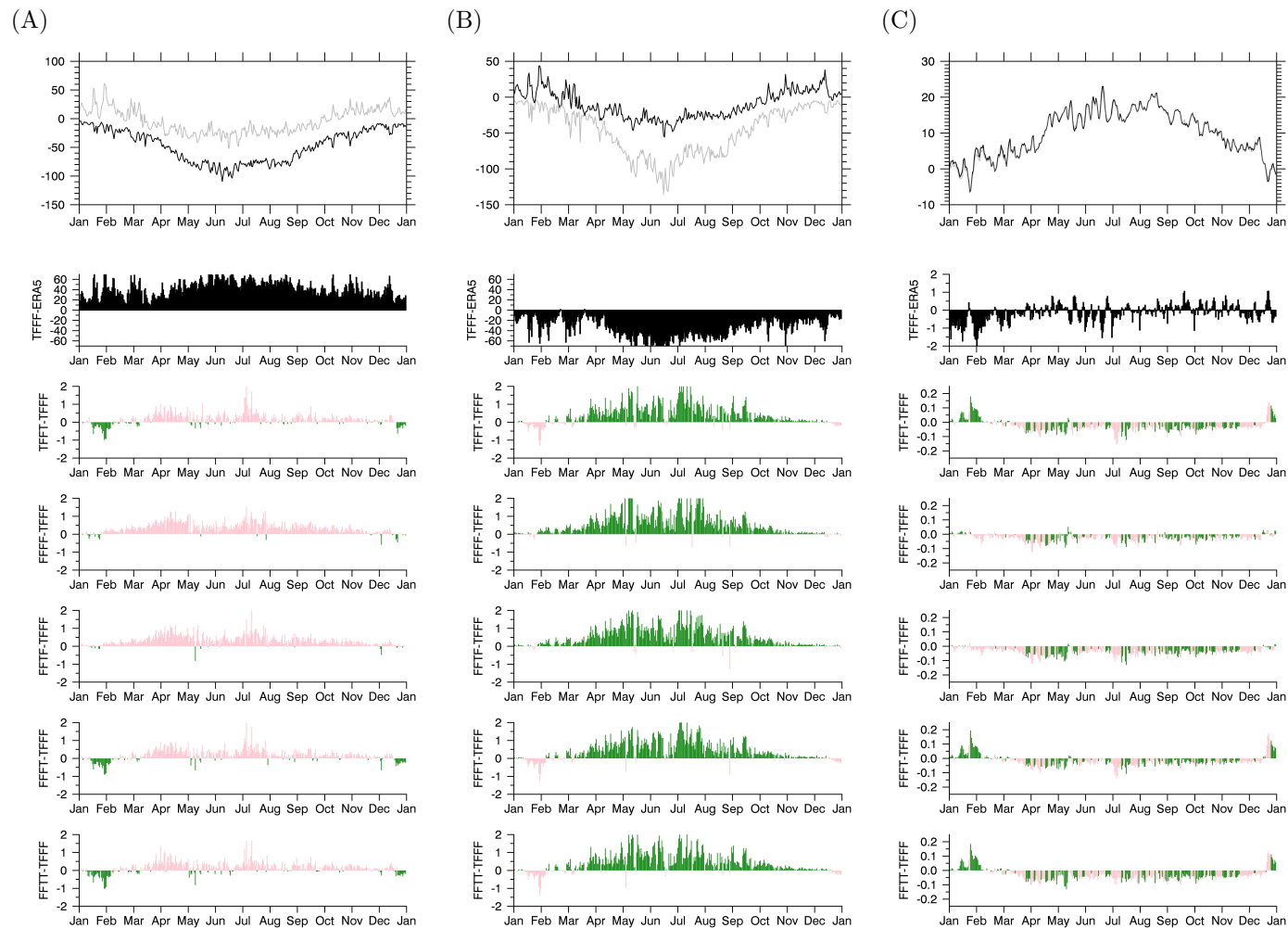


Abb. 69: Vergleich des latenten und sensiblen Wärmeflusses sowie der mittleren Temperatur der verschiedenen Modellläufe von Experimentplan 2 (vgl. Abb. 45) mit den ERA5-Validationsdaten für GER-11 (Aufbau analog zu Abb. 65 mit auf  $\pm 2 \frac{\text{W}}{\text{m}^2}$  bzw.  $\pm 0.25 \text{ }^\circ\text{C}$  begrenzter y-Achse) (A) latenter Wärmefluss (Angaben in  $\frac{\text{W}}{\text{m}^2}$ ) / (B) sensibler Wärmefluss (Angaben in  $\frac{\text{W}}{\text{m}^2}$ ) / (C) mittlere 2 m-Temperatur (Angaben in  $^\circ\text{C}$ )

### Zusammenfassung und Einschätzung der Erkenntnisse durch die Wärmeflussdifferenzanalyse

Insgesamt ist bei dieser Analyse zu beachten, dass es sich bei den zum Vergleich herangezogenen ERA5-Daten nicht um tatsächliche Messungen, sondern um Reanalysedaten handelt. Dass sich diese teilweise von Beobachtungsdaten unterscheiden, wird exemplarisch anhand der Temperaturvariablen in Kapitel 6.5.2 gezeigt. Dennoch ist ein Vergleich mit diesen Daten zur Einschätzung der Änderungen notwendig. Zudem haben mehrere Studien gezeigt, dass die ERA5-Daten die Wärmeflüsse generell gut abbilden und mit der Qualität von Satelliten gemessenen Daten zu vergleichen sind. Küsten und gebirgige Regionen, wo tendenziell die höchsten Differenzen zu den REMO-Daten festgestellt wurden, besitzen laut verschiedener Studien auch größere Unterschiede zu den jeweils zur Validation verwendeten Satellitendaten (Urraca et al., 2018; Tang et al., 2021). Besonders in topographisch komplexen Gebieten werden die Strahlungsvariablen durch Hangneigung, -ausrichtung und Schattenbildung beeinflusst, was in den Modellen durch die gemittelten Topographien nicht abgebildet wird (M. D. Müller & Scherer, 2005). Auch Martens et al. (2020) stellen beispielsweise fest, dass der latente Wärmefluss in ERA5 überschätzt wird. Das bedeutet, dass der in den vorherigen Analysen festgestellte Fehler von REMO für die latente Wärme beim Vergleich mit gemessenen Daten nicht so groß ausfallen dürfte. Martens et al. (2020) merken an, dass die Flüsse von ERA5 insgesamt so gut abgebildet werden, um trotzdem als Vergleichsdaten hilfreich zu sein.

Durch die Datensatzänderungen haben sich räumlich gesehen einige Änderungen der Wärmefluss- und Strahlungsvariablen ergeben. Allerdings zeigen die genannten Beispiele keine allgemeine Verbesserung oder Verschlechterung an. Die Differenzen sind sowohl im zeitlichen als auch im räumlichen Vergleich stark variabel. Bei der Analyse hat sich herausgestellt, dass durch die neuen Datensätze zwar einige Variablen in manchen Regionen besser, gleichzeitig andere Variablen oder andere Regionen auch schlechter modelliert werden. Eine eindeutige Beurteilung der Änderungen anhand der vorliegenden Validationswerte von ERA5 kann daher nicht vorgenommen werden. Obwohl die dargestellten Änderungen der einzelnen Modellläufe im Vergleich zu den Absolutwerten und im Vergleich zu den ERA5-Validierungsdaten sehr gering erscheinen, kann sich die veränderte Energiemenge über den Zeitraum aufsummieren und zu deutlicheren Änderungen im Atmosphäre-Boden-System führen. Wie sich beispielsweise die neuen Bodenwärmestromdichten auf die Bodentemperaturen auswirken, wird im nächsten Kapitel untersucht.

### Bodentemperaturveränderung

Dass in diesem Kapitel nur auf die Veränderungen eingegangen wird, die durch den Experimentplan 2 ausgelöst werden, liegt daran, dass diese Modellversionen alle die gleiche Bodendatengrundlage und die gleiche hydrologische und thermale Bodenparametrisierung verwenden. Unterschiede, die nachgewiesen werden können, können daher nur durch den neu eingeführten Sand- und Tonanteil statt der Texturklasse, den fünf verschiedenen statt nur einem konstanten Bodenparametrisierungswert oder der numerischen Methode der Bodentemperaturberechnung ausgelöst werden. Weitere Analysen der Bodentemperaturen für zusätzliche Modellversionen werden in Kapitel 6.5.2.2 vorgenommen (dort auch im Vergleich zu Validationsdaten), während

nachfolgend zunächst nur auf die Unterschiede zwischen den Modellläufen eingegangen wird.

### **Räumliche und zeitliche Mittelwerte der verschiedenen Bodentemperaturen im Vergleich**

In Abbildung 70A sind die jährliche Durchschnittstemperatur der obersten Bodenschicht (TD3) des Modelllaufs id03\_TFFF für EUR-44 sowie das entsprechende tägliche Flächenmittel der fünf verschiedenen Bodentemperaturen dargestellt. Im zeitlichen Mittel ist der Nord-Süd-Gradient, der nur durch höher gelegene Gebirge wie z. B. die Alpen unterbrochen wird, klar erkennbar. Die Durchschnittstemperatur der obersten Bodenschicht ist kleiner, je weiter nördlich eine Region liegt. Betrachtet man die Zeitreihen der unterschiedlichen Bodenschichten über das Jahr, fällt auf, dass alle fünf Kurven grob einer verschobenen Sinuskurve entsprechen. Die Amplitude jeder Temperaturkurve ist dabei abhängig von der Tiefe der Bodenschicht. Je tiefer die Schicht liegt, desto kleiner ist die Temperaturamplitude. Außerdem verschiebt sich das jeweilige Temperaturmaximum, welches bei Bodenschicht 1 und 2, die insgesamt sehr ähnlich sind, noch bei Ende Juli liegt, für Bodenschicht 3 auf Anfang August, für Bodenschicht 4 auf Anfang Oktober und für die fünfte Bodenschicht auf Mitte Dezember. Zudem dreht sich die Reihenfolge der Bodentemperaturen um. Im Winter ist die unterste Bodenschicht am wärmsten und die oberste Bodenschicht am kältesten. Während im Sommer die oberen Schichten viel wärmer als die unteren Bodenschichten sind.

Vergleicht man diesen Lauf id03\_TFFF mit der Modellversion id03\_FFTT, bei der alle drei neuen Schalter geändert werden (vgl. Abb. 45), räumlich und zeitlich, erhält man Abbildung 70B. Das Differenzmuster zeigt, dass im Großteil des EUR-44-Gebiets durchschnittliche Unterschiede von ca.  $\pm 0.5^\circ\text{C}$  vorliegen. Während im westlichen Teil des Untersuchungsgebiet vermehrt negative Werte erreicht werden, ist die Bodentemperatur für id03\_FFTT im Vergleich zu id03\_TFFF im östlichen Bereich zumeist höher. Auffällig sind die positiven Differenzen um das Schwarze Meer herum, welche in Teilen von Rumänien und Serbien sogar höhere Beträge von bis zu  $1^\circ\text{C}$  haben. In den nicht dargestellten Differenzkarten der anderen Bodenschichten sind die positiven und negativen Differenzen ähnlich verteilt, wenn auch der durchschnittliche Betrag der Differenzen geringer ausfällt, was ebenso die Zeitreihenanalyse beweist. Daher wird auf diese Karten nicht näher eingegangen.

Die als Zeitreihe dargestellten räumlichen Mittelwerte der Differenz zwischen Lauf id03\_FFTT und id03\_TFFF (vgl. Abb. 70B unten) zeigen, dass der Unterschied zwischen den Modellversionen nicht konstant, sondern zeitlich variabel ist. Dies kann zu einem gewissen Teil am Einfluss des Niederschlags und dem daraus resultierenden Wassergehalt im Boden liegen, der große Auswirkungen auf Bodenparameter hat. Es ist zu erkennen, dass sowohl positive als auch negative Differenzen auftreten. Die stärksten Änderungen werden in den oberen Bodenschichten nachgewiesen. Die Temperaturen der beiden unteren Bodenschichten sind dagegen annähernd gleich für beide Modellläufe. Vor allem im Mai und Juni nimmt der Unterschied zwischen den Läufen FFTT und TFFF ein Ausmaß von  $1\text{--}2^\circ\text{C}$  an, welches in der Differenzkarte nicht vorkommt, weil sich Differenzen mit unterschiedlichen Vorzeichen teilweise ausgleichen. Daher ist sowohl die Analyse der räumlichen als auch der zeitlichen Mittel wichtig, um die Auswirkungen

der vorgenommenen Änderungen besser verstehen zu können.

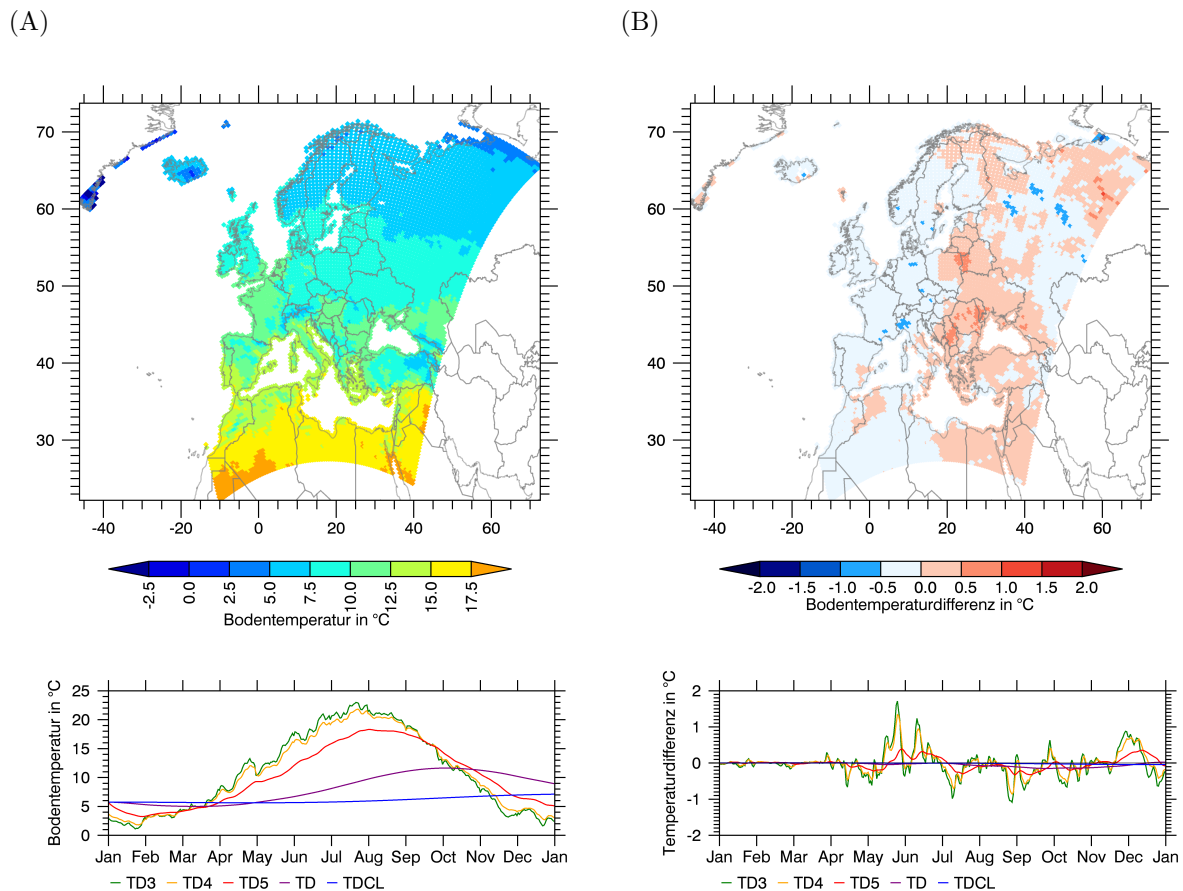


Abb. 70: Zeitliches und räumliches Mittel der Bodentemperaturen im Vergleich für EUR-44 (Datengrundlage: jeweils tägliche Werte des Jahres 2000, Modellversionen vgl. Abb. 45 und dazugehöriger Text)

(A) Absolutwerte des Laufs id03\_TFFF

(B) Differenzwert der Läufe id03\_FFTT und id03\_TFFF

oben: zeitliches Mittel der obersten Bodenschicht (TD3)

unten: jeweiliges räumliches Mittel der fünf Bodenschichten (TD3, TD4, TD5, TD, TDCL)

Analog zu Abbildung 70 zeigt Abbildung 71 die Bodentemperatur der gleichen Modellläufe (id03\_FFTT und id03\_TFFF) für GER-11. Wie bereits bei der größeren Auflösung festgestellt werden konnte, haben die höchstgelegenen Regionen wie die Alpenregion oder die noch dargestellten Ausläufer der nordwestlichen Karpaten mit 0–5 °C jährlicher Durchschnittstemperatur die geringsten Bodentemperaturwerte. Die höchsten Bodentemperaturen liegen im südöstlichen Teil Frankreichs, Norditalien und Balkan bei über 12.5 °C. Die insgesamt Werteverteilung von GER-11 stimmt sehr gut mit dem Muster der größeren Auflösung (Abb. 70A) überein. Auffällig am Differenzmuster (Abb. 71B) ist die leichte räumliche Abhängigkeit der Differenz zu den Absolutwerten. Die angesprochenen höheren Durchschnittstemperaturen in der Po-Ebene Italiens sowie in einigen Teilen des südöstlichen Untersuchungsgebiets, die beim Modelllauf id03\_TFFF zwischen 12.5–15.0 °C erreichen, werden bei der Modellversion mit der Einstellung FFTT um ca. 0.1–0.2 °C verringert. Dieses Phänomen konnte bei der EUR-44-Analyse nicht beobach-

tet werden. Im restlichen GER-11-Gebiet ist der Betrag der Differenz noch kleiner und am Differenzmuster sind keine weiteren Merkmale abzuleiten.

An den Zeitreihen der Bodentemperaturen des TFFF-Laufs (Abb. 71A unten) lassen sich die gleichen Merkmale wie beim Mittel des EUR-Gebiets mit größerer Auflösung feststellen. Sowohl die Höhe der Temperaturen als auch der Verlauf stimmen trotz unterschiedlicher Bezugsgebiete, über die gemittelt wird, sehr gut überein. Lediglich die Variabilität der täglichen Temperaturen in den obersten drei Bodenschichten ist beim GER-11-Lauf deutlich höher als bei EUR-44. Dadurch wird die charakteristische Kurve mit einem Maximum in der Mitte des Jahres, die bei EUR-44 in den oberen Schichten gut abgebildet ist, bei GER-11 schlechter dargestellt. Die Zeitreihen der Differenzen zwischen den Läufen id03\_FFTT und id03\_TFFF (Abb. 71B unten) hingegen unterscheiden sich deutlich von denen von EUR-44. Die maximalen Beträge, die bei EUR-44 noch bei ca.  $\pm 2^\circ\text{C}$  lagen, minimieren sich bei GER-11 auf ca.  $\pm 0.2^\circ\text{C}$ . Außerdem ist anders als bei den EUR-44-Zeitserien zu erkennen, dass positive Differenzen vor allem in den kalten Monaten und negative Differenzen vor allem im Sommer auftreten. Zudem stimmen die Kurvenverläufe der oberen drei Bodenschichten sehr gut mit der in Abbildung 69C dargestellten Abweichung der mittleren 2 m-Temperatur beider Läufe überein, was den großen Zusammenhang der Bodentemperatur mit der Lufttemperatur bestätigt. Die Temperatur der beiden unteren Bodenschichten hingegen weist aufgrund der sich andeutenden Sinusschwingung vor allem bei der vierten Schicht (TD) auf eine Verschiebung der Extremwerte der tatsächlichen Werte hin. Im Vergleich mit den tatsächlichen Werten sieht man, dass höhere Bodentemperaturen in Lauf TFFF von FFTT kälter modelliert werden und umgekehrt niedrigere Temperaturen positive Differenzen haben, was jedoch aufgrund des geringen Betrags der Differenz kaum Auswirkungen hat. Aus dieser Flächenmittelanalyse kann dennoch abgeleitet werden, dass der Modelllauf id03\_FFTT für das GER-11-Gebiet zu einer Verringerung der jährlichen Bodentemperaturamplitude in den einzelnen Schichten führt.

Um die Auswirkungen der einzelnen Datensatz- und Parametrisierungsänderungen noch besser verstehen zu können und mögliche Effekte, die durch die Bildung von Flächenmitteln entstehen, zu reduzieren, wird bei der nächsten Analyse auf einen konkreten Beispielpixel (Würzburg) eingegangen. Dieser ist in Abbildung 71B mit einem Rechteck umrandet, liegt in der Mitte des Untersuchungsgebiets und kennzeichnet sich durch mittlere Absolut- und Differenzwerte.

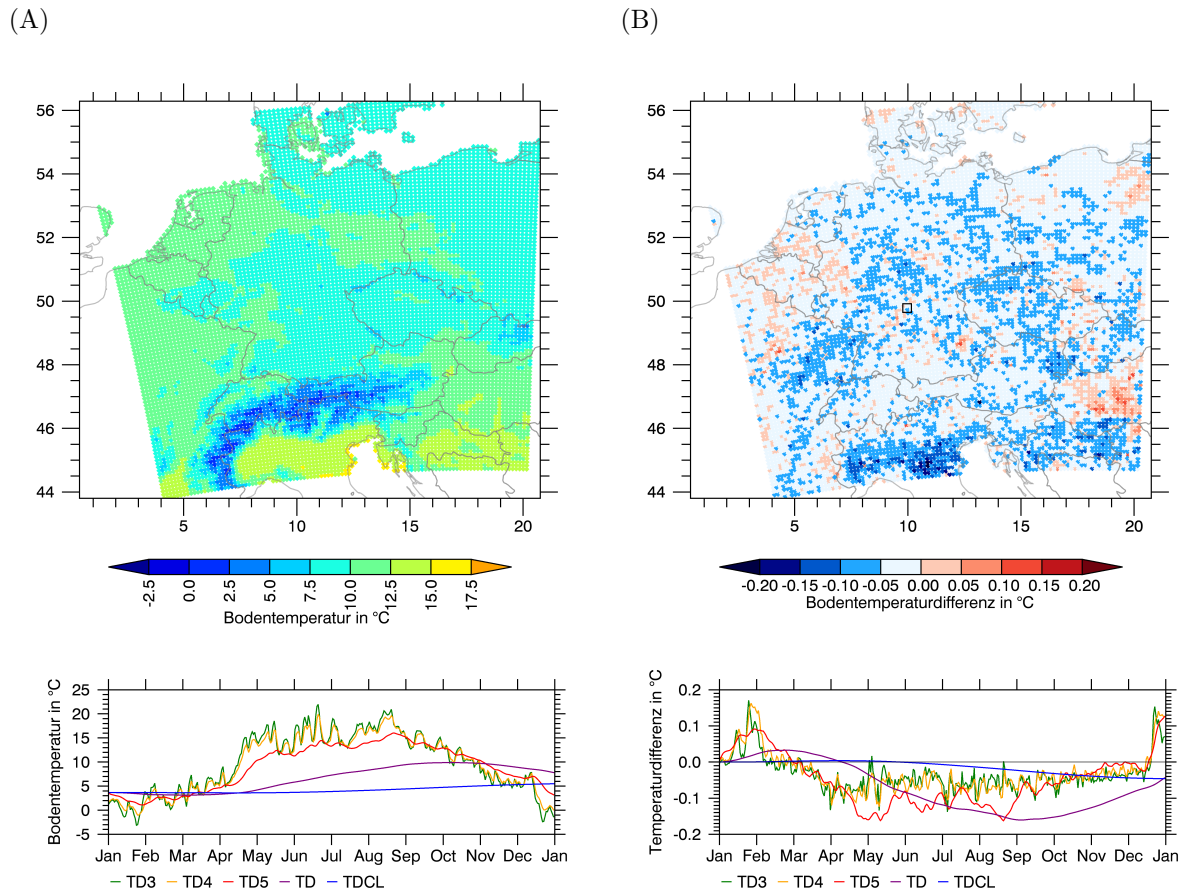


Abb. 71: Zeitliches und räumliches Mittel der Bodentemperaturen im Vergleich für GER-11 (Datengrundlage: jeweils tägliche Werte des Jahres 2000, Modellversionen vgl. Abb. 45 und dazugehöriger Text)  
 (A) Absolutwerte des Laufs id03\_TFFF  
 (B) Differenzwert der Läufe id03\_FFTT und id03\_TFFF  
 oben: zeitliches Mittel der obersten Bodenschicht (TD3)  
 unten: jeweiliges räumliches Mittel der fünf Bodenschichten (TD3, TD4, TD5, TD, TDCL)  
 (Das eingezeichnete schwarze Rechteck in der Mitte des Untersuchungsgebiets kennzeichnet die Lage des für andere Analysen verwendeten Pixels „Würzburg“.)

### Vergleich der jährlichen Bodentemperaturprofile des Experimentplans 2 für einen ausgewählten Standort

In Abbildung 72 sind die modellierten Bodentemperaturen aller fünf Schichten des Laufs id03\_TFFF und die jeweilige Differenz der fünf Testläufe des Experimentplans 2 dargestellt. Bei den Werten handelt es sich um die Monatsmittelwerte der Kaltstartläufe vom Jahr 2000 für den Beispielpixel Würzburg, der aus dem GER-11-Lauf ausgeschnitten wurde (vgl. schwarzes Rechteck in Abb. 71B). Wie an den Zeitreihen der Bodentemperaturen (vgl. Abb. 71 unten) bereits festgestellt wurde, sind zum einen die Jahressgänge der einzelnen Bodenschichten zeitlich zueinander verschoben und zum anderen deren Amplituden geringer, je tiefer die Schicht liegt.

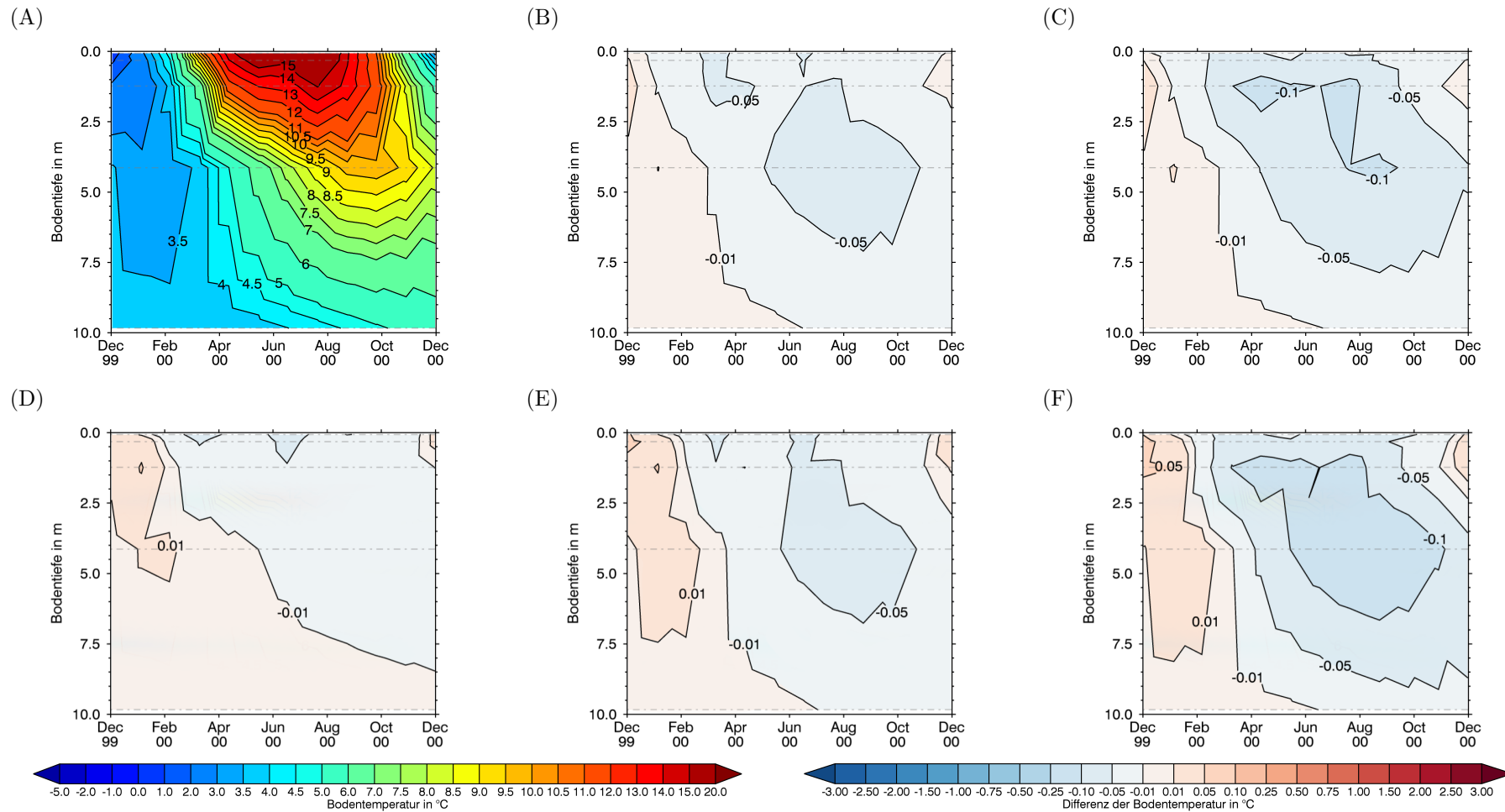


Abb. 72: Bodentemperaturprofile der Modellläufe des Experimentplan 2 (Parametrisierungsunterschiede vgl. Abb. 45) im Vergleich für die Monatsmittelwerte des Jahres 2000 für den Beispielpixel „Würzburg“ (vgl. schwarzes Rechteck in Abb. 71B)

Alle Differenzabbildungen (Abb. 72B-72F) weisen ein ähnliches Muster auf. Auffällig ist dabei, dass dieses Differenzmuster dem Muster der tatsächlichen Werte (vgl. Abb. 72A) folgt. Wärmere Bodentemperaturen werden in allen neuen Läufen etwas kühler, die kälteren Temperaturen etwas wärmer. Zwischen den Ergebnissen der einzelnen Schalter lässt sich kein großer Unterschied erkennen. Nur das Differenzmuster von Lauf TFFT (vgl. Abb. 72D), der wie der Originallauf noch die Bodentextur statt dem Sand- und Tonanteil verwendet, zeigt im Vergleich zu den anderen Mustern sowohl betragsmäßig kleinere Differenzen als auch einen „flacheren Winkel“ bei der negativen Temperaturdifferenz. Das Verwenden von fünf verschiedenen Bodenparametern und dem neuen Bodentemperaturschema, welches als Kombination von Abbildung 72C und 72E angesehen werden kann, ist in Abbildung 72F zu sehen. Diese Änderungen führen zu den höchsten negativen Temperaturdifferenzen von  $0.1\text{ °C}$  im Sommer in der vierten Bodenschicht und im Herbst in der oberen Hälfte der fünften Schicht. Insgesamt wirken sich alle Änderungen dieses Experimentplans nicht so stark auf das Modellergebnis der Bodentemperaturen aus. Die hier erläuterten Ergebnisse dieses Beispielpixels können zudem auf das Mittel des gesamten GER-11-Untersuchungsgebiets übertragen werden. Bei der Analyse der Bodentemperaturen des EUR-44-Gebietes (nicht graphisch dargestellt) zeigt das Differenzmuster je zwei positive und zwei negative Temperaturanomalien. Ferner sind besonders die positiven Anomalien mit Beträgen von durchschnittlich bis zu  $0.75\text{ °C}$  wärmeren Bodentemperaturen im Mai in der obersten Bodenschicht deutlich ausgeprägter als beim Deutschlandlauf. Die vorgenommenen Änderungen haben daher durchaus einen signifikanten Effekt auf die modellierten Bodentemperaturen. Da es sich bei diesen Modelldaten aber um Kaltstartläufe über ein Jahr handelt, was zu besonders großen Differenzen in den unteren beiden Bodenschichten führen würde (vgl. Abb. 8), kann der Vergleich mit den ERA5-Daten für die Modellversionen dieses Experimentplans keine aussagekräftigen Ergebnisse über die Richtung der Effekte liefern. Eine Beurteilung der Änderung der Bodentemperaturen findet daher mit zwei Warmstartläufen und den ERA5-Validationsdaten in Kapitel 6.5.2.2 statt.

### 6.5.2 Vergleich der Validationsdatensätze untereinander und mit den Modelldaten

Um die Güte eines Modells zu prüfen, wird das Modell mit Beobachtungsdaten validiert. Die verwendeten Beobachtungsdaten (CRU, E-OBS, ERA5, DWD) wurden in Kapitel 4.5 vorgestellt. Generell ist zu erwähnen, dass sich diese ähneln, aber dennoch Unterschiede vor allem bedingt durch die notwendige Aufbereitung wegen unterschiedlicher räumlicher Auflösung existieren. Daher folgt zunächst ein grundlegender Vergleich der verwendeten klimatischen Beobachtungsdaten untereinander. In Abbildung 73 sind die Kerndichteschätzfunktionen (vgl. Kapitel 4.2) der vier Validationsdatensätze für die vier klimatologischen Variablen T2M, T2MAX, T2MIN und RR dargestellt, welche für die nachfolgenden Vergleiche verwendet werden. Zudem werden die vier Jahreszeiten unterschieden. Die Abbildung bezieht sich auf die bereits auf das  $0.44^\circ$  interpolierten Koordinatengitterwerte des EUR-44-Gebietes. Die Werte entsprechen den jeweiligen Monatsmitteln (für die Temperaturen in  $^\circ\text{C}$ ) bzw. -summen (für den Niederschlag in mm) von 2000-2018.



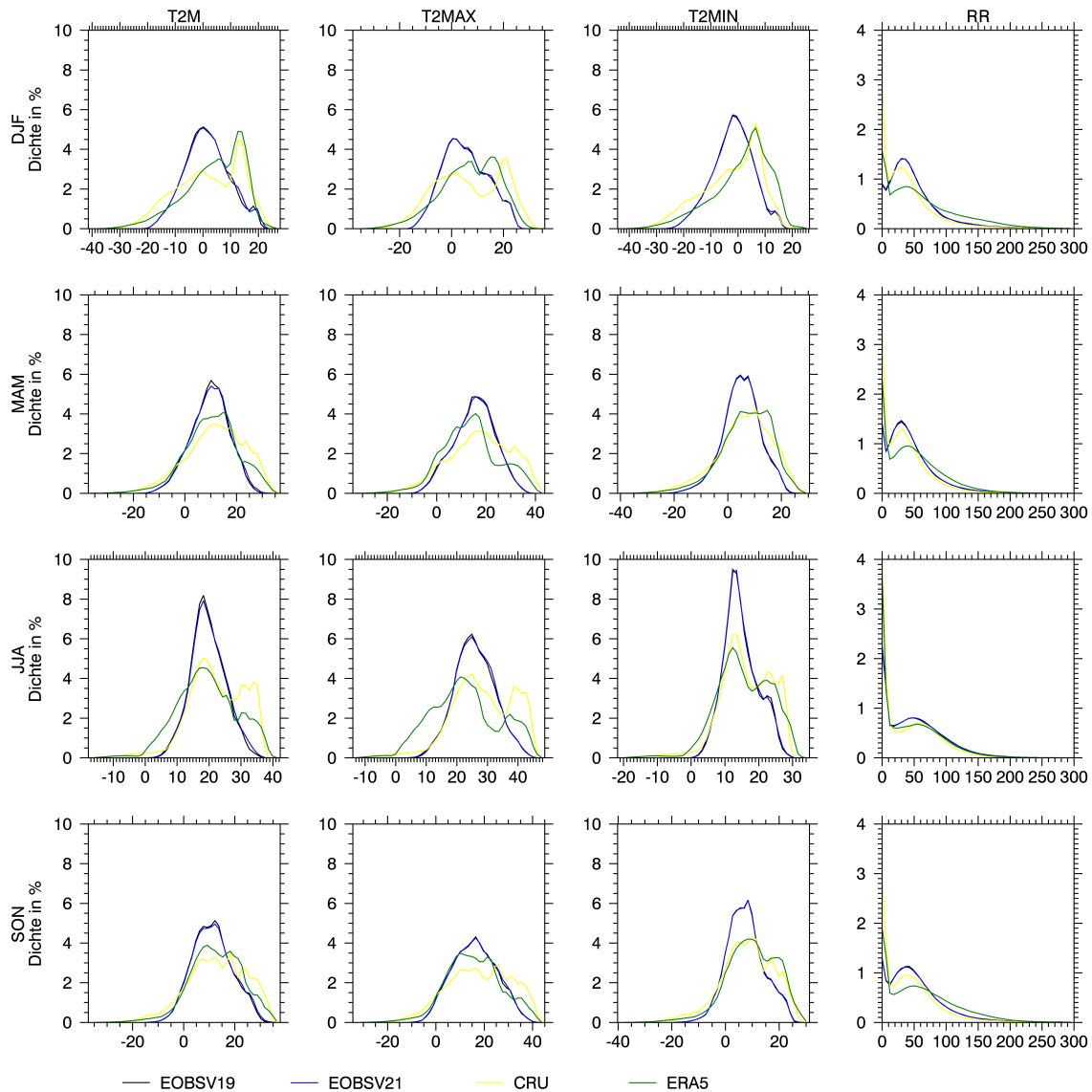


Abb. 73: Kernfunktionsdichteschätzung der klimatologischen Validationsdatensätze pro Jahreszeit für EUR-44

Die Datengrundlage bilden die Monatsmittel (T2M, T2MAX und T2MIN) bzw. -summen (RR) aller Landpunkte des EUR-44-Gebiets von 2000-2018. Die jeweils an die Jahreszeit und Variable angepassten Werte der x-Achse sind für die Temperaturen (1.-3. Spalte) in  $^{\circ}\text{C}$  und für den Niederschlag (4. Spalte) in mm angegeben.

Zunächst lässt sich an den Dichteschätzfunktionen feststellen, dass die beiden E-OBS-Versionen (V19 und V21) fast identisch sind. Kleine Unterschiede erkennt man nur bei T2M in den Maxima der Funktionen bei MAM, JJA und SON. Beim Vergleich mit den anderen Datensätzen CRU und ERA5 fällt auf, dass bei allen drei Temperaturvariablen die E-OBS-Daten eher einer unimodalen Verteilung und die beiden anderen eher einer bimodalen Verteilung mit einem Hauptmaximum und einem meist wesentlich niedrigeren Nebenmaximum folgen. Besonders deutlich ist dies in allen drei Temperaturvariablen im Sommer zu erkennen. Während E-OBS ein klares Hauptmaximum bei ca.  $20^{\circ}\text{C}$  (T2M),  $25^{\circ}\text{C}$  (T2MAX) und  $14^{\circ}\text{C}$  (T2MIN) erreicht, wird sowohl bei CRU als auch bei ERA5 ein weiterer Extrempunkt in den einzelnen Funktionen bei jeweils ca.  $15^{\circ}\text{C}$  höheren Temperaturen erreicht. Ferner erreichen sowohl CRU als auch ERA5 jeweils eine größere

Variationsbreite in allen Temperaturvariablen und allen Jahreszeiten. Besonders im Winter erkennt man beispielsweise eine starke Verschiebung des Maximums von CRU und ERA5 im Vergleich zu beiden E-OBS-Daten um ca. 10–15 °C zu höheren Temperaturen. In den anderen Jahreszeiten stimmen die Maxima der Funktionen relativ gut überein. Auch zwischen CRU und ERA5 gibt es Unterschiede in der jeweiligen Verteilung. Während die Verläufe der Funktionen relativ gut übereinstimmen (außer T2MAX in DJF und JJA), variiert die Höhe jedoch je nach Variable und Jahreszeit stärker. Im Gegensatz zu den Temperaturvariablen unterscheiden sich die Funktionsverläufe beim Niederschlag zwischen den verschiedenen Jahreszeiten und zwischen den verschiedenen Datensätzen kaum. Die Niederschlagsverteilung erreicht bei allen Teilabbildungen und Daten ihr Maximum beim Wert 0. Anschließend nehmen die Häufigkeiten bis zu einem gewissen Niederschlagswert ab, um dann bis zu einem zweiten Maximum bei ca. 30–50 mm (je nach Jahreszeit) noch einmal anzusteigen. Die Funktionen der jeweiligen Validationsdaten stimmen besonders für JJA sehr gut überein. Bei DJF, MAM und SON ist jedoch zu erkennen, dass ERA5 höhere Dichtewerte für höhere Niederschlagswerte (>100 mm) hat.

Die gleiche Analyse wurde auch für die Daten gemacht, welche für die GER-11-Modellläufe zur Validation benutzt werden (vgl. Anhang K). Die Unterschiede zwischen den Validationsdatensätzen sind hier sehr viel kleiner. Dies liegt vor allem am Deutschlandgebiet und nicht an der Auflösung bzw. Interpolation der Daten, denn eine zusätzliche nicht dargestellte Analyse der Daten des Deutschlandgebietes, welches aus dem EUR-44-Gebiet ausgeschnitten wurde (GER-44), zeigt annähernd die gleichen Kerndichteschätzfunktionen wie die Ergebnisse von GER-11. Insgesamt lässt sich an den Werten des GER-11-Gebietes ebenfalls feststellen, dass beide E-OBS-Datensatzversionen sehr ähnlich sind und auch die ERA5-Daten für dieses Gebiet die gleichen Funktionsverläufe wie diese aufweisen. Wie bereits beim EUR-44-Gebiet zeigt ERA5 etwas mehr Niederschlag. Auffällig an der GER-11 Dichtefunktion für RR ist jedoch der sehr viel kleinere Wert für 0 mm Niederschlag in allen Jahreszeiten. Ein weiteres markantes Merkmal ist bei den Kurven aller Temperaturen in den Übergangsjahreszeiten MAM und SON zu erkennen. Hier liegt jeweils eine bimodale Verteilung vor, während bei den anderen beiden Jahreszeiten jeweils eine unimodale Verteilung erkennbar ist. Dieser Effekt tritt nicht im EUR-44-Gebiet auf. Es ist an der Abbildung in Anhang K ebenfalls sichtbar, dass die gesamte T2MIN-Verteilung von ERA5 in allen Jahreszeiten etwas in Richtung höherer Temperaturwerte verschoben ist im Vergleich zu den E-OBS-Daten. Bei T2MAX hingegen liegt sie tendenziell etwas weiter in Richtung niedrigerer Temperaturen. Besonders im unteren bis mittleren Wertebereich in MAM und JJA unterscheiden sich die Kerndichteschätzfunktionen von T2MAX stärker als in allen anderen Variablen und Jahreszeiten.

Generell zeigt diese Analyse, dass ausgewählten Validationsdatensätze zwar generell relativ gleiche Wertebereiche und Verteilungen in beiden Untersuchungsgebieten besitzen, dass aber auch je nach Jahreszeit und Variable Unterschiede auftreten. Einzig die beiden E-OBS-Versionen sind nahezu identisch. Zudem muss beachtet werden, dass es sich hier nur um den Vergleich der Kerndichteschätzfunktionen handelt. Räumliche und zeitliche Muster werden bei dieser Analyse nicht berücksichtigt. Daher werden in den nachfolgenden Vergleichen mit den Modelldaten meist mehrere Validationsdatensätze herangezogen.

### 6.5.2.1 Analyseergebnisse der kurzen Modellläufe der drei Experimentpläne

Anhand der vier Variablen T2MIN, T2M, T2MAX und RR, die in allen Beobachtungsdaten vorliegen, werden nun die Modelldaten der verschiedenen Experimentpläne mit den Beobachtungsdaten verglichen. Bei diesen kurzen Läufen handelt es sich um Kaltstartläufe, was für den Vergleich mit den Beobachtungsdaten berücksichtigt werden muss. Dennoch lässt sich anhand der folgenden Abbildung eine Einschätzung der Auswirkungen der verschiedenen Änderungen vor allem im Vergleich zu der jeweils ebenfalls dargestellten bisherigen Modellversion id01 (Kaltstartversion) durchführen.

#### Quartile des RMSE und Korrelationskoeffizienten

Abbildung 74 zeigt die Ergebnisse des Experimentplan 1. Die Abbildung ist so aufgebaut, dass sowohl die verschiedenen Jahreszeiten als auch die vier verschiedenen Variablen separat für die drei verfügbaren Beobachtungsdaten und die verschiedenen Modellläufe aufgetragen sind. Zudem sind einerseits der RMSE und andererseits der Korrelationskoeffizient abzulesen. Des Weiteren ist nicht der Mittelwert, sondern der Median als Mittelpunkt und das entsprechende 1. und 3. Quartil aller Werte als „Whisker“ dargestellt. Als Datengrundlage werden die Tagesdaten der jeweiligen Datensätze verwendet. Die Abbildung ist außerdem so erstellt, dass die jeweilige x-Achse, auf der der RMSE aufgetragen ist, den Wertebereich über alle Zeitabschnitte erfasst. Das heißt, als Minimumwert wird für jede Variable der minimale Wert aller 1. Quartile und als Maximalwert der maximale Wert aller 3. Quartile verwendet. Analog dazu wird für die y-Achse, auf der der Korrelationskoeffizient angezeigt wird, jeweils das Minimum bzw. Maximum aller 1. bzw. 3. Quartile pro Zeiteinheit benutzt. Dieses Vorgehen dient wegen des möglichst kleinen Wertebereichs einer besseren Vergleichsmöglichkeit der verschiedenen Daten. Je besser die Modelldaten mit den jeweiligen Validationsdaten übereinstimmen, desto weiter oben und links liegen die Punkte und desto kürzer sind die jeweiligen Whisker.

Es fällt auf, dass die resultierenden Werte von RMSE und  $r$  stark von den verwendeten Beobachtungsdatensatz abhängen. Zwar sind die beiden E-OBS-Versionen nahezu identisch, ERA5 hingegen erreicht in allen Jahreszeiten und Variablen signifikant andere Werte, was bereits bei der Verteilungsanalyse der Validationsdaten (vgl. Abb. 73) gezeigt werden konnte. Im Gegensatz dazu ist zwischen den Modellläufen kaum ein Unterschied zu erkennen. Dadurch bilden sich in Abbildung 74 jeweils zwei Cluster. Alle ERA5-Clusterwerte haben in allen vier Variablen einen höheren Korrelationskoeffizienten und einen niedrigeren RMSE-Wert. Diese Werte stimmen also mehr mit den Modelldaten überein als die der E-OBS-Cluster.

Betrachtet man die jeweiligen IQR der einzelnen Teilabbildungen, stellt man fest, dass besonders der RMSE-IQR von T2M im Winter bei allen drei Validationsdatensätzen sowohl im Vergleich zu den anderen Jahreszeiten als auch im Vergleich zu den jeweils anderen Variablen sehr groß ist, wenngleich beachtet werden muss, dass T2M insgesamt die kleinsten RMSE-Werte und den kleinsten Wertebereich hat (0.79–2.22 °C im Vergleich zu 2.51–6.45 °C bei T2MAX und 1.85–5.62 °C bei T2MIN). Es fällt auf, dass T2MIN beim RMSE besser abschneidet, während T2MAX höhere

Korrelationskoeffizienten erreicht. Daraus lässt sich schließen, dass die modellierte T2MAX den Verlauf der Beobachtungsdaten besser abbilden kann, jedoch einen höheren systematischen Fehler hat als T2MIN. Wie bereits in den vorherigen Analysen festgestellt, erreicht der Niederschlag in allen Zeitabschnitten die geringsten r-Werte.

Unter Zuhilfenahme der in Anhang L dargestellten Medianwerte von EOBSV21, aus denen Abbildung 74 zum Teil erstellt wurde, ist zu erkennen, dass bei der Variable T2MIN die Modellläufe id01\_AST, id01\_ASF, id01\_GSF, id01\_GST und id01\_GHF sowohl beim RMSE als auch bei r etwas besser abschneiden als die restlichen Läufe, während AFT, GFF und GHT jeweils höhere RMSE- und niedrigere r-Werte erreichen. Bei T2M im DJF lässt sich ebenfalls ein kleiner Unterschied zwischen den Modellläufen erkennen. Hier unterliegen die Läufe id01\_GFT und id01\_GHT den anderen Versionen etwas deutlicher. Insgesamt schneiden alle Modellläufe, die mit Soilgrid berechnet werden, beim Korrelationskoeffizienten für T2M über den Gesamtzeitraum am besten ab. Anhand der Werte und Farbmuster der Tabelle im Anhang ist allerdings klar sichtbar, dass die Werte der jeweiligen Modellläufe zwischen den verschiedenen Zeiträumen und vor allem zwischen den verschiedenen Variablen größeren Schwankungen unterliegt als zwischen den Modellversionen und dass sich keine Modellversion überall durch gute Werte durchsetzen kann.

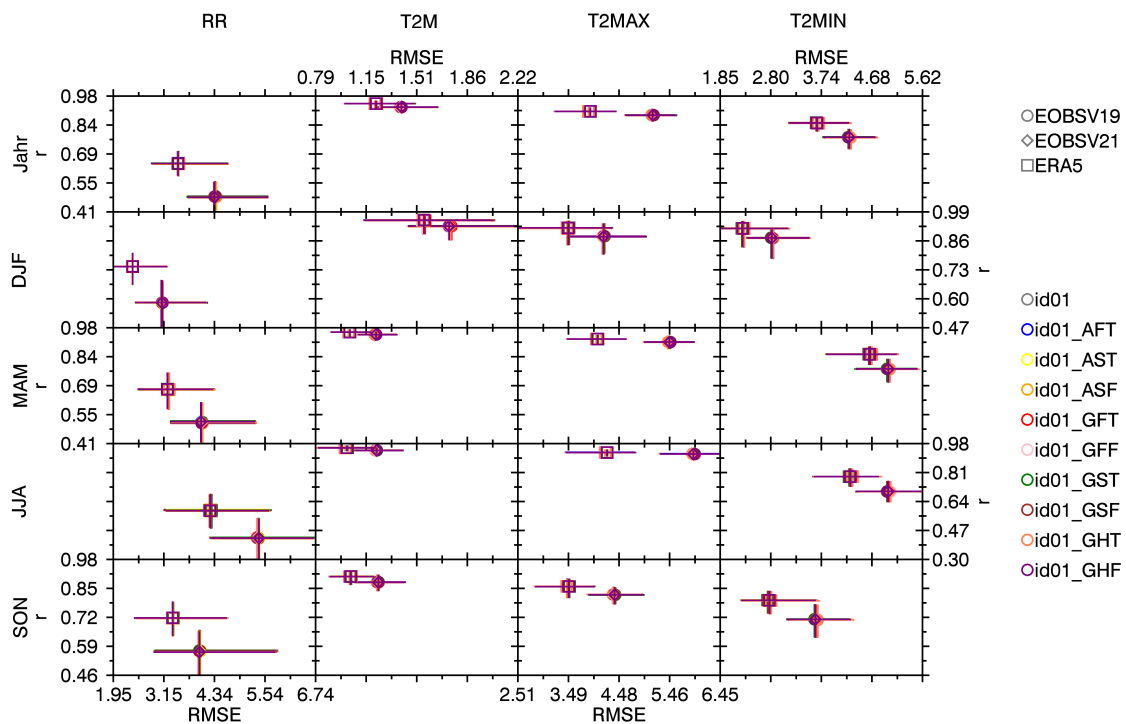


Abb. 74: Übersicht des 1., 2. und 3. Quartils des jeweiligen RMSE und Korrelationskoeffizienten der Modellläufe des ersten Experimentplans (vgl. Abb. 44 und dazugehörigen Text für Erklärung der Modellversionen) im Vergleich zu verschiedenen Validationsdaten für RR, T2M, T2MAX und T2MIN pro Jahr und Jahreszeit

Datengrundlage sind die entsprechenden Tagesdaten der Läufe des GER-11-Gebiets. Auf der x-Achse sind jeweils die RMSE-Werte aufgetragen (mit der Einheit mm für RR und °C für T2M, T2MAX und T2MIN). Die y-Achse stellt den Korrelationskoeffizienten r dar. Zu beachten ist, dass sich der Wertebereich der x- und y-Achse jeweils für die Variablen bzw. die Jahreszeiten ändert, um einen besseren Vergleich der Modellläufe zu ermöglichen.

Erstellt man die gleiche Abbildung mit den Läufen des zweiten Experimentplans und nur im Vergleich zu EOBSV21, erhält man Abbildung 75. Auch hier bestätigen sich die verschiedenen Qualitätskriterien der vier verschiedenen Variablen mit den fast exakt gleichen Wertebereichen und den geringen Auswirkungen der verschiedenen Modellläufe auf den RMSE-Wert und den Korrelationskoeffizienten.

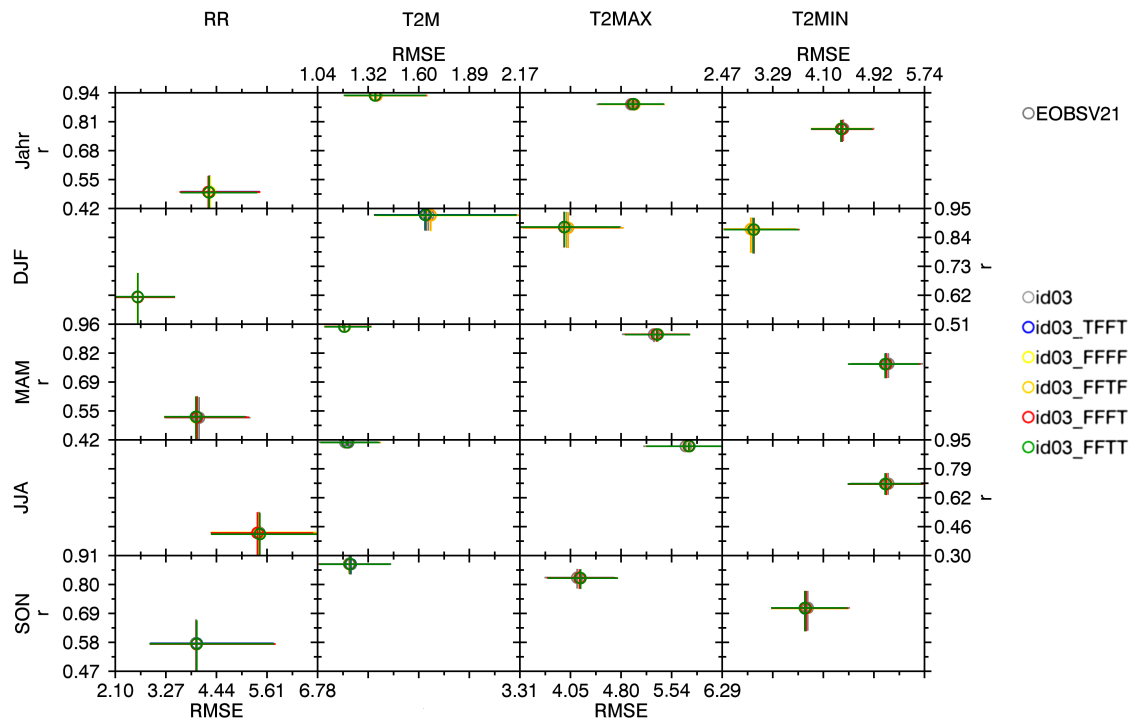


Abb. 75: Übersicht des 1., 2. und 3. Quartils des jeweiligen RMSE und Korrelationskoeffizienten der Modellläufe des zweiten Experimentplans (vgl. Abb. 45 und dazugehörigen Text für Erklärung der Modellversionen) im Vergleich zu den EOBSV21-Daten für RR, T2M, T2MAX und T2MIN pro Jahr und Jahreszeit

Datengrundlage sind die entsprechenden Tagesdaten der Läufe des GER-11-Gebiets. Auf der x-Achse sind jeweils die RMSE-Werte aufgetragen (mit der Einheit mm für RR und °C für T2M, T2MAX und T2MIN). Die y-Achse stellt den Korrelationskoeffizienten  $r$  dar. Zu beachten ist, dass sich der Wertebereich der x- und y-Achse jeweils für die Variablen bzw. die Jahreszeiten ändert, um einen besseren Vergleich der Modellläufe zu ermöglichen.

Unterschiede zwischen den verschiedenen Modellläufen lassen sich in der Abbildung beispielsweise beim Niederschlag im Sommer erkennen. Hier erreicht der Lauf mit der Kombination FFFT einen kleineren RMSE als die anderen Läufe. Bei allen Temperaturvariablen kann man die größten Unterschiede im DJF feststellen. Der Lauf FFFF liefert bei T2M und T2MAX schlechtere, bei T2MIN jedoch bessere Ergebnisse. Lauf FFTT ist bei T2MIN immer (bis auf DJF) minimal besser als die restlichen Läufe. Wie bei der Analyse der Datensatzänderungen (vgl. Abb. 74) sind die resultierenden Änderungen der Schalterkombinationen für die vier dargestellten Variablen sehr klein.

Abbildung 76, welche die Ergebnisse von Experimentplan 3 ebenfalls nur im Vergleich zu EOBSV21 enthält, zeigt jedoch, dass größere Änderungen bei dieser Analyse- und Darstellungsart möglich sind. Zwar ergeben die vorgenommenen Modifikationen mit Ausnahme des id11-Laufs auch in

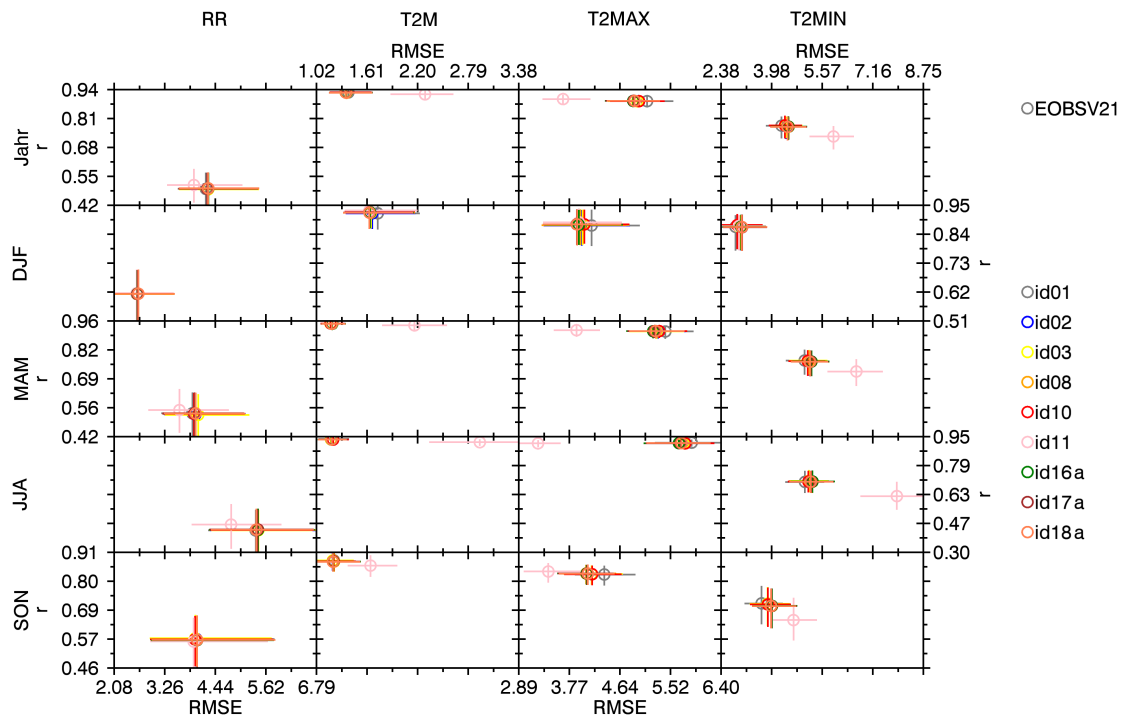


Abb. 76: Übersicht des 1., 2. und 3. Quartils des jeweiligen RMSE und Korrelationskoeffizienten einiger Modellläufe des dritten Experimentplans (vgl. Abb. 46 und dazugehörigen Text für Erklärung der Modellversionen) im Vergleich zu den EOBVS21-Daten für RR, T2M, T2MAX und T2MIN pro Jahr und Jahreszeit

Datengrundlage sind die entsprechenden Tagesdaten der Läufe des GER-11-Gebiets. Auf der x-Achse sind jeweils die RMSE-Werte aufgetragen (mit der Einheit mm für RR und °C für T2M, T2MAX und T2MIN). Die y-Achse stellt den Korrelationskoeffizienten  $r$  dar. Zu beachten ist, dass sich der Wertebereich der x- und y-Achse jeweils für die Variablen bzw. die Jahreszeiten ändert, um einen besseren Vergleich der Modellläufe zu ermöglichen.

diesem Experimentplan kaum signifikante Änderung in den Modelldaten, im id11-Lauf bewirkt die Verwendung des 5Layer-Schemas jedoch deutliche Veränderungen sowohl im RMSE-Wert als auch im Korrelationskoeffizient bei allen Variablen. Auffällig ist, dass sich dieser Lauf aber in allen Variablen im DJF nicht so stark von den anderen Läufen abhebt. Der Niederschlag im Frühjahr und Sommer sowie die T2MAX in allen anderen Jahreszeiten werden durch diesen Lauf mit dem 5Layer-Schema besser abgebildet. Die Variablen T2M und T2MIN werden dagegen besonders im Sommer deutlich schlechter modelliert und die IQR des RMSE wird ebenfalls deutlich größer.

Die anderen dargestellten Modellläufe unterscheiden sich nicht so stark voneinander. Allerdings verbessern sich die T2M und die T2MAX aller Läufe in allen Jahreszeiten vor allem hinsichtlich des RMSE-Wertes im Vergleich zum bisherigen Lauf id01 etwas. Die T2MIN wird jedoch schlechter abgebildet. Bei dieser Variable ist bei genauem Hinsehen auch zu erkennen, dass die Läufe bis einschließlich id10 jeweils näher am bisherigen Lauf liegen als die Läufe id16a, id17a und id18a.

Die gleichen Analysen auf dem EUR-44-Gebiet, welche hier nicht separat dargestellt sind, zeigen insgesamt größere Wertebereiche sowohl bei den RMSE-Werten als auch beim Korrelationsko-

effizienten, bei dem für T2MIN teilweise sogar leicht negative Werte für RR erreicht werden. Aufgrund der breiteren Streuung liegen die beiden Validationsdatencluster (EOBS und ERA5) dichter beieinander und, anders als bei den dargestellten Ergebnissen von GER-11, überlagern sich die verschiedenen Median- und IQR-Werte. Daher sind die größeren Unterschiede zwischen den einzelnen Modellläufen trotzdem schlechter zu erkennen. Die Aussagen über die jeweilige Lage der Qualitätskriterien und die unterschiedlichen IQR, welche bei den GER-11-Läufen getroffen wurden, bestätigen sich jedoch.

### **Ergebnisse des Vergleichs der Overall-Accuracy-Werte**

Neben dem Korrelationskoeffizienten und dem RMSE sind noch weitere Indizes zum Vergleichen der Modelldaten untereinander und im Vergleich zu Beobachtungsdaten berechnet worden (vgl. Kapitel 6.1.3). In Tabelle 16 sind die durchschnittlichen OA-Mittelwerte der Variablen (T2M, T2MAX, T2MIN und RR) für die Läufe des Experimentplan 1 (vgl. Abb. 44) im Vergleich zu den jeweiligen Beobachtungsdatensätzen für EUR-44 und GER-11 dargestellt. Grundsätzlich fällt auf, dass die spaltenweise Einfärbung, welche das Ranking der einzelnen Modellläufe darstellt, bei allen Validationsdatensätzen sehr ähnlich ist. Der Wert des Laufs mit dunkelgrüner Einfärbung schneidet jeweils im Vergleich zu den jeweiligen Validationsdaten am besten und der Lauf mit dunkelroter Einfärbung am schlechtesten ab. Die anderen Modellläufe werden abhängig von diesen beiden Extremwerten mit verschiedenen Farben hinterlegt und sind daher nicht für alle Validationsdatensätze gleich. Anhand der eingetragenen Werte ist auch hier sehr deutlich, dass, wie bei den vorherigen Analysen bereits festgestellt, die Modellläufe untereinander weniger Variabilität aufweisen als die Werte, die durch die jeweiligen Validationsdaten erzeugt werden. Die höchsten OA-Werte und damit die höchste Übereinstimmung liefert ERAI bei den EUR-44-Daten, was aufgrund des Modellantriebs mit ERA-Interim-Daten nicht überrascht. Danach folgen die ERA5-Werte, die relativ ähnliche Werte wie ERAI erreichen und anschließend folgen die Werte, die mit den beiden EOBS-Version erzeugt werden, wobei der Wert für EOBSV21 für alle Läufe höher ist als der für EOBSV19. Dies gilt auch für die auf zeitlicher Basis erzeugten Werte des GER-11-Gebiets (Tab. 16c). Bei der räumlichen Variante wurden höhere Übereinstimmungen für EOBSV19 mit den GER 11-Modelldaten berechnet.

Vergleicht man das Abschneiden der einzelnen Modellläufe des Experimentplan 1 im Bezug auf die OA miteinander, lassen sich einige Merkmale ableiten. Während bei den Modellläufen des EUR-44-Gebietes (Tab. 16a und 16b) zu erkennen ist, dass die jeweilige Einstellung des LSOIL-Schalters T im Vergleich zu F für alle Datensatzkombinationen und alle Validationsdatensätze immer etwas schlechter abschneidet, sieht man diesen Effekt beim GER-11-Gebiet nicht. Bei diesem Untersuchungsgebiet sticht jedoch besonders bei der Berechnung auf zeitlicher Basis (Tab. 16c) das bessere Abschneiden der ASTER-Daten im Vergleich zu den Läufen mit GTOPO heraus. Auf räumlicher Basis (Tab. 16d) ist dies aufgrund der minimal schlechteren Werte des id01\_AFT-Laufs und der besseren Werte des id01\_GST-Laufs weniger gut nachzuvollziehen. Auffällig ist auch, dass beim EUR-44-Gebiet die räumlichen Werte jeweils niedriger sind als die zeitlichen Werte und dass sich das beim GER-11-Gebiet umkehrt.

Tab. 16: Übersicht der durchschnittlichen Overall-Accuracy-Werte (in %) für Experimentplan 1 (vgl. Abb. 44 für Erklärung der Modellversionen) für beide Untersuchungsgebiete (Datengrundlage: Tagesmittel von T2M, T2MAX, T2MIN und -summen von RR)

(a) EUR-44 (zeitliche Basis)

(b) EUR-44 (räumliche Basis)

(c) GER-11 (zeitliche Basis)

(d) GER-11 (räumliche Basis)

Ranking (dunkelgrün = Rang 1; dunkelrot = Rang 10) erfolgt spaltenweise.

(a)

	EOBSV19	EOBSV21	ERA5	ERA1	Mittelwert
id01	84.69	84.70	88.13	88.45	86.49
id01_GFT	84.43	84.43	87.71	88.23	86.20
id01_GFF	84.57	84.58	87.90	88.37	86.36
id01_GST	84.34	84.35	87.80	88.19	86.17
id01_GSF	84.48	84.50	87.88	88.29	86.29
id01_GHT	84.43	84.43	87.70	88.24	86.20
id01_GHF	84.60	84.61	87.92	88.36	86.37
id01_AFT	84.32	84.34	87.76	88.21	86.16
id01_AST	84.63	84.65	88.10	88.30	86.42
id01_ASF	84.86	84.89	88.31	88.58	86.66
Mittelwert	84.54	84.55	87.92	88.32	86.33

(b)

	EOBSV19	EOBSV21	ERA5	ERA1	Mittelwert
id01	71.41	71.45	75.20	76.13	73.55
id01_GFT	70.97	71.01	74.52	75.77	73.07
id01_GFF	71.36	71.40	74.96	75.97	73.42
id01_GST	70.74	70.77	74.42	75.59	72.88
id01_GSF	70.90	70.94	74.50	75.71	73.01
id01_GHT	71.07	71.10	74.70	75.87	73.19
id01_GHF	71.31	71.34	74.81	75.92	73.35
id01_AFT	70.12	70.16	73.96	75.31	72.39
id01_AST	71.20	71.24	74.72	75.87	73.26
id01_ASF	71.62	71.64	75.33	76.32	73.73
Mittelwert	71.07	71.10	74.71	75.84	73.18

(c)

	EOBSV19	EOBSV21	ERA5	Mittelwert
id01	79.87	79.90	82.46	80.74
id01_GFT	79.84	79.87	82.42	80.71
id01_GFF	79.87	79.90	82.47	80.75
id01_GST	79.87	79.90	82.47	80.75
id01_GSF	79.86	79.89	82.46	80.73
id01_GHT	79.85	79.88	82.46	80.73
id01_GHF	79.85	79.88	82.47	80.73
id01_AFT	79.90	79.93	82.50	80.78
id01_AST	79.90	79.93	82.49	80.77
id01_ASF	79.91	79.94	82.50	80.79
Mittelwert	79.87	79.90	82.47	80.75

(d)

	EOBSV19	EOBSV21	ERA5	Mittelwert
id01	82.62	82.55	85.22	83.46
id01_GFT	82.62	82.55	85.20	83.46
id01_GFF	82.62	82.55	85.22	83.46
id01_GST	82.64	82.57	85.24	83.48
id01_GSF	82.63	82.56	85.22	83.47
id01_GHT	82.61	82.53	85.19	83.44
id01_GHF	82.62	82.55	85.21	83.46
id01_AFT	82.63	82.55	85.22	83.47
id01_AST	82.66	82.59	85.23	83.49
id01_ASF	82.64	82.57	85.22	83.48
Mittelwert	82.63	82.56	85.22	83.47

Zusammenfassend zeigt die Analyse der Ergebnisse von Experimentplan 1, dass der neue Topographiedatensatz ASTER in Kombination mit dem Bodendatensatz Soilgrid besonders bei Verwendung des Sand- und Tonanteils (LSOIL=F) in beiden Untersuchungsgebieten sowie auf räumlicher und zeitlicher Basis für alle Vergleichsdaten im Mittel über die vier Variablen etwas besser abschneidet als der Lauf mit den bisherigen Datensätzen. Die OA-Werte für die einzelnen Variablen, die nicht dargestellt sind, weichen zwar betragsmäßig leicht vom dargestellten Mittelwert ab, die Abstufung der Validationsdaten und die Reihenfolge der Modellläufe unterscheidet sich aber in den meisten Fällen nicht. Die beschriebenen „Rankingmuster“ entsprechen ebenfalls denen des Kappa-Koeffizienten. Dieser liegt bei den GER-11-Werten je nach Validierungsdatensatz zwischen 0.56 und 0.68 und für EUR-44 zwischen 0.37 und 0.73, was einer guten Abbildung entspricht. Die Ergebnisse der Monatswerte zeigen auch sehr ähnliche Muster, sodass auf diese nicht näher eingegangen wird.



Die in Tabelle 17 dargestellten Ergebnisse der OA-Analyse für die Modellläufe von Experimentplan 2 (vgl. Abb. 45) zeigen ebenfalls, dass die Modelländerungen für die beiden Untersuchungsgebiete zu leicht unterschiedlichen Reihenfolgen führen. Wie zuvor bei der Auswertung des Experimentplan 1 wird mit dem Verwenden des Sand- und Tonanteil statt der Textur (Lauf FFFF und TFFF) für das EUR-44-Gebiet eine Verbesserung erzielt, beim GER-11-Gebiet ist dies jedoch nicht zu erkennen. Bei beiden Untersuchungsgebieten wird ein höherer OA-Wert erreicht, wenn das neue Bodentemperaturschema in Kombination mit LSOIL=T (Lauf TFFF im Vergleich zu Lauf TFFT) verwendet wird. Bei LSOIL=F hingegen führt die Nutzung des neuen Schemas bei EUR-44 zu einer Verschlechterung und bei GER-11 nur zu minimalen Änderungen (vgl. Lauf FFFF und Lauf FFFT). Die isolierte Änderung hinsichtlich fünf verschiedener Bodenschichtparameter (L5LAYER\_TOTAL=T) resultiert bei allen vier Analysen (Tab. 17a-17d) in niedrigeren OA-Werten (vgl. Lauf FFFT). In Kombination mit dem neuen Bodentemperaturschema werden dagegen bessere Genauigkeiten sowohl für EUR-44 als auch für GER-11 erreicht. Dieser Lauf FFFT schneidet bei beiden EUR-44-Analysen besser als der bisherige Lauf TFFF ab (Rang 3) und bei den beiden GER-11-Analysen im Mittel ebenfalls etwas besser als Lauf TFFF (Rang 2), wenngleich der Unterschied zwischen den verschiedenen Läufen insgesamt sehr gering ist.

Tab. 17: Übersicht der durchschnittlichen Overall-Accuracy-Werte (in %) für Experimentplan 2 (vgl. Abb. 45 für Erklärung der Modellversionen) für beide Untersuchungsgebiete (Datengrundlage: Tagesmittel von T2M, T2MAX, T2MIN und -summen von RR)

(a) EUR-44 (zeitliche Basis)

(b) EUR-44 (räumliche Basis)

(c) GER-11 (zeitliche Basis)

(d) GER-11 (räumliche Basis)

Ranking (dunkelgrün = Rang 1; dunkelrot = Rang 6) erfolgt spaltenweise.

(a)

	EOBSV19	EOBSV21	ERA5	ERA1	Mittelwert
id03_TFFF	84.29	84.30	87.66	88.18	86.11
id03_TFFT	84.59	84.58	88.02	88.38	86.39
id03_FFFF	84.79	84.80	88.13	88.56	86.57
id03_FFTF	84.32	84.34	87.77	88.25	86.17
id03_FFFT	84.41	84.42	87.77	88.26	86.22
id03_FFTT	84.52	84.54	87.99	88.37	86.36
Mittelwert	84.49	84.50	87.89	88.33	86.30

(b)

	EOBSV19	EOBSV21	ERA5	ERA1	Mittelwert
id03_TFFF	70.81	70.83	74.38	75.60	72.91
id03_TFFT	71.30	71.33	74.95	76.02	73.40
id03_FFFF	71.68	71.72	75.33	76.30	73.76
id03_FFTF	70.93	70.97	74.46	75.67	73.01
id03_FFFT	70.91	70.95	74.41	75.67	72.98
id03_FFTT	71.05	71.08	74.80	75.87	73.20
Mittelwert	71.11	71.15	74.72	75.86	73.21

(c)

	EOBSV19	EOBSV21	ERA5	Mittelwert
id03_TFFF	79.89	79.92	82.56	80.79
id03_TFFT	79.88	79.92	82.58	80.79
id03_FFFF	79.87	79.90	82.56	80.78
id03_FFTF	79.86	79.89	82.55	80.77
id03_FFFT	79.87	79.91	82.56	80.78
id03_FFTT	79.88	79.92	82.58	80.79
Mittelwert	79.88	79.91	82.56	80.78

(d)

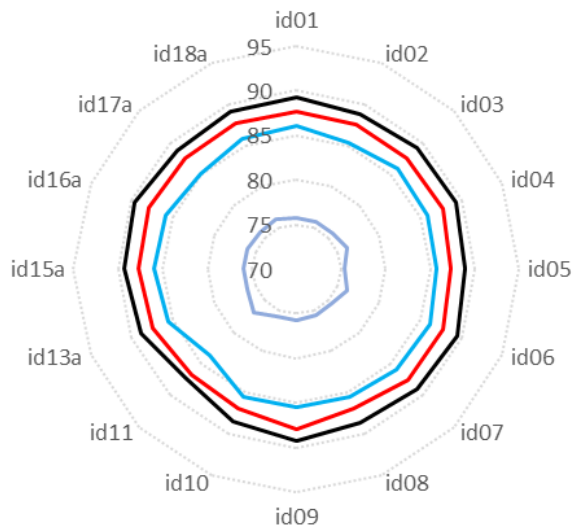
	EOBSV19	EOBSV21	ERA5	Mittelwert
id03_TFFF	82.59	82.52	85.21	83.44
id03_TFFT	82.60	82.53	85.23	83.45
id03_FFFF	82.58	82.51	85.20	83.43
id03_FFTF	82.56	82.50	85.20	83.42
id03_FFFT	82.59	82.52	85.22	83.44
id03_FFTT	82.59	82.52	85.22	83.44
Mittelwert	82.58	82.52	85.21	83.44

Für die OA-Analyse von Experimentplan 3 (vgl. Abb. 46) sind in Abbildung 77 die aus den entsprechenden Tabellen erstellten Netze verschiedener Variablen für EOBSV21 und ERA5 für beide Untersuchungsgebiete dargestellt. Anhand der Netzabbildung können theoretisch Veränderungen sowohl zwischen den verschiedenen Variablen als auch zwischen den einzelnen Modellläufen gut detektiert werden. Im vorliegenden Fall variieren die Werte, die auf zeitlicher Basis berechnet wurden, zwischen den verschiedenen Variablen und je nach Validationsdatensatz deutlich mehr als zwischen den verschiedenen Modellläufen. Insgesamt lässt sich erkennen, dass alle jeweiligen Variablenwerte von EOBSV21 unterhalb derer von ERA5 liegen, was ebenfalls bei der Analyse der beiden anderen Experimentpläne mit den Variablenmittelwerten festgestellt worden ist. Mit vereinzelt Ausnahmen spiegeln alle dargestellten zeitlichen Werte die auf räumlicher Basis berechneten Werte wider, sodass auf diese nicht gesondert eingegangen wird. In allen vier Teilabbildungen ist abzulesen, dass T2M bei allen Modellläufen am besten mit den Validationsdaten übereinstimmt. Ferner bildet REMO die T2MAX besser ab als T2MIN. Die Ausnahme hiervon ist in Abbildung 77B zu sehen. Die OA-Werte von T2MAX und T2MIN sind auf dem gleichen Niveau (ca. 90 %) und für Lauf id09 und id11 ist der T2MIN-Wert höher. Für den Niederschlag werden OA-Werte von ca. 75 % mit EOBSV21 und ca. 80 % mit ERA5 für beide Gebiete erreicht.

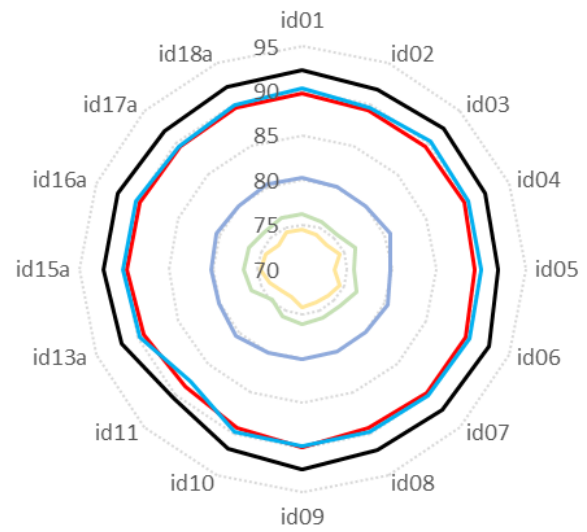
Die größten Veränderungen sind bei allen Variablen bei Lauf id11 (aktivierter L5LAY-Schalter) festzustellen, was die Analyse des RMSE und Korrelationskoeffizienten bestätigt (vgl. Kapitel 6.5.2.1). Sowohl bei EOBSV21 als auch bei ERA5 werden T2M sowie T2MIN schlechter und RR für beide Untersuchungsgebiete besser. Ferner ist bei den ERA5-Daten zu sehen, dass für die beiden Windkomponenten geringere Übereinstimmungen mit diesem Lauf vorliegen. T2MAX wird bei EUR-44, wie die beiden anderen Temperaturen, schlechter, bei GER-11 jedoch besser modelliert. Beim Niederschlag fällt zudem Lauf id05 negativ auf. Bei diesem Modelllauf werden statt konstanter Bodenparameter (wie in Lauf id04) für alle fünf Bodenschichten variable Parameter verwendet. Dies führte bereits bei der Analyse von Experimentplan 2 zu schlechteren Modellergebnissen. Für Lauf id06, bei dem nicht die eigene PTF (vgl. Kapitel 5.1.1) sondern die Parametrisierungen von de Vries (1963) und Farouki (1981b) für die Bodenparameter verwendet wird (vgl. Tab. 14), steigt die OA trotz des weiterhin aktiven L5LAYER\_TOTAL-Schalters auf das vorherige Niveau an. Unter Zuhilfenahme der nicht dargestellten Tabellen, aus denen die Netzabbildungen erstellt wurden, ist außerdem zu erkennen, dass auch alle drei Temperaturvariablen bei Lauf id06 bei EUR-44 höhere OA-Werte als die vorherigen und nachfolgenden Läufe erreichen. Unter anderem deswegen ist diese Kombination für die längeren Modellläufe als erste „Zwischenstufe“ ausgewählt worden. Ein weiterer kleiner „Sprung“ ist zwischen Lauf id13a und Lauf id15a zu erkennen. Alle Variablen erzielen höhere OA-Werte bei EUR-44 und auch bei GER-11 werden für T2M, T2MAX und RR ab Lauf id16a die höchsten Übereinstimmungen im Vergleich zu ERA5 gemessen. Der Wert von T2MIN nimmt bei GER-11 dagegen ab. Beim EUR-44-Gebiet ist Lauf id17a, bei dem das neue Bodentemperaturschema verwendet wird, in allen Variablen schlechter als der vorherige Lauf. Bei Lauf id18a, für den die neue Topographie in diesem Experiment zum ersten Mal eingesetzt wird, werden die OA-Werte jedoch wieder besser. Den positiven Effekt von ASTER in Kombination mit Soilgrid, welcher ab Lauf id03 genutzt wird, hat die Analyse von Experimentplan 1 bereits gezeigt.

Die beiden Windkomponenten U und V, die nur für ERA5 zum Vergleich verfügbar sind, liegen bei EUR-44 (Abb. 77B) zwischen 73–77 % und damit deutlich niedriger als die Werte für GER-11 (Abb. 77D), die mit 85–90 % sehr hohe OA-Werte erreichen. Auffällig ist zudem, dass bei EUR-44 die West-Ost-Komponente (U) und bei GER-11 die Süd-Nord-Komponente (V) für alle Läufe schlechtere Übereinstimmungen mit den Werten von ERA5 hat. Anders als bei den vorgestellten Temperaturvariablen und dem Niederschlag erkennt man besonders bei EUR-44 größere Variationen zwischen den verschiedenen Modellläufen. Es fällt auf, dass beispielsweise Lauf id02 und Lauf id03 für beide Gebiete die beiden Windkomponenten schlechter modellieren als der bisherige Lauf (id01). Sowohl Modelllauf id04 als auch Modelllauf id06 haben dagegen höhere OA-Werte als Lauf id01. Alle darauffolgenden Läufe, in denen hauptsächlich verschiedene hydrologischen Parametrisierungen getestet werden, erreichen nicht mehr so hohe Übereinstimmungen. Die Werte für Lauf id09 bei EUR-44 und für Lauf id08 bei GER-11 sind etwas höher als ihr jeweiliger Vor- und Nachgängerlauf, sind aber nicht so hoch wie die von Lauf id06. Ab Lauf id13a werden die OA-Werte wieder höher. Besonders Lauf id15a, bei dem die volumetrische Wärmekapazität neu parametrisiert und der als Kaltstart nur für EUR-44 gerechnet wurde, erreicht bei beiden Komponenten die höchste Übereinstimmung mit den Validationsdaten ERA5. Ähnliche Analysen mit anderen Qualitätsindizes führten zu sehr ähnlichen Mustern, weswegen Lauf id15a als zweiter Zwischenschritt für die längeren Warmstartmodellläufe ausgewählt worden ist. Generell sind die beiden Windkomponenten in ihren Rängen sehr ähnlich. Zwei auffällige Ausnahmen sind Modelllauf id03 und Modelllauf id18a bei GER-11. Während Lauf id18a für EUR-44 sowohl bei U als auch bei V mit Rang 3 bzw. Rang 5 gut abschneidet und auch für GER-11 bei U den zweithöchsten OA-Wert erreicht, hat der OA-Wert für V den vorletzten Rang. Dieses schlechtere Abschneiden der Süd-Nord-Komponente bei GER-11 ist auch mittels erzeugter OA-Karten (nicht graphisch dargestellt) nicht eindeutig erklärbar.

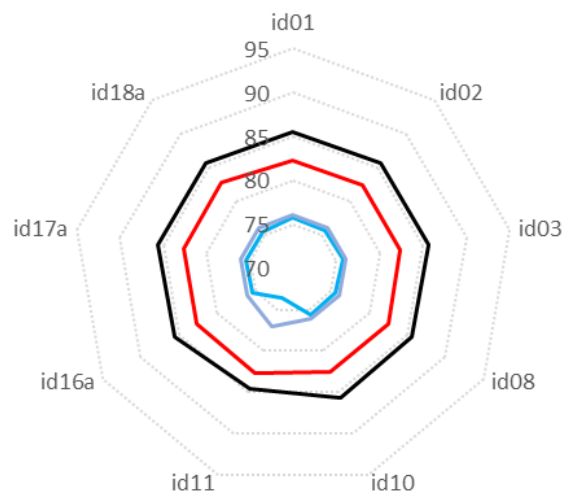
(A)



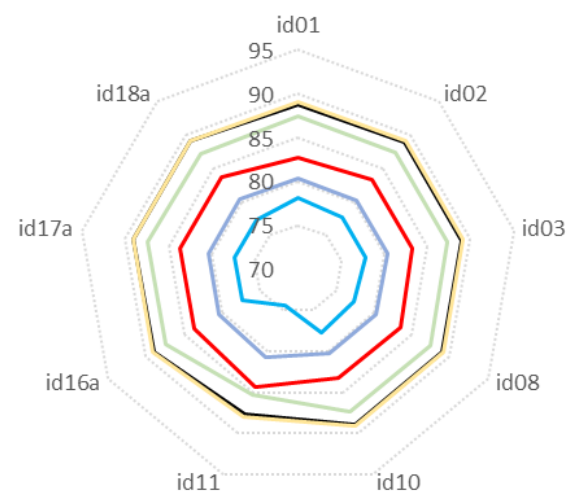
(B)



(C)



(D)



— T2M — T2MAX — T2MIN — RR — U — V

Abb. 77: Darstellung der Overall-Accuracy (in %) verschiedener Variablen pro Modelllauf des Experimentplans 3 (vgl. Abb. 46 für Erklärung der unterschiedlichen Modellparametrisierungen) für EUR-44 und GER-11

- (A) EUR-44 mit EOBSV21
- (B) EUR-44 mit ERA5
- (C) GER-11 mit EOBSV21
- (D) GER-11 mit ERA5

## **Zusammenfassung der Erkenntnisse zu den kurzen Modellläufen der drei Experimentpläne**

Im Vergleich zu den Validationsdaten treten für die drei Temperatur- und die Niederschlagsdaten sowohl beim Austausch der Topographie- und Bodendaten (Experimentplan 1) als auch bei der Verwendung der unterschiedlichen Schalterkombinationen (Experimentplan 2) kaum Unterschiede auf. Auch die Auswirkungen der verschiedenen Änderungen von Experimentplan 3 zeigen sich durch die vorgenommenen Analysen nicht so deutlich. Einzige Ausnahme ist der Modelllauf id11, bei dem durch das Verwenden des 5Layer-Schemas signifikante Unterschiede zu allen anderen Läufen direkt erkennbar sind. Alle getesteten Änderungen wirken sich kaum auf die Modellperformanz der vier Variablen, für die Beobachtungsdaten vorliegen, auf dem kurzen Modellzeitraum von einem Jahr, aus und werden teilweise zusätzlich durch die räumliche bzw. zeitliche Mittelwertbildung nicht sichtbar. Die Verbesserungen oder Verschlechterungen durch die neuen Daten oder neuen Parametrisierungen sind weder über die zeitliche noch über die räumliche Dimension konstant. Die gebildeten Reihenfolgen verschiedener Modellläufe hinsichtlich vorgestellter Qualitätskriterien ändert sich teilweise pro Variable, pro Untersuchungsgebiet und pro Saison, sodass auf der Grundlage dieser Daten nur eine Vorauswahl der Modellläufe für zusätzliche Warmstartläufe über eine längere Zeitspanne (Modellversion id01, id06, id15a und id18a) vorgenommen worden ist.

### **6.5.2.2 Analyseergebnisse der langen Modellläufe**

Für die längeren Läufe wurden die Jahre 2000-2018 gewählt, da diese von allen Beobachtungsdaten abgedeckt werden und einige Läufe weiter gerechnet oder für mehrere Analysen verwendet und somit Rechen- und Speicherkapazitäten eingespart werden können.

### **Einordnung der klimatischen Bedingungen in den Testjahren anhand der Beobachtungsdaten**

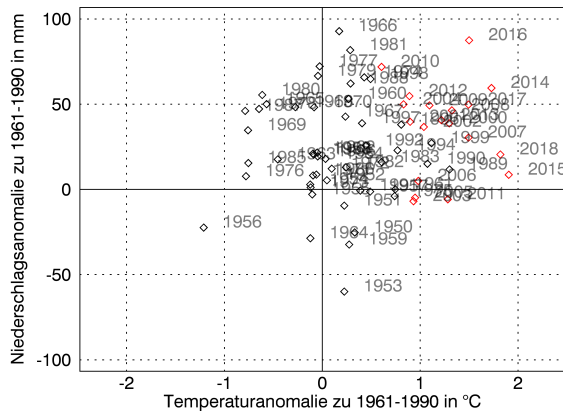
Um eine grobe Einschätzung hinsichtlich der klimatischen Situation dieser Testjahre zu machen, zeigt Abbildung 78A die Anomalien der 2 m-Temperatur und der Niederschlagssumme pro Jahr (Abb. 78A) und pro Saison und Jahr (Abb. 78B-78E) im Vergleich zur Klimareferenzperiode 1961-1990 basierend auf den Flächenmitteln der 0.25° EOBSV19-Daten. Die modellierten Testjahre sind rot markiert. Es fällt bei der jährlichen Auswertung sofort auf, dass alle Testjahre in der positiven Temperaturanomaliehälfte liegen. Außerdem sind die meisten Jahre auch Teil der positiven Niederschlagsanomaliehälfte. Nur 2003, 2005 und 2011 sind in diesem Jahrtausend über das ganze Jahr gesehen etwas zu trocken gewesen. Diese Niederschlagsdefizite sind aber im Vergleich zu denen in den Jahren 1950/51/53/56/59 und 1964 sehr klein.

Betrachtet man hingegen die einzelnen Jahreszeiten ergibt sich ein differenziertes Bild. Im Winter (Abb. 78B) sind hohe Temperaturanomalien von bis zu 3 °C für 2007 und 2016 gemessen worden. Gleichzeitig sind 2003, 2006, 2008 und 2010 etwas kühler als der Mittelwert der Referenzperiode

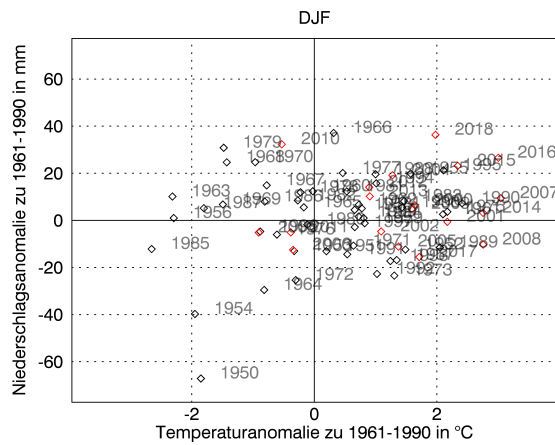
gewesen. Im Sommervergleich (Abb. 78B) liegt wieder ein Großteil aller Jahre (und alle Testjahre) in der positiven Temperaturanomalielhälfte. Es stechen vor allem 2003, 2010 und 2018 heraus, die mehr als  $1.7^{\circ}\text{C}$  zu warm waren. In dieser Abbildung wird außerdem deutlich, dass die erhöhten Temperaturen vermehrt mit zu trockenen Sommern einhergehen und nicht wie aus der Gesamtjahresübersicht zu vermuten ist, eher feuchte Jahre sind. Dass das Jahr 2003 ein so trockenes Jahr war, liegt nicht nur an dem zu trockenen Sommer, denn 2015 fehlte beispielsweise mit  $-22\text{ mm}$  doppelt so viel Regen wie 2003 mit  $-11\text{ mm}$ , sondern auch am viel zu trockenen Frühjahr, in dem 2003 bereits  $-20\text{ mm}$  fehlten (Abb. 78C). Auch 2009 und 2011 starteten mit einem sehr trockenen und warmen Frühling, aber der Sommer fiel in beiden Jahren feuchter aus, sodass eine andauernde Trockenperiode wie 2003 vermieden werden konnte, obwohl auch diese Jahre vergleichsmäßig warm waren. Der Herbst 2011 war jedoch ebenfalls zu trocken und zu warm (Abb. 78E).

Somit liegen in den Testjahren besonders viele warme Jahre vor, wovon gleichzeitig einige zu trocken waren. Da durch den prognostizierten Klimawandel in Zukunft diese jetzt noch extremen Jahre zur Normalität werden können, sollte das Klimamodell diese Bedingungen gut simulieren können. Deshalb eignen sich die ausgewählten Testjahre zur Überprüfung der Modellqualität.

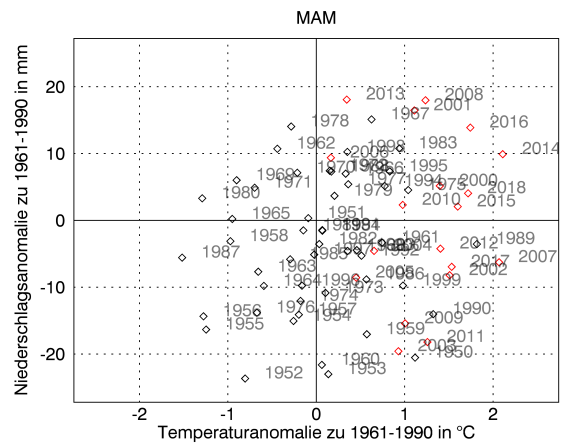
(A)



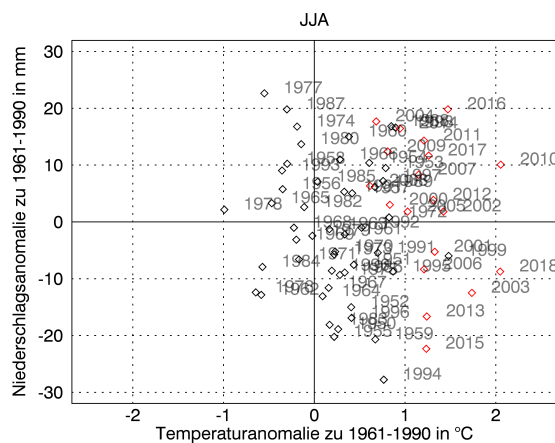
(B)



(C)



(D)



(E)

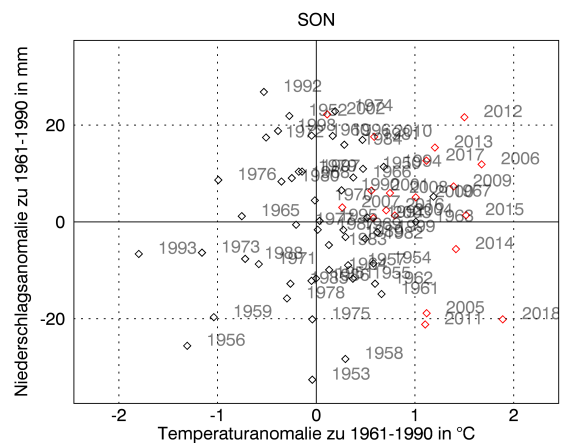


Abb. 78: Durchschnittliche Temperatur- und Niederschlagsanomalien der Jahre 1950-2018 im Vergleich zur Referenzperiode 1961-1990 laut EOBSV19 in 0.25° Auflösung (Jahre seit 2000 sind rot markiert.)

(A) Jahr / (B) DJF / (C) MAM / (D) JJA / (E) SON

(Die x- und y-Achse der Einzelabbildungen sind für eine bessere Unterscheidung der Jahreswerte an die jeweilige Zeitspanne angepasst.)

### Quartile des RMSE und Korrelationskoeffizienten

In Abbildung 79 sind, wie bereits für die kurzen Läufe (vgl. Abb. 74, Abb. 75 und Abb. 76), die jeweiligen 1., 2. und 3. Quartile der RMSEs und Korrelationskoeffizienten der vier Variablen RR, T2M, T2MAX und T2MIN für die gesamte Zeitperiode von 2000-2018 sowie für die jeweiligen Jahreszeiten dargestellt. Alle vier Warmstartläufe, die für das GER-11 vorliegen, werden mit den drei Validationsdatensätzen (EOBSV19, EOBSV21 und ERA5) verglichen. Im Gegensatz zu den vorherigen Analysen der kurzen Läufe ist beim Niederschlag ein kleiner Unterschied anhand der „Whisker“ der beiden E-OBS-Versionen zu erkennen. Die Modelldaten stimmen etwas besser mit der neueren Version (EOBSV21) überein. Ebenso sind in allen Teilabbildungen größere Unterschiede zwischen den verschiedenen Läufen zu sehen als bei der Betrachtung der kurzen Läufe. Die Lage der jeweiligen Quantile, die IQR und der Unterschied zwischen dem ERA5- und den E-OBS-Clustern ist jedoch gleich, sodass die bei den vorherigen Analysen der kurzen Läufe getroffenen Aussagen nicht aufgrund des kleineren Stichprobenumfangs oder aufgrund der Auswahl des Jahres 2000 hinterfragt werden müssen. Die Absolutwerte dieser Abbildung können jedoch nicht mit den anderen Analysen verglichen werden, da es sich hier um Monatswerte handelt.

Beim Niederschlag erkennt man nahezu keinen Unterschied zwischen den verschiedenen Modellläufen. Für T2M hingegen wird eine klare Verbesserung im Winter hinsichtlich des RMSE durch alle neuen Modellläufe erreicht. Die Verbesserung im DJF überträgt sich sogar auf den Jahreswert, obwohl in den anderen Jahreszeiten kaum Änderungen zu erkennen sind. Auch bei der Maximaltemperatur wirken sich die vorgenommenen Änderungen am Modell positiv auf die Performanz im Vergleich zu den Validationsdaten aus. Da insgesamt die Temperaturen ansteigen, bedeutet das jedoch gleichzeitig, dass die T2MIN-Werte schlechter modelliert werden. Bei dieser Variable sind auch die größten Unterschiede zwischen den Modellläufen zu erkennen, da sich nicht nur der RMSE-Wert wie bei den anderen Variablen, sondern auch der Korrelationskoeffizient verändert. Der Lauf id06 schneidet am schlechtesten ab. Der finale Lauf id18a hingegen liegt sehr nahe beim originalen Lauf id01.



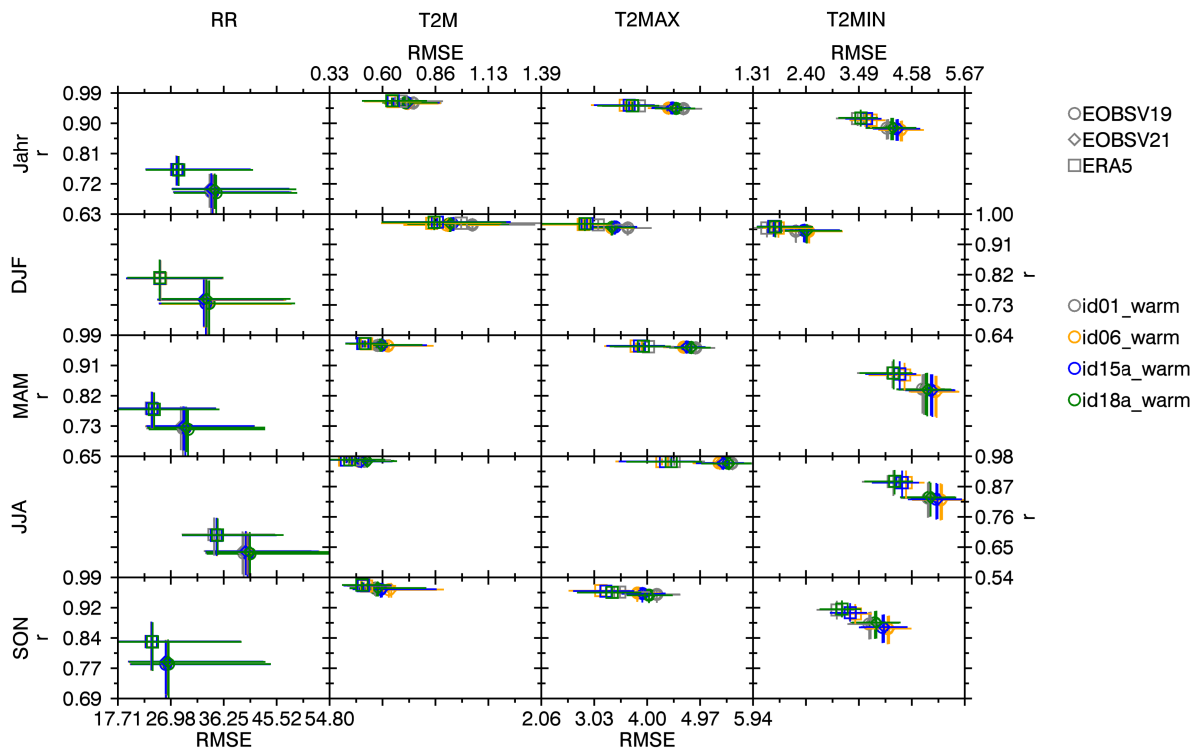


Abb. 79: Übersicht des 1., 2. und 3. Quartils des jeweiligen RMSE und Korrelationskoeffizienten der vier Warmstartläufe (Erläuterung der Parametrisierungsunterschiede vgl. Kapitel 6.2) im Vergleich zu den Validationsdaten für RR, T2M, T2MAX und T2MIN pro Jahr und Jahreszeit

Datengrundlage sind die entsprechenden Monatsdaten von 2000-2018 der Läufe des GER-11-Gebiets. Auf der x-Achse sind jeweils die RMSE-Werte aufgetragen (mit der Einheit mm für RR und °C für T2M, T2MAX und T2MIN). Die y-Achse stellt den Korrelationskoeffizienten  $r$  dar. Zu beachten ist, dass sich der Wertebereich der x- und y-Achse jeweils für die Variablen bzw. die Jahreszeiten ändert, um einen besseren Vergleich der Modellläufe zu ermöglichen.

### Ergebnisse der Qualitätsindizes

Neben dem Korrelationskoeffizienten und dem RMSE sind auch die anderen Qualitätsindizes, die in Kapitel 6.1.3 und 6.1.5 vorgestellt worden sind, für alle Läufe im Vergleich zu allen Validationsdaten berechnet worden. Wie bereits festgestellt wurde, ähneln sich die Ergebnisse der einzelnen Validationsdatensätze zwar nicht immer im Betrag aber in der Reihenfolge (vgl. Tab. 16, Tab. 17 und Abb. 79), sodass in Tabelle 18 nur die Werte im Vergleich zu ERA5 dargestellt sind, da dieser Validationsdatensatz zusätzlich zu den Temperatur- und Niederschlagsdaten auch Vergleichswerte für die beiden Windkomponenten enthält. Alle Werte für die einzelnen Indizes sind auf der Grundlage der räumlichen Monatsmittel bzw. -summen der Jahre 2000 bis 2018 der angegebenen Warmstartläufe für das EUR-44- (nur Lauf id01) bzw. für das GER-11-Gebiet (Lauf id01, Lauf id06, Lauf id15a und Lauf id18a) gebildet worden. Ein direkter Vergleich der beiden id01 Läufe ist daher aufgrund der unterschiedlichen Gebiete ungeeignet, bietet aber einen ersten Anhaltspunkt, da das GER-11-Gebiet in den vorherigen Analysen jeweils ungefähr den mittleren Werten und Abweichungen des EUR-44-Gebiets entsprochen hat. Anhand der Einfärbung sind die jeweiligen spaltenweisen Rangfolgen der Modellversionen abzulesen. Der Wert

des Laufs mit dunkelgrüner Einfärbung schneidet jeweils im Vergleich zu den Validationsdaten am besten und der Lauf mit dunkelroter Einfärbung am schlechtesten ab. Die anderen Läufe werden abhängig von diesen beiden Extremwerten mit verschiedenen Farben hinterlegt und sind daher nicht für alle Variablen und Indizes gleich. Zudem muss beachtet werden, dass das Einfärben der Werte mit den berechneten Originalwerten auf höherer Genauigkeit stattfindet, sodass trotz zweier gleich dargestellter Werte (vier Nachkommastellen) unterschiedliche Farbgebungen auftreten können.

Zunächst ist anhand von Tabelle 18 zu erkennen, dass die Einfärbungen von jeweils NSE und RSR (Tab. 18a und 18b), von PBIAS und RMSE (Tab. 18c und 18d) sowie vom Korrelations- und Kappa-Koeffizienten (Tab. 18e und 18f) für alle Modellläufe und Variablen sehr gut übereinstimmen. Die Rangfolge der Läufe dieser drei Gruppen variiert jedoch stark.

Der in Tabelle 18a dargestellte NSE (vgl. Formel 6.23) zeigt keine negative, sondern höhere positive Werte, sodass alle Modelldaten die Validationsdaten sehr viel besser abbilden, als das der Mittelwert der Beobachtungsdaten tun würde. Auch beim RSR in Tabelle 18b, bei dem es sich um den mit der Standardabweichung normierten RMSE handelt (vgl. Formel 6.22), liegen mit Ausnahme der T2MIN-Werte und dem RR-EUR-44-id01-Wert alle Werte unterhalb der 0.5-Grenze, die sehr gute Ergebnisse anzeigt. Sowohl beim NSE als auch beim RSR fließt die Varianz bzw. die Standardabweichung der Validationsdaten in die Berechnung ein, sodass auch die verschiedenen Variablen miteinander verglichen werden können. Beide Indizes zeigen an, dass T2M am besten von REMO abgebildet wird. Auffällig ist außerdem, dass sowohl RR als auch die beiden Windkomponenten bessere Indexwerte als T2MAX und T2MIN erreichen. Werden die verschiedenen Modellversionen miteinander verglichen, schneidet Lauf id01 für GER-11 bei beiden Windkomponenten am besten ab. Die maximale Temperatur wird von Lauf id06 am besten abgebildet. Für T2M und RR hingegen erreicht Lauf id18a die höchsten bzw. niedrigsten Werte für NSE bzw. RSR.

In Tabelle 18c ist der jeweilige PBIAS dargestellt. Es fällt sofort auf, dass einzelnen Variablen Werte in unterschiedlichen Größenordnungen besitzen. Dies liegt an den jeweils verwendeten Einheiten der Variablen, da für diesen Index der jeweilige Fehler ins Verhältnis zu den tatsächlichen Werten gesetzt wird (vgl. Formel 6.25). Ferner muss beachtet werden, dass die Einfärbung hier abhängig vom Betrag des PBIAS vorgenommen wurde, da das Vorzeichen anders als bei den anderen Indizes nicht die Stärke, sondern die Richtung des Fehlers anzeigt. Es lässt sich daher ableiten, dass T2MIN von allen Modellversionen im gleichen Maß überschätzt, wie T2MAX unterschätzt wird. Bei T2M sind die prozentualen Modellbias dagegen sehr viel kleiner, wobei GER-11-id06 zwar als einziger Lauf die ERA5-Werte überschätzt, aber gleichzeitig den geringsten Betrag erreicht und damit die Validationsdaten besser abbildet als die anderen Läufe. Der Niederschlag wird bei EUR-44-id01 ebenfalls überschätzt und bei allen GER-11-Läufen unterschätzt. Mit ca. 7 % PBIAS liefert Lauf id18a von allen GER-11-Versionen den geringsten prozentualen Bias-Wert. Diese Reihenfolge steht im Gegensatz zu dem in Tabelle 18d dargestellten RMSE-Wert für RR. Für diesen Index schneidet Lauf id18a wegen des größeren Wertes am schlechtesten ab. Zusätzlich fällt beim Vergleich dieser beiden Qualitätsindizes auf, dass für die U-

Tab. 18: Übersicht der Qualitätsindizes verschiedener Variablen für die langen Warmstartläufe (Erläuterung der Parametrisierungsunterschiede vgl. Kapitel 6.2) ermittelt zwischen den Modelldaten und ERA5 (Datengrundlage: Monatsmittel von T2M, T2MAX, T2MIN, U, V und -summen von RR 2000-2018)

(a) NSE / (b) RSR

(c) PBIAS (in %) / (d) RMSE (in °C bzw. mm bzw.  $\frac{m}{s}$ )

(e) Korrelationskoeffizient / (f) Kappa-Koeffizient

Ranking (dunkelgrün = Rang 1; dunkelrot = Rang 5) erfolgt spaltenweise. Die Werte werden auf räumlicher Basis für beide Untersuchungsgebiete ermittelt, sodass sich die jeweiligen EUR-44- und GER-11-Werte nicht auf das gleiche Gebiet beziehen.

(a)

	T2M	T2MAX	T2MIN	RR	U	V
EUR-44 (id01)	0.9967	0.7813	0.7335	0.7481	0.8308	0.9480
GER-11 (id01)	0.9968	0.7756	0.7227	0.8058	0.9560	0.9529
GER-11 (id06)	0.9979	0.8044	0.6771	0.8227	0.9538	0.9468
GER-11 (id15a)	0.9979	0.7976	0.6913	0.8218	0.9538	0.9484
GER-11 (id18a)	0.9981	0.7895	0.7083	0.8342	0.9485	0.9485

(b)

	T2M	T2MAX	T2MIN	RR	U	V
EUR-44 (id01)	0.0573	0.4677	0.5162	0.5019	0.4113	0.2280
GER-11 (id01)	0.0565	0.4737	0.5266	0.4407	0.2098	0.2171
GER-11 (id06)	0.0462	0.4423	0.5682	0.4211	0.2150	0.2307
GER-11 (id15a)	0.0454	0.4499	0.5556	0.4222	0.2149	0.2272
GER-11 (id18a)	0.0435	0.4588	0.5401	0.4072	0.2270	0.2269

(c)

	T2M	T2MAX	T2MIN	RR	U	V
EUR-44 (id01)	0.0306	0.9925	-1.0002	-6.7417	-16.8135	-51.5257
GER-11 (id01)	0.0807	1.1606	-1.0322	9.3529	-13.2302	30.3578
GER-11 (id06)	-0.0061	1.0785	-1.1242	8.0218	-13.2818	33.7859
GER-11 (id15a)	0.0174	1.0984	-1.0970	8.0925	-13.5451	32.9470
GER-11 (id18a)	0.0409	1.1180	-1.0694	7.1950	-16.2213	33.4866

(d)

	T2M	T2MAX	T2MIN	RR	U	V
EUR-44 (id01)	1.2846	3.1385	3.1513	30.2518	0.9309	0.7900
GER-11 (id01)	0.7309	3.4504	3.1556	34.2756	0.3940	0.3823
GER-11 (id06)	0.6851	3.2298	3.3898	34.4655	0.3983	0.3935
GER-11 (id15a)	0.6876	3.2825	3.3187	34.4697	0.3974	0.3917
GER-11 (id18a)	0.6867	3.3455	3.2292	34.7510	0.4076	0.4103

(e)

	T2M	T2MAX	T2MIN	RR	U	V
EUR-44 (id01)	0.8132	0.7947	0.7298	0.5514	0.8239	0.7940
GER-11 (id01)	0.9392	0.8928	0.8022	0.7398	0.9721	0.9570
GER-11 (id06)	0.9372	0.8895	0.8026	0.7376	0.9715	0.9562
GER-11 (id15a)	0.9389	0.8912	0.8049	0.7377	0.9717	0.9564
GER-11 (id18a)	0.9427	0.8835	0.8128	0.7367	0.9718	0.9567

(f)

	T2M	T2MAX	T2MIN	RR	U	V
EUR-44 (id01)	0.6688	0.6557	0.6333	0.4020	0.5943	0.5774
GER-11 (id01)	0.8543	0.8351	0.7436	0.6016	0.8275	0.8174
GER-11 (id06)	0.8530	0.8392	0.7398	0.6003	0.8266	0.8153
GER-11 (id15a)	0.8547	0.8392	0.7421	0.6001	0.8271	0.8153
GER-11 (id18a)	0.8561	0.8366	0.7469	0.5999	0.8274	0.8164

und V-Komponente sehr ähnliche RMSE-Werte ausgegeben werden. Bei den PBIAS-Werten sieht man jedoch, dass die U-Komponente durch alle Modellläufe überschätzt, die V-Komponente aber noch deutlicher (mit Ausnahme des einzigen EUR-44-Laufs) unterschätzt wird. Die Tatsache, dass die Windkomponenten sehr variabel sind und Modell- und Validationsdaten unterschiedliche Vorzeichen besitzen können, erschwert die Interpretation dieser PBIAS-Werte. Beispielsweise zeigt die Überschätzung der U-Komponente nur, dass der U-Wert vom Modell numerisch größer ist als der von ERA5. Das kann jedoch sowohl bedeuten, dass der Validationswert kleiner ist (bei positiven Vorzeichen der beiden U-Werte), gleichzeitig aber auch darauf hinweisen, dass der Validationswert einen größeren Betrag als das Modell hat (bei negativen Vorzeichen der beiden U-Werte).

Für die beiden Koeffizienten (Tab. 18e und 18f) wird eine klare Verbesserung durch die erhöhte Auflösung der GER-11-Werte erreicht, wenngleich die unterschiedlichen Bezugsgebiete berücksichtigt werden müssen. Außerdem lässt sich anhand beider Tabellen erkennen, dass vor allem RR und die beiden Windkomponenten am besten von GER-11-id01 abgebildet werden, was jedoch nur für die Windkomponenten mit den Aussagen der anderen Indizes übereinstimmt. Für T2M bestätigen beide Koeffizienten, dass der Lauf id18a höhere Übereinstimmungen mit ERA5 hat als die anderen drei GER-11-Läufe. Diese Reihenfolge für T2M wurde auch bei NSE und RSR beobachtet. Anders als bei den bisherigen vier Qualitätsindizes wird laut der beiden Koeffizienten die T2MIN von GER-11-id18a besser modelliert als durch GER-11-id01.

Insgesamt zeigt diese Übersicht der verschiedenen Indizes, dass das Ranking der Modellläufe anhand eines einzelnen Qualitätskriteriums irreführend sein kann. Zudem müssen die verschiedenen Variablen einzeln betrachtet und bewertet werden, da auch hier unterschiedliche Reihenfolgen für die Modellläufe vorliegen. Beim Vergleich der unterschiedlichen Modellwerte lässt sich feststellen, dass die Indizes jeweils sehr ähnlich sind. Dennoch kann festgehalten werden, dass der finale Modelllauf id18a in den meisten Fällen eine leichte Verbesserung im Vergleich zu den anderen Modellversionen liefert.

### **Vergleich der Bodentemperaturen**

Um die Auswirkungen aller Modelländerungen speziell auf die Bodentemperaturen zu testen, werden die beiden Warmstarts des ursprünglichen Modelllaufs (id01\_warm) und des finalen Modelllaufs (id18a\_warm) mit Validierungsdaten verglichen. Es gilt jedoch zu beachten, dass die Validierungsdaten für die Temperatur der Bodenschichten, die aus den Reanalysedaten von ERA5 stammen, nur für die oberen vier Bodenschichten zur Verfügung stehen. Damit wird nur der Bereich von der Oberfläche bis in eine Tiefe von 2.89 m abgedeckt, was nur ca. 30% des Modellbodens entspricht. Zudem stellt die Bodentemperatur der vierten Schicht von ERA5 die Temperatur in 1.945 m Tiefe dar, während bei REMO der Wert in der vierten Schicht für die Temperatur in einer Tiefe von 2.683 m steht. Daher wird zunächst auf die interpolierten vertikalen Profile der Temperaturdaten geschaut.

### Vergleich jährlicher Bodentemperaturprofile zu Validationsdaten für einen ausgewählten Standort

In Abbildung 80 ist die tatsächliche mittlere Monatstemperatur pro Bodenschicht für ERA5 (Abb. 80A), die jeweilige Differenz zu Modelllauf id01 (Abb. 80C) und id18a (Abb. 80B) und die Differenz von id18a zu id01 (Abb. 80D) für das Jahr 2018 für Würzburg (vgl. schwarzes Rechteck in Abb. 71B) dargestellt. Das Jahr 2018 wurde ausgewählt, weil es das letzte Jahr in der modellierten Zeitreihe ist und so der Unterschied zwischen den Modellläufen, der eventuell durch ein zusätzliches „nachträgliches Einschwingen“ wegen der veränderten Bodenkarte eintreten und sich auf die Bodentemperaturen auswirken könnte, am geringsten wäre. Außerdem wurde nicht die ganze Modellperiode von 19 Jahren gezeigt, weil sich die Differenzmuster der einzelnen Jahre sehr stark ähneln und die Unterschiede innerhalb eines Jahres größer sind als zwischen den einzelnen Jahren. Daher werden die Muster exemplarisch anhand von 2018 dargestellt. Gleiches gilt für die Auswahl dieses kleinen räumlichen Ausschnittes. Die in diesem Ausschnitt festgestellten Merkmale lassen sich auf das Mittel des gesamten Modellgebiets übertragen.

Die unterschiedlichen Jahregänge der einzelnen Bodenschichten, die bereits beim räumlichen Mittel der Modellzeitreihe (vgl. Abb. 71A unten) nachgewiesen wurde, sind beim Profil der ERA5-Daten ebenfalls zu erkennen. Durch die Verwendung von Monatsdaten und der proportionalen Darstellung der Schichttiefen ist die größere Variabilität der oberen Bodenschichten jedoch aus dieser Darstellung nicht ableitbar. Es lässt sich aber feststellen, dass zwischen dem Zeitpunkt des jeweiligen Minimums (Ende Januar) und Maximums (Ende Juli) der Temperatur der obersten Bodenschicht und dem der dargestellten vierten Schicht ca. ein Monat liegt. Zudem kennzeichnen die eng beieinander liegenden Thermoisoplethen die Zeiten und Tiefen, in denen sich der Boden schnell erwärmt (MAM) oder abkühlt (Oktober, November und Dezember).

Die Differenzmuster der Ergebnisse beider REMO-Läufe und der Validationsdaten ERA5 (Abb. 80B-Abb. 80D) verhalten sich konträr zu den tatsächlichen Werten (Abb. 80A), was besonders gut in der 3. und 4. Bodenschicht zu sehen ist. Kalte Temperaturen werden überschätzt, während wärmere Bodentemperaturen vom Modell unterschätzt werden. Auffällig ist, dass teilweise ein „vertikaler“ Vorzeichenwechsel bei der Differenz stattfindet, der vor allem am Anfang und Ende des dargestellten Zeitraums zu erkennen ist. Beide Modellversionen sind im Winter für die oberen Schichten zu kalt und im Frühling, Sommer und Herbst zu warm. In den unteren Schichten hingegen wird die Bodentemperatur von beiden Modellen in den ersten vier Monaten des Jahres überschätzt und in den restlichen Monaten unterschätzt. Dieses Muster ist bei beiden Modellläufen gleich. Für den Betrag der Differenz ergeben sich jedoch für die Modellversion id18a viel kleinere Werte. Vor allem in der vierten Schicht werden Anfang des Jahres bis zu einem halben Grad kühlere Temperaturen und im Sommer bis zu einem Grad wärmere Temperaturen im Vergleich zu id01 modelliert, sodass sich die beiden extremen Anomalien zu den ERA5-Daten abschwächen. Dennoch sind mit 2.5 °C im ersten Halbjahr und 3 °C im zweiten Halbjahr die Differenzen der vierten Bodenschicht zu den Validierungsdaten wesentlich höher, als die um  $\pm 0.5$  °C schwankenden Temperaturdifferenzen in der obersten Schicht. Eine leichte ganzjährige Erwärmung in Lauf id18a im Vergleich zu id01 in der obersten Schicht, die

im Winter jedoch etwas stärker ausfällt, führt zu einer Reduzierung der negativen Anomalie, aber gleichzeitig auch zu einer minimalen Erhöhung des Fehlers im Sommer. Ein ähnlicher Temperaturanstieg von Modelllauf id18a zu id01 findet auch in der fünften Modellschicht statt, welche jedoch nicht im Vergleich zu den Beobachtungsdaten gesetzt werden kann. Insgesamt führen die Änderungen dieses Experimentplans zu wesentlich höheren Temperaturdifferenzen als die von Experimentplan 2 (vgl. Abb. 72).



### **Vergleich des Wertebereichs aller Bodenschichten mit Validationsdaten für einen ausgewählten Standort**

Um den Wertebereich der verschiedenen Modellläufe für unterschiedliche Jahreszeiten mit den Beobachtungsdaten zu vergleichen, sind in Abbildung 81 die jeweiligen Boxplots der verfügbaren Validationsdaten und der jeweiligen Warmstartläufe id01, id06, id15a und id18a für die fünf REMO-Bodenschichten dargestellt. Die Werte sind aus den jeweiligen monatlichen Mittelwerten von 2000-2018 für den Würzburger Beispieldata (vgl. schwarzes Rechteck in Abb. 71B) berechnet. Trotz der Tatsache, dass ERA5 Daten für vier Bodenschichten liefert, werden nur die oberen drei Schichtwerte zum direkten Vergleich herangezogen, weil diese nicht interpolierten Bodenschichtmittelmittelwerte ungefähr auf den gleichen Tiefen liegen wie die Schichtmittelpunkte von REMO (nur 0.25–13.55 cm Unterschied). Der Wert der vierten ERA5-Schicht würde dagegen mehr als 70 cm höher liegen als der von REMO.

Es fällt auf, dass die Unterschiede der beiden Beobachtungsdaten (ERA5 und DWD), die zusammen nur für die erste Bodenschicht vorliegen, je nach Saison stark variieren. Besonders ausgeprägt ist der Unterschied im Sommer. Während hier alle Modellläufe sehr gut mit den ERA5-Daten übereinstimmen, liegt der untere Whisker der DWD-Daten beim Median der restlichen Daten. Der Wertebereich der DWD-Daten ist um ca. 4 °C nach oben verschoben, denn die jeweiligen Whiskerlängen und die IQR nimmt ähnliche Größen an wie bei den Boxplots der anderen Daten. Im Winter unterscheiden sich diese Boxplotkenngrößen deutlich. Der Wertebereich von ERA5 ist größer als der von DWD und die Modellversionen umfassen alle einen noch größeren Bereich, sind jedoch zueinander sehr ähnlich. Beide Beobachtungsdatensätze weisen jeweils höhere Quartile aus. In den Übergangsjahreszeiten hingegen liegen die Modellläufe mit dem Median jeweils zwischen dem kälteren ERA5- und dem wärmeren DWD-Wert.

Die Temperaturwerte für Bodenschicht 2 und 3 verhalten sich außer im Winter vom Verhältnis zu den ERA5-Daten jeweils gleich. In Frühling liegen alle Modelltemperaturen über denen von ERA5 und unterscheiden sich zwischen den verschiedenen Modellversionen kaum. Auch im Sommer ist in der 2. Bodenschicht kaum ein Unterschied zwischen den Modellläufen zu erkennen. Die Mediane der Bodentemperatur liegen jedoch unter dem von ERA5. Zudem ist die modellierte Temperatur der 3. Bodenschicht kälter als die der Validationsdaten. Die Modellwerte zeigen hier aber, dass zunächst eine sukzessiv höhere Differenz zu den Validationswerten erzielt wird. Der Median von Lauf id18a liegt jedoch näher als der des bisherigen Modelllaufs id01. Gleichzeitig sieht man aber auch, dass der durch die Whiskerlängen abgedeckte Wertebereich hier, anders als in den anderen Jahreszeiten, stark unterschätzt wird. Der Herbst ähnelt vom Verhältnis der Mediane zueinander den Winterwerten. Interessanterweise sind sich die Läufe id01 und id18a und besonders die Läufe id06 und id15a ähnlich, obwohl id01 und id18a bezüglich der Parametrisierung die größten Unterschiede aufweisen. In den Abbildungen für Bodenschicht 4 und 5 ist dieses Phänomen ebenfalls bei fast allen Jahreszeiten zu erkennen. Die Quartile des id01-Laufs liegen jeweils zwischen denen von id18a und den anderen beiden Läufen. Es fällt jedoch auf, dass die Whiskerlängen von id18a für alle Jahreszeiten länger sind als die der restlichen Läufe. In diesem Lauf wird eine größere Amplitude der tiefen Bodentemperatur erreicht, das heißt, im Winter



und Frühjahr wird es kälter, im Sommer und Herbst wird es deutlich wärmer, was ebenso in Abbildung 80D anhand des Jahres 2018 festgestellt wurde.

Die Analyse, die alle im GER-11-Gebiet befindlichen Pixel berücksichtigt und deren Ergebnisse hier nicht dargestellt sind, liefert die gleichen Erkenntnisse wie die vorgestellte Analyse des Einzelpixels, wenngleich die Absolutwerte stärker streuen und die Unterschiede aufgrund des vielfach größeren Stichprobenumfangs in den Abbildungen weniger gut zu erkennen sind.

Die größeren Unterschiede zwischen den einzelnen Modellläufen in den tieferen Bodenschichten lassen sich zunächst damit erklären, dass der Einfluss der Atmosphäre, die für alle Modellläufe mit den gleichen Daten angetrieben wird und an der keine direkten Änderungen vorgenommen worden sind, im Vergleich zum Einfluss der Bodenprozesse auf die Bodentemperaturen, welche in den verschiedenen Läufen stark verändert sind, mit zunehmender Tiefe generell abnimmt. Dadurch können hier die vorgenommenen Unterschiede in der Bodenparametrisierung deutlicher werden. Außerdem sind die unteren Bodenschichten sehr viel dicker. Eine kleine Änderung der Eingangsparameter für diese Schichten gilt für mehr Bodenvolumen und resultiert in größeren Unterschieden der Energiemengenänderung und damit größeren Unterschieden in der Bodentemperatur als in den nur wenige Zentimeter dicken oberen Bodenschichten. Ungünstigerweise liegen eben für diese tieferen Bodenschichten, wo sich signifikantere Unterschiede zwischen den einzelnen Parametrisierungen hervorheben, keine Beobachtungsdaten vor.

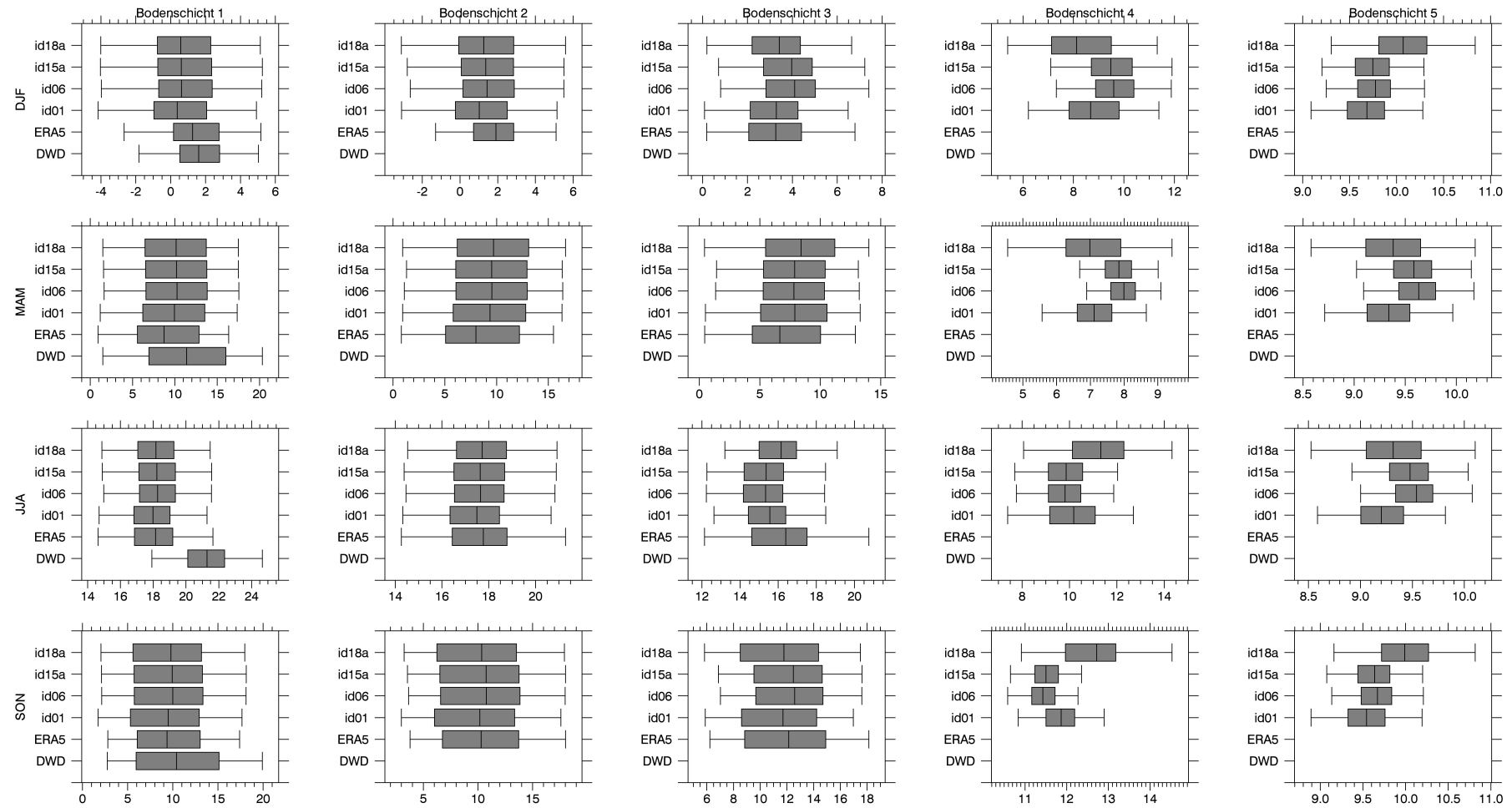


Abb. 81: Boxplot der mittleren Bodentemperaturen pro Bodenschicht und Jahreszeit für den Beispielpixel (wie in Abb. 80) für verschiedene Beobachtungs- und Modelldaten (Parametrisierungsunterschiede vgl. Kapitel 6.2) der Jahre 2000-2018  
 Der jeweilige Wertebereich der Bodentemperatur (x-Achse, Angaben jeweils in °C) wurde sowohl an die Bodenschicht als auch an die Jahreszeit angepasst, um die Unterschiede der Modellläufe und Beobachtungsdaten in den Einzelabbildungen besser kenntlich zu machen.

### Räumlicher Vergleich der Modell- und Validationsdaten für die Temperatur der obersten Bodenschicht

Um ebenfalls eine Einschätzung der räumlichen Modellqualität der Bodentemperatur vornehmen zu können, werden die Daten der obersten Bodenschicht für Deutschland in einer Auflösung von  $0.11^\circ$  mit den vorgestellten Beobachtungsdaten verglichen. Abbildung 82 zeigt die beobachteten Bodentemperaturen der DWD-Daten und den Vergleich sowohl mit den Bodendaten von ERA5 als auch mit den Daten der vier Warmstartläufe. In der obersten Reihe der Abbildung sind zum einen der Mittelwert aller Daten und zum anderen jeweils die Mittelwerte über die vier Jahreszeiten dargestellt. In den weiteren Reihen folgt die jeweilige Differenz von ERA5 und von den vier Modellläufen zu den DWD-Daten. Generell lässt sich feststellen, dass sich die Differenzmuster pro Jahreszeit stärker unterscheiden als die Differenzmuster zwischen den verschiedenen Modellläufen. Insgesamt werden die DWD-Bodentemperaturen bis auf wenige Ausnahmen sowohl von ERA5 als auch von den Modelldaten unterschätzt. In den jährlichen Differenzen liegt bei allen fünf Vergleichsdaten eine Abweichung von  $-1^\circ$  bis  $-2^\circ\text{C}$  (vereinzelt in Süddeutschland sogar  $-3^\circ\text{C}$ ) vor. Durch die genauere Analyse der einzelnen Jahreszeiten zeigt sich, dass die großen Differenzbeträge durch besonders große Abweichungen von teilweise über  $-4^\circ\text{C}$  im Sommer zustande kommen, die in Regionen auftreten, wo die DWD-Daten die höchsten Bodentemperaturen anzeigen.

Unterschiede zwischen den Modellläufen sind kaum zu erkennen. Während Modellversion id15a und id18a aufgrund der Klassenbreite der Differenzkarte von  $1^\circ\text{C}$  nahezu identisch sind, zeigen sich leichte Unterschiede im Muster zwischen Lauf id15a und id01 im Herbst und etwas größere Unterschiede zu id06 in DJF und SON. Im nördlichen Bereich von Deutschland ergeben sich mit id06 positive Differenzen zu den DWD-Daten statt negative wie bei id01, id15a und id18a. Der Betrag der Differenz bleibt jedoch gleich. Eine Verbesserung der Modellbodentemperatur im Vergleich zu den DWD-Daten ist dagegen im Herbst in Südwestdeutschland an der Grenze zu Frankreich zu erkennen. Während beim Modelllauf id01 noch größere Flächen mit Differenzen von  $-2^\circ\text{C}$  auftreten, nehmen diese mit id06, id15a und id18a ab.

Besonders auffällig ist ein Großteil der Mittelgebirge im DJF. Für sie werden in allen Modellläufen die höchsten Differenzen zu den DWD-Daten erreicht. Die gleichen Flächen treten im Vergleich zu den ERA5-Daten jedoch nicht auf. Stattdessen orientiert sich die Differenz zwischen DWD und ERA5 an den Absolutwerten der DWD-Daten. Das heißt, die wärmeren Flächen (v. a. Nord- und Westdeutschland) sind in ERA5 noch wärmer, während die kälteren Flächen in ERA5 noch kältere Werte haben. Die Beträge der DJF-Differenzen zwischen den beiden Validationsdatensätzen liegen im Bereich von  $\pm 1^\circ\text{C}$  und nehmen damit bei weitem nicht das Ausmaß an wie im Frühjahr oder Sommer. Da das auffällige Differenzmuster zwischen DWD- und Modelldaten nur im Winter auftritt, könnte es an einer vorhandenen Schneedecke liegen, die in den DWD-Daten zu Fehlern führen kann (vgl. Kapitel 4.5.1.5).

Die starke Ähnlichkeit zwischen ERA5 und allen Modellläufen könnte neben den atmosphärischen Rahmenbedingungen durch die gleiche Bodentexturdatengrundlage bei Lauf id01 verstärkt werden. Beide verwenden FAO-Bodendaten und sogar die Einteilung in fünf Texturklassen ist gleich. Bei den DWD-Daten dagegen werden (wie in Kapitel 4.5.1.5 beschrieben) die Bodentemperaturen

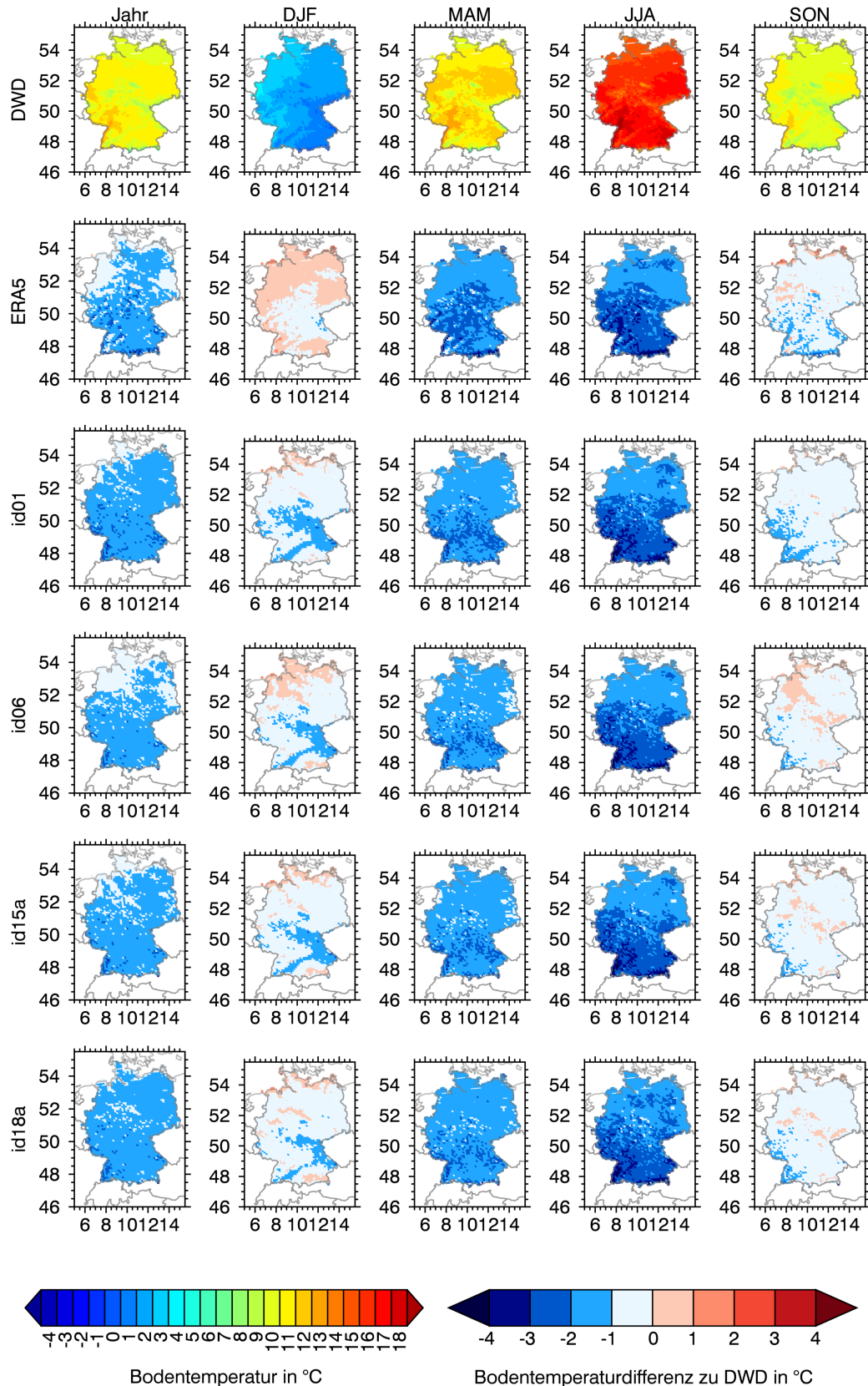


Abb. 82: Räumlicher Vergleich der jährlichen und saisonalen Bodentemperaturen zwischen den Validationsdaten und unterschiedlichen Modellläufen (Warmstarts vgl. Kapitel 6.2) für Deutschland

1. Reihe: Absolutwerte der DWD-Daten
2. Reihe: Differenz ERA5-DWD
- 3.-6. Reihe: jeweilige Differenz zwischen id01/ id06/ id15a/ id18a - DWD

durch das Anwenden eines Bodenmodells auf die Texturklasse „sandiger Lehm“ normiert. Ein räumlicher Vergleich der Fehlermuster (RMSE und Differenz) zwischen DWD und Modelldaten mit den Bodenparametern sowohl optisch (vgl. Abb. 82 mit Abb. 55, Abb. 56, Abb. 57 und den anderen nicht dargestellten Parametern) als auch numerisch durch die Berechnung von räumlichen Korrelationskoeffizienten zeigt jedoch keinen direkten Zusammenhang. Der betragsmäßig höchste Korrelationswert wird zwischen dem RMSE-Wert und DLAMBDA mit ca. 0.30 erreicht. Die Korrelationsanalyse zwischen den absoluten Differenzwerten und den Parameterwerten liefert sehr viel kleinere und nicht signifikante Werte.

### **Zusammenfassung der Erkenntnisse über Änderungen in den verschiedenen Bodentemperaturen**

Generell wird durch die Analyse der Bodentemperaturen deutlich, dass es zwar sowohl räumliche als auch zeitliche Unterschiede zwischen den einzelnen Modellläufen gibt, diese jedoch nicht so groß sind, wie die in Kapitel 6.4 gezeigten Änderung der Bodenparameter vermuten lassen. Die vorgenommenen Änderungen der konstanten Bodenparameter führt nicht zu systematischen Änderungen der Bodentemperaturen, da diese neben den Konstanten von weiteren zeitlich variablen Faktoren wie dem jeweiligen Wassergehalt und den Energieflüssen hochgradig abhängig sind. Außerdem konnte festgestellt werden, dass der Unterschied zwischen und zu den vorliegenden Beobachtungsdaten wesentlich größer ist als die Variabilität der Modellläufe. Das Fehlen von geeigneten Beobachtungsdaten speziell für die unteren Bodenschichten ist besonders ungünstig, da hier die betragsmäßig größten Änderungen durch die neu eingeführten Daten und Parametrisierungen auftreten. Anhand der vorliegenden Validationsdaten kann aber abgeleitet werden, dass mit Modelllauf id18a im Vergleich zum bisherigen Modelllauf id01 eine leichte Verbesserung bei der Darstellung der Wintertemperaturen erzielt wird. Des Weiteren ist in den tieferen Bodenschichten zu erkennen, dass durch Lauf id18a eine größere Temperaturamplitude vor allem aufgrund der veränderten thermalen Parametrisierung modelliert wird. Aus den Analysen geht auch hervor, dass die kombinierten thermischen und hydrologischen Parameteränderungen (Experimentplan 2 und 3) größere Änderungen hervorrufen als die Änderung der räumlichen Verteilung durch die Datensatzänderungen (Experimentplan 1). Eine separate Validierung des Wassergehaltes, um die Auswirkungen der verschiedenen Modellversionen ins Verhältnis zu den thermischen Auswirkungen zu setzen, konnte nicht durchgeführt werden, da sich der in REMO berechnete absolute Wassergehalt auf die jeweilige Wurzeltiefe des einzelnen Pixels bezieht, die Validierungsdaten (z. B. ERA5) jedoch den Wassergehalt für Bodenschichten ausweisen. Letztendlich weisen jedoch alle Analysen darauf hin, dass nicht nur der Einsatz von neuen besser aufgelösten Datensätzen, sondern auch eine verbesserte Abbildung der im Boden stattfindenden Prozesse wichtig zur grundsätzlichen Verbesserung des Modells ist.

### **Veränderung der Modellqualität anhand ausgewählter Extremereignissen**

Um die verschiedenen Veränderungen, die am Modell durchgeführt wurden, auch in verschiedenen klimatischen Extremsituationen zu evaluieren, werden nachfolgend der „Rekordsommer 2003“

und der relativ kalte und schneereiche Winter 2005/06 genauer analysiert. Neben den jeweiligen mittleren Temperaturen und Niederschlagssummenwerten auf unterschiedlichen Zeitskalen werden hierfür auch klimatische Indizes verwendet. Weit verbreitet sind die sogenannten 27 Expert Team on Climate Change Detection and Indices (ETCCDI), die verschiedene Extremsituationen wie Hitzewellen, Starkniederschlagsereignisse oder verschiedene Perzentilwerte abbilden können. Die auf Tagesdaten basierenden Indizes werden weltweit sowohl für Beobachtungsdaten als auch für Modelldaten eingesetzt (Alexander et al., 2006; Sillmann et al., 2013). In dieser Arbeit werden die in Tabelle 19 beschriebenen Indizes zur Analyse der Extremereignisse verwendet.

Tab. 19: Übersicht und Definition der verwendeten Klimaindizes

Index-Abkürzung	Beschreibung
SU	Sommertag: $T2MAX > 25\text{ }^{\circ}\text{C}$
TR	tropische Nacht: $T2MIN > 20\text{ }^{\circ}\text{C}$
FD	Frosttag: $T2MIN < 0\text{ }^{\circ}\text{C}$
R1mm	Regentag: $RR > 1\text{ mm}$
R20mm	Anzahl der Tage mit mehr als 20mm Regen: $RR > 20\text{ mm}$
CSU	Höchste Anzahl der Tage in einer Sommerperiode, die durch mehr als 5 Tage mit $T2MAX > 25\text{ }^{\circ}\text{C}$ gekennzeichnet ist
CFD	Höchste Anzahl der Tage in einer Frostperiode, die durch mehr als 5 Tage mit $T2MIN < 0\text{ }^{\circ}\text{C}$ gekennzeichnet ist
CDD	Höchste Anzahl der Tage in einer Trockenperiode, die durch mehr als 5 Tage mit $RR < 1\text{ mm}$ gekennzeichnet ist
CWD	Höchste Anzahl der Tage in einer Regenperiode, die durch mehr als 5 Tage mit $RR > 1\text{ mm}$ gekennzeichnet ist

### Hintergrund und Auswirkungen der Hitzewelle 2003

Der als „Jahrhundertsommer“ titulierte Sommer 2003 war der wärmste Sommer der letzten 500 Jahre in Europa und gilt mit ca. 22 000–70 000 Todesopfern und einem wirtschaftlichen Schaden von mehr als 13 Mrd. US-Dollar als eine der schlimmsten Naturkatastrophen überhaupt in Europa. Besonders stark von der Hitzewelle und deren Auswirkungen waren Frankreich und die südeuropäischen Länder wie Italien, Spanien und Portugal betroffen. Als Höhepunkt der sowohl in der Dauer als auch in den Maximaltemperaturen extremen Hitzewelle gelten die Tage vom 01.-13. August. Erst am 14. August wurde seit dem 01. August kein historischer Hitzerekord mehr gebrochen. Wegen des großflächigen und stabilen Hochdruckgebiets „Michaela“ über Mitteleuropa, welches wegen der Anordnung der Druckgebilde Teil einer sogenannten Omegawetterlage war, konnten keine Tiefdruckgebiete die immer weiter steigende Hitze unterbrechen (G. Müller et al., 2003; DWD, 2004; Schär & Jendritzky, 2004; Robine et al., 2008; García-Herrera et al., 2010).

Dass Änderungen an der Bodenparametrisierung und deren Auswirkung auf die Bodenvariablen auch Einfluss auf die Modellierung dieser extremen Hitzeereignisse haben können, zeigen beispielsweise Whan et al. (2015). Sie untersuchen mittels eines Wasserbilanzmodells und verschiedener Beobachtungsdaten den Zusammenhang des Bodenwassergehaltes auf die Maximaltemperaturen

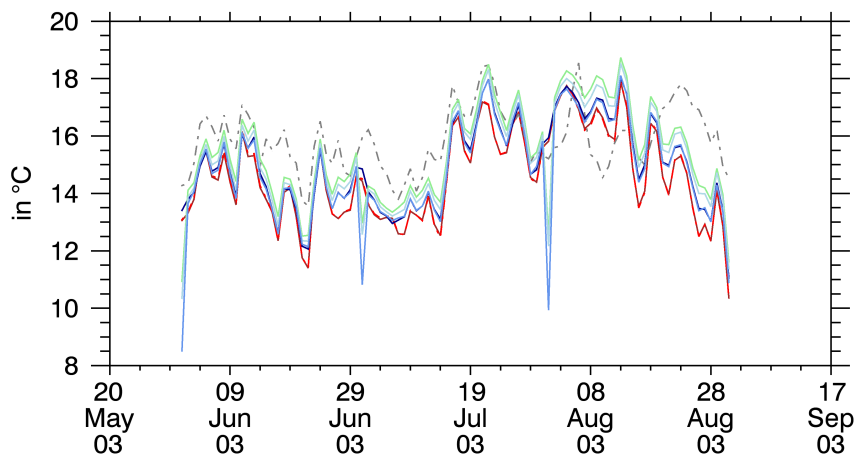
für diese Hitzewelle 2003. Wie sich die beschriebenen Änderungen dieser Arbeit auf die von REMO modellierten Werte für den Sommer 2003 auswirken, wird nachfolgend anhand verschiedener räumlicher und zeitlicher Analysen mit den ERA5 validiert.

### **Modellperformanz für Sommer 2003**

In Abbildung 83 sind die räumlichen Mittelwerte des GER-11-Gebiets sowohl für ERA5-Daten in den beiden Auflösungen ( $0.44^\circ$  und  $0.11^\circ$ ) als auch für die beiden Modellläufe id01 (GER-11 und GER-44) sowie für id06, id15a und id18a für GER-11 als Zeitreihen zu sehen. Die ERA5-Werte sind für beide Auflösungen ( $0.11$  und  $0.44^\circ$ ) in dieser Abbildung dargestellt, weil die Interpolation zu unterschiedlichen Ergebnissen führen kann. Zwar sind bei den Zeitreihen der hier dargestellten Minimumtemperatur und des Niederschlags kaum sichtbare Unterschiede zwischen den beiden Kurven zu erkennen, bei den aus den gleichen Daten abgeleiteten Kennwerten in Abbildung 84 werden dagegen je nach Auflösung und Kenntage unterschiedliche Werte erreicht, welche den Vergleich mit Modelldaten beeinflussen. Bei der Analyse der Zeitreihen fällt zunächst auf, dass sich GER-44 deutlich von allen anderen Zeitreihen unterscheidet. Die Reihe erreicht zumeist höhere Minimumtemperaturen als die GER-11-Läufe und liegt immer über den ERA5-Daten. Der Verlauf stimmt aber bis auf wenige Ausnahmen (z. B. um den 15. Juni oder um den 20. August) relativ gut mit den Validationsdaten überein. Durch die GER-11-Läufe werden die Validationsdaten jedoch noch besser abgebildet. Trotzdem überschätzen alle GER-11-Daten die minimale Temperatur zu nahezu jedem Zeitpunkt. Ferner lässt sich an den meisten Tagen der dargestellten Periode eine Rangfolge der GER-11-Läufe bilden. Der erste Zwischenschrittlauf (id06) modelliert die höchsten Temperaturen, gefolgt vom zweiten Zwischenschrittlauf (id15a) und die beiden anderen Läufe (id01 und id18a) liefern die geringsten Minimaltemperaturen. Auffällig ist, dass Lauf id018a, der im Bezug auf die Datensätze und die Parametrisierungen den größten Unterschied zu id01 aufweist, annähernd direkt auf diesem bisherigen Lauf id01 liegt. Ausnahmen davon können am Beginn der Zeitreihe am 01. Juli und am 01. August ausgemacht werden, die alle durch deutlich geringere Minimumtemperaturen bei Modelllauf id18a gekennzeichnet sind. Bei den beiden Ereignissen im Juli und August fällt zudem auf, dass auch die beiden Zwischenschrittläufe einen großen Temperatursprung nach unten aufweisen, obwohl die ERA5-Daten und auch die bisherige Modellversion id01 in beiden Fällen für beide Tage nur einen leichten Abfall bzw. Anstieg der Minimumtemperatur zeigen. Dieses Merkmal ist weder bei der mittleren noch bei der maximalen Temperatur, sondern nur bei der dargestellten Minimumtemperatur zu beobachten. Beim Niederschlag lässt sich feststellen, dass die größten Unterschiede zwischen den Zeitreihen ebenfalls durch die Modellauflösung entstehen. Obwohl auch bei der GER-44-Kurve der Verlauf, mit der Ausnahme einer Periode Mitte August, relativ gut mit den Validationsdaten übereinstimmt, liegen alle GER-11-Läufe näher an der Validationszeitreihe. Ein Unterschied zwischen den verschiedenen Modellläufen ist kaum zu sehen. Größere Abweichungen zwischen Modell- und Validationsdaten treten vermehrt bei mittleren Niederschlagssummen (2–4 mm) auf. Der für die Minimumtemperatur erwähnte markante Unterschied zwischen Lauf id01 und den anderen Läufen für einen Tag im Juli und einen Tag im August fällt zumindest im Juli mit einem größeren Niederschlagsereignis (v. a. im Alpen und Voralpenland) zusammen. Durch diesen hohen

Wassereintrag und den vorgenommenen Änderungen in der Bodenparametrisierung, die stark vom Wassergehalt im Boden abhängt, und besonders durch das neue Bodentemperaturschema das in Lauf id18a verwendet wird, kann der große Temperaturabfall von fast 4 °C erklärt werden. Die höhere Wärmekapazität durch einen höheren Bodenwassergehalt im Boden führt wegen des Wärmeaustausches zwischen Boden und Atmosphäre und wegen der höheren Evapotranspiration zu geringeren Lufttemperaturen. Beim zweiten Abfall der Minimumtemperatur am 01. August, der eine noch größere Differenz erzeugt, ist im Niederschlag hingegen kein Signal am Tag direkt zu erkennen.

(A)



(B)

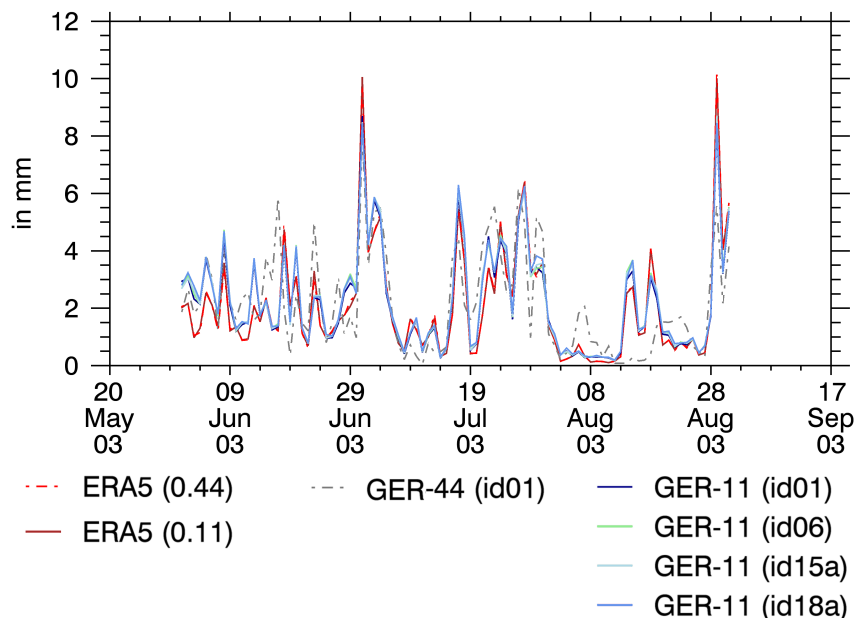


Abb. 83: Vergleich der Zeitreihen des Flächenmittels vom GER-11-Gebiet täglicher Werte von Juni bis August 2003 für die verschiedenen Warmstartläufe (Parametrisierungsunterschiede vgl. Kapitel 6.2) und ERA5

- (A) Minimumtemperatur (T2MIN)
- (B) Niederschlag (RR)



Vergleicht man alle längeren Warmstartläufe mit den ERA5-Validationsdaten für verschiedene klimatische Wärme- und Trocken-Indizes einerseits für die drei Sommermonate (JJA) von 2003 und andererseits für den Höhepunkt der Hitzewelle (01.-13. August), erhält man Abbildung 84. Wie bereits bei den Zeitreihen der Flächenmittel deutlich wurde, ist zu erkennen, dass der größte Unterschied zwischen GER-44 und allen anderen Läufen und Validationsdaten vorliegt. Bei der Betrachtung des ganzen Sommers modelliert GER-44 in allen dargestellten Indizes (mit Ausnahme der Regentage) zu viele Tage (Abb. 84A), in der Periode Anfang August hingegen zu wenige Tage (Abb. 84B). Unter Zuhilfenahme der Zeitreihe der Minimaltemperatur (Abb. 83A) erklärt sich dieser Aspekt, denn bis auf Anfang August liegt die Temperatur für GER-44 immer über den Temperaturen von allen GER-11-Läufen. Bei der Anzahl der Regentage (RR1mm) ist zwischen allen Modellläufen in Abbildung 84A nahezu kein Unterschied zu erkennen. Alle Läufe unterschätzen die Regentage um ca. fünf Tage im Vergleich zu ERA5 für Juni bis August 2003. Bei Abbildung 84B dagegen überschätzt nur GER-44 die Anzahl der Regentage, während alle GER-11-Läufe die ERA5-Regentage nahezu perfekt abbilden. Der größte Unterschied zwischen den GER-11-Modellläufen wird bei der Anzahl der Tropennächte (TR) erreicht. In beiden Zeiträumen überschätzt GER-11-Lauf id06 die Anzahl am meisten, gefolgt von Lauf id15a, der im Vergleich zu ERA5 auch zu hohe Werte liefert. Die Anzahl der Tropennächte von Modelllauf id18a liegt zwar ebenfalls über der Anzahl der Validationsdaten, die Differenz zwischen id18a und ERA5 und die Differenz zwischen Lauf id01, der die Anzahl unterschätzt, und ERA5 ist aber annähernd gleich. Insgesamt liegen die Werte für alle Indizes und beide Zeitperioden sehr nahe beieinander. Auffällig ist außerdem, dass obwohl in beiden Zeiträumen die Anzahl der Sommertage (SU) sehr hoch ist, die höchste Anzahl der Tage in einer Sommerperiode (CSU) sehr klein ist. Dies liegt daran, dass diese durch mehr als fünf Tage mit  $T2MAX > 25\text{ °C}$  gekennzeichnet ist (vgl. Tab. 19) und bereits ein Tag mit einer Maximaltemperatur von unter  $25\text{ °C}$  die kumulierte Häufigkeit unterbricht, sodass erst wieder fünf Tage höhere Temperaturen erreichen müssen, bevor die Zählung erneut beginnt.

Der Unterschied zwischen den beiden ERA5-Daten kommt durch die Interpolation der vier Ausgangsvariablen auf die beiden Zielaufösungen  $0.44^\circ$  und  $0.11^\circ$  zustande. Trotz der Tatsache, dass beide aus dem gleichen originalen ERA5-Datensatz mit  $0.25^\circ$  erzeugt wurden, unterschieden sie sich so stark, dass es vor allem bei den Klimaindizes, die auf Grundlage von Grenzwerten erstellt werden, aber auch vereinzelt bei den Zeitreihen in den dargestellten Werten zu erkennen ist. Auf dieses „Problem“ wird in Kapitel 7 noch einmal separat eingegangen.

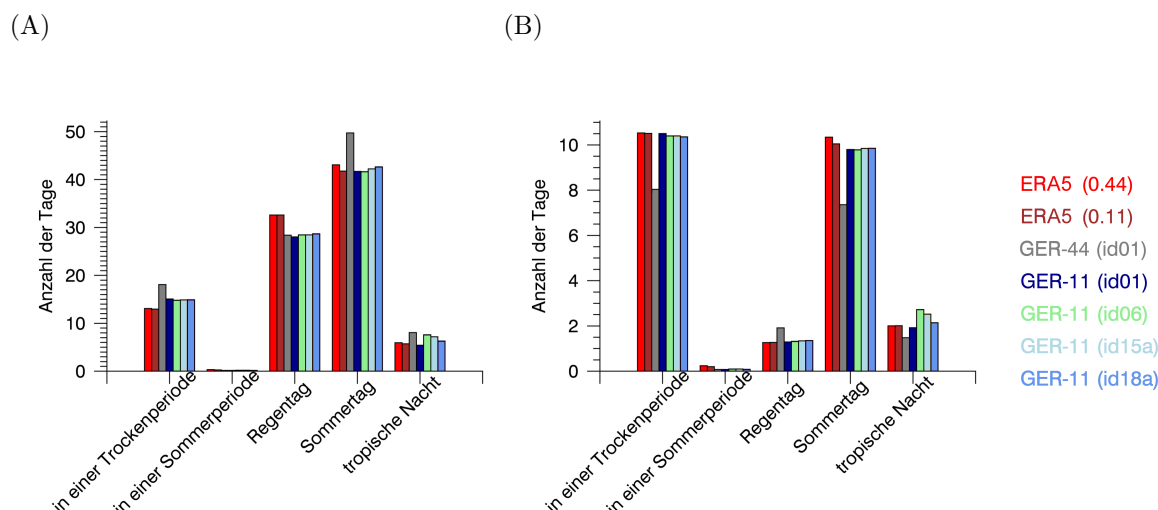


Abb. 84: Vergleich zwischen Validationsdaten und Modellläufen (Parametrisierungsunterschiede vgl. Kapitel 6.2) anhand verschiedener Kenntage in Deutschland für den Sommer 2003 (Die GER-44-Daten sind die auf das GER-11-Gebiet ausgeschnittenen EUR-44-Daten, sodass sich alle dargestellten Zahlen auf das GER-11-Gebiet beziehen.)

(A) Juni 2003 bis August 2003

(B) 01. August bis 18. August 2003

Nach der Analyse der Flächenmittel wird auch eine Analyse der zeitlichen Mittelwerte durchgeführt. Abbildung 85 zeigt zum einen die mittlere Minimumtemperatur von Juni bis August 2003 für das GER-11-Gebiet für ERA5 in  $0.11^\circ$  und  $0.44^\circ$  Auflösung (Abb. 85A und Abb. 85C) und zum anderen die jeweilige Differenz zum korrespondierenden Gebiet vom Warmstart des id01-Modelllaufs von EUR-44 sowie die jeweiligen Differenzen zu den Warmstarts von GER-11. Wie bereits bei der Zeitreihendarstellung, die sich durch die räumliche Mittelung dieser Daten ergeben hat, zu erkennen war (vgl. Abb. 83A), überschätzt der Modelllauf id01 von EUR-44 die Minimumtemperatur annähernd im gesamten Gebiet um  $1\text{--}2^\circ\text{C}$  und damit mehr als alle vier GER-11-Läufe. Dennoch sind die räumlichen Muster von EUR-44- und GER-11-Lauf id01 bis auf den Betrag der Differenz sehr ähnlich. Markant ist z. B. das Differenzmuster in Polen, das sich durch eine höhere Differenz zu ERA5 abhebt und welches sich durch die Bodenkarte bzw. die bisherige Bodenparametrisierung (vgl. Muster in Abb. 55B und Abb. 57B) ergibt.

Vergleicht man die verschiedenen GER-11-Läufe miteinander, ist festzustellen, dass mit Modelllauf id06 (Abb. 85D) besonders im Norden Deutschlands und in Polen großflächig höhere Differenzen erzeugt werden, weil hier die Minimumtemperatur noch stärker überschätzt wird als mit dem bisherigen Lauf id01 (Abb. 85B). Auch die zuvor unterschätzten Regionen, wie der abgebildete Teil Frankreichs und das bayerische Voralpenland, die sich in den Validationsdaten durch höhere Minimaltemperaturen von der Umgebung abgrenzen (vgl. Abb. 85A), erreichen in GER-11-id06 teilweise eine positive Differenz durch diese Erwärmung im Vergleich zu Lauf id01. Der zweite Zwischenschrittlauf id15a ähnelt dagegen mehr Lauf id01 als Lauf id06, obwohl beide

neuen Läufe beispielsweise Soilgrid statt FAO als Bodendatengrundlage verwenden. Betrachtet man das Differenzmuster von Lauf id18a (Abb. 85G), stellt man fest, dass sich, anders als die vorherige Zeitreihen- und Indizesanalyse suggeriert haben, die beiden Läufe id01 und id18a durchaus unterscheiden. Besonders im nördlichen Teil Deutschlands, in dem unter anderem die neue Bodengrundlage (vgl. Abb. 50) einfließt, die zwar bei den anderen beiden Läufen ebenfalls verwendet wird, aber für Lauf id18a in Kombination mit allen Parametrisierungsänderungen und dem neuen Temperaturschema größere Änderungen hervorruft, dreht sich das Vorzeichen der Differenz zu den Validationsdaten sogar um. Im Alpenbereich hingegen werden (wie bereits bei der Analyse der längeren Zeiträume) keine signifikanten Änderungen durch die neuen Modellversionen erreicht. Dass sich vor allem das Differenzmuster von Abb. 85G nicht deutlicher von den anderen unterscheidet, obwohl hier die ASTER-Topographie statt der GTOPO-Topographie verwendet wird, liegt daran, dass jeweils für beide Topographien Höhenkorrekturen für die ERA5-Daten vorgenommen wurde (vgl. Kapitel 4.5.2).

In den nicht dargestellten Karten der mittleren Temperatur für Juni bis August 2003 ist dieser Unterschied zwischen den vier GER-11-Versionen ebenfalls zu erkennen. Bei den mittleren Maximumtemperaturen lässt er sich dagegen nicht nachweisen. Beim Niederschlag ist nur der sich in den Zeitreihen bereits angedeutete Unterschied zwischen den beiden Auflösungen erkennbar (im Großteil des Gebiets weniger Überschätzung durch  $0.11^\circ$  als durch  $0.44^\circ$ ), die vier Differenzmuster der GER-11-Läufe zeigen keinen signifikanten Unterschied.

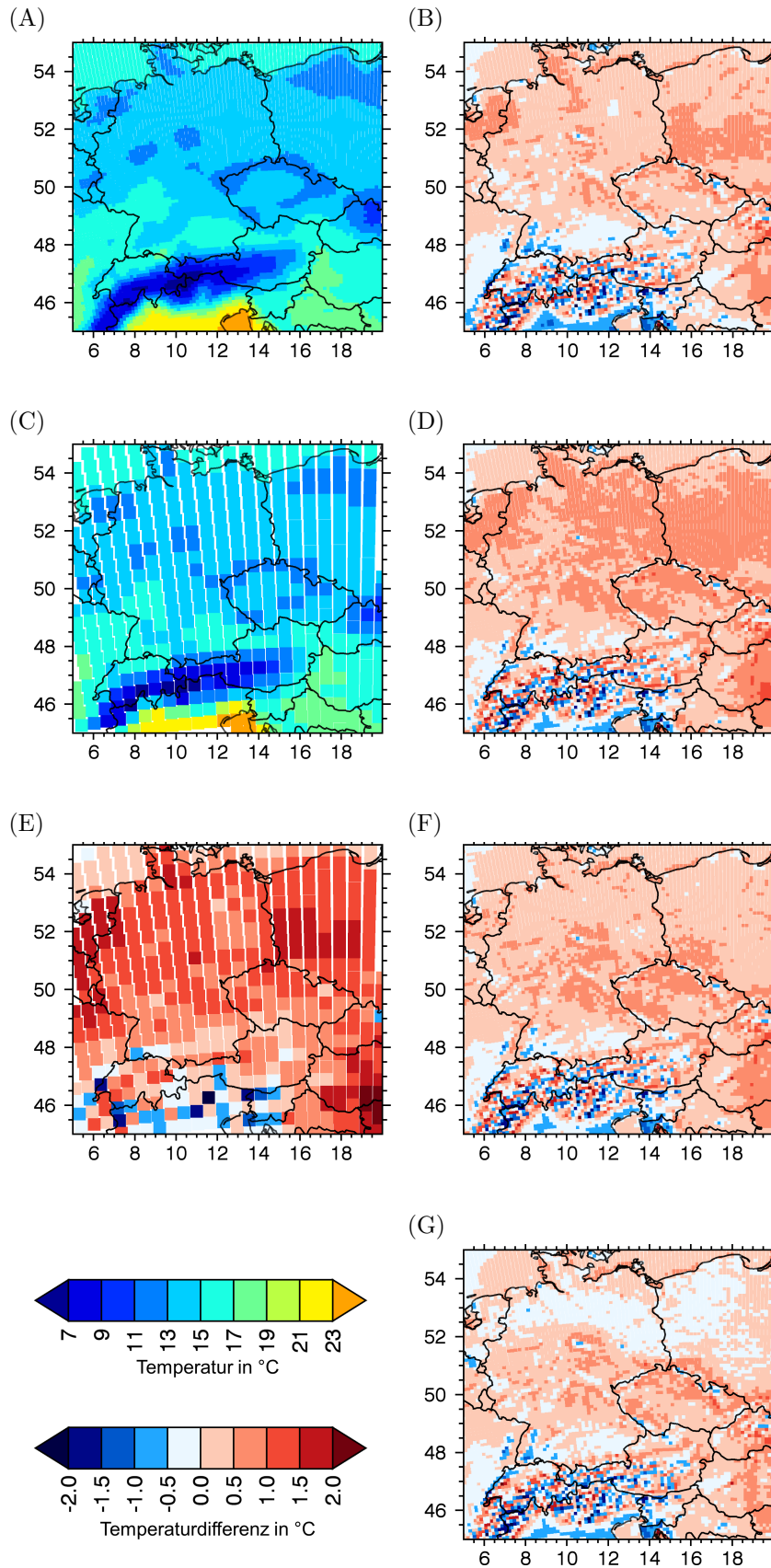


Abb. 85: Räumlicher Vergleich der mittleren Minimumtemperatur zwischen Validationsdaten und Modellläufen (Parametrisierungsunterschiede vgl. Kapitel 6.2) für Juni bis August 2003  
 (A) ERA5 (0.11) / (B) GER-11 (id01) - ERA5 (0.11) / (C) ERA5 (0.44) /  
 (D) GER-11 (id06) - ERA5 (0.11) / (E) EUR-44 (id01) - ERA5 (0.44) /  
 (F) GER-11 (id15a) - ERA5 (0.11) / (G) GER-11 (id18a) - ERA5 (0.11)

### Hintergrund und Auswirkungen des Winters 2005/06

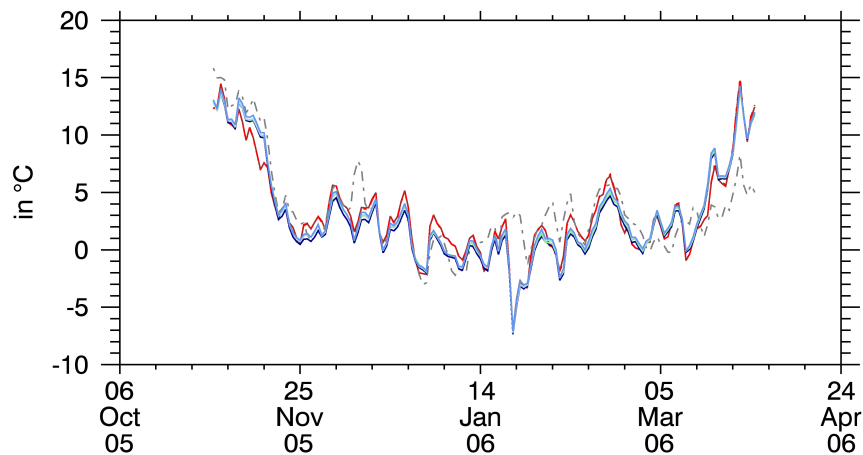
Der Winter 2005/06 war ein ungewöhnlich langer und vor allem im Osten und Süden Deutschlands strenger Winter. Schon Mitte November bildeten sich die ersten flächendeckenden geschlossenen Schneedecken, die aufgrund von fehlenden Tauphasen über den Winter immer dicker wurden. Alle Wintermonate - besonders der Januar und März 2006 - waren deutschlandweit zu kalt im Vergleich zur Klimanormalperiode. Vor allem Anfang Januar und Anfang Februar traten zudem Perioden mit sehr viel Neuschnee innerhalb kurzer Zeit auf. Die große Schneelast führte beispielsweise Ende November 2005 zum Einsturz vieler Strommasten in Münster und Anfang 2006 zu hoher Einsturzgefahr von Gebäuden in Bayern. Zudem war es selbst Anfang März 2006 noch kalt, sodass es am 04. März vom Main bis ins Alpenvorland und am 12. März in Norddeutschland aufgrund von Schneefall zu Verkehrsbehinderungen auf den Straßen kam. Trotzdem war der Winter 2005/06 insgesamt kein außergewöhnlich nasser, sondern in den tieferen Lagen zu trockener und außerdem zu sonniger Winter (DWD, 2006; Deuschländer & Wichura, 2006; Wetter Online, 2006; DWD, 2007). Wie REMO diesen Winter nachbilden kann und wie sich die vorgenommenen Änderungen auf die Temperaturen, den Niederschlag und andere Kenntage auswirken, wird in den folgenden Unterkapitel analysiert.

### Modellperformanz für Winter 2005/06

Abbildung 86 zeigt die Flächenmittel der täglichen Maximaltemperatur (Abb. 86A) und der täglichen Niederschlagssumme (Abb. 86B) des GER-11-Gebiets für die beiden ERA5-Auflösungen sowie für alle Warmstartläufe von November 2005 bis März 2006. Analog zu Abbildung 83 handelt es sich bei dem GER-44-Lauf um die aus dem EUR-44-Lauf id01 ausgeschnittenen Daten, sodass sich alle dargestellten Werte auf das gleiche Gebiet (GER) beziehen.

Alle REMO-Versionen stimmen für die Maximaltemperatur sehr gut mit den ERA5-Daten überein (Abb. 86A). Sowohl die Werte als auch der Verlauf der Modell-Zeitreihen liegen für alle Versionen eng bei den Validationsdaten. Die grob aufgelöste Modellversion GER-44-id01 zeigt aber größere Differenzen als die GER-11-Läufe. Anfang Dezember 2005 ist ferner für GER-44-id01 zu sehen, dass die Maximaltemperatur ansteigt, während sie bei allen andern Zeitreihen fällt. Zudem wird bei der gröber aufgelösten Modellversion der März 2006 deutlich zu kalt modelliert. Vergleicht man die verschiedenen GER-11-Modellläufe miteinander, sind nur wenige Unterschiede zu erkennen. Durch die leichte aber konstante Erwärmung liegt Lauf id18a von Ende November 2005 bis Anfang März 2006 näher an den Validationsdaten als die anderen Läufe, da die Maximumtemperatur in diesem Zeitraum bisher unterschätzt wurde. Zu Beginn und zum Ende der Zeitreihe dagegen schneidet dieser Lauf entweder etwas schlechter oder genauso gut wie die anderen GER-11-Läufe ab. Auch beim Niederschlag ähneln sich die Zeitreihen (vgl. Abb. 86B). Während sich beim GER-44-id01 Lauf jedoch besonders Ende Dezember 2005 sowie im Januar und März 2006 größere positive Abweichungen von den Validationsdaten beobachten lassen, liegen alle GER-11-Modellversionen sehr eng an den ERA5-Kurven bzw. unterschätzen diese leicht. Generell ist kaum ein Unterschied zwischen den GER-11-Läufen zu erkennen.

(A)



(B)

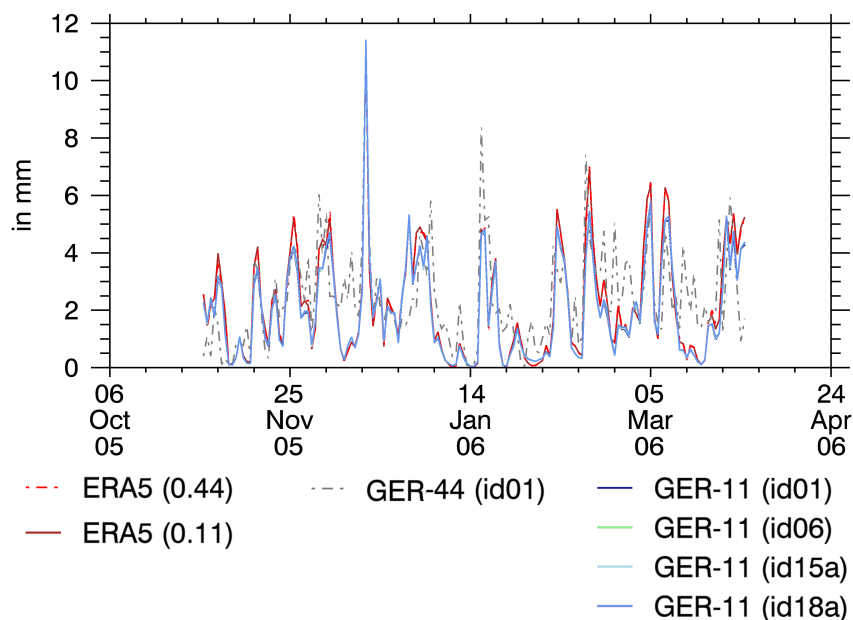


Abb. 86: Vergleich der Zeitreihen des Flächenmittels vom GER-11-Gebiet täglicher Werte von November 2005 bis März 2006 für die verschiedenen Warmstartläufe (Parametrisierungsunterschiede vgl. Kapitel 6.2) und ERA5  
 (A) Maximumtemperatur (T2MAX)  
 (B) Niederschlag (RR)

Die unterschiedlichen Ergebnisse der Modellläufe und der Validationsdaten für verschiedene Kälte- und Niederschlagsindizes für den Winter 2005/06 sind in Abbildung 87 dargestellt. Bei der Variable RR1mm, ist zu erkennen, dass wie bei Abbildung 84B und anders als Abbildung 84A der gröber aufgelöste Modelllauf (GER-44 id01) die Anzahl der Regentage überschätzt, während alle GER-11-Läufe die Validationsdaten deutlich unterschätzen. Der negative Bias aller GER-11-Läufe ist dabei ca. dreimal so groß wie der positive Bias von GER-44. Betrachtet man die höchste Anzahl der Tage in einer Regenperiode (CWD) (Definition siehe Tab. 19) ist diese Abstufung ebenfalls zu erkennen, wengleich die Absolutwerte mit ca. zehn Tagen deutlich geringer sind als

bei RR1mm. Bei der Anzahl der Starkregentage (RR20mm), die im Winter weniger als fünf Tage beträgt, ist in dem dargestellten Flächenmittel kein Unterschied zwischen den Modellversionen zu erkennen und alle Läufe überschätzen die Validationsdaten leicht.

Bei der Betrachtung der thermischen Indizes fällt ferner auf, dass die GER-44-Daten zwar die Anzahl der Frosttage (FD) (vgl. Definition in Tab. 19) im Vergleich zu den Validationsdaten deutlich unterschätzen, dass jedoch dieser Modelllauf die höchste Anzahl der Tage in einer Frostperiode (CFD) am besten abbildet. Umgekehrt überschätzen alle GER-11-Modellversionen die maximale Anzahl an kumulierten Frosttagen um bis zu 18 Tage, der tatsächliche FR-Wert wird dagegen sehr gut modelliert. Ebenso lassen sich bei den Frosttagen, die von der jeweiligen Temperatur und nicht vom Niederschlag abhängen, die größten Unterschiede zwischen den verschiedenen Modellversionen nachweisen, was bereits bei der Analyse des Hitzesommers 2003 (Kapitel 6.5.2.2) festgestellt werden konnte. Sowohl bei CFD als auch bei FD schneidet der finale GER-11-Modelllauf id18a sehr gut im Vergleich zu ERA5 ab.

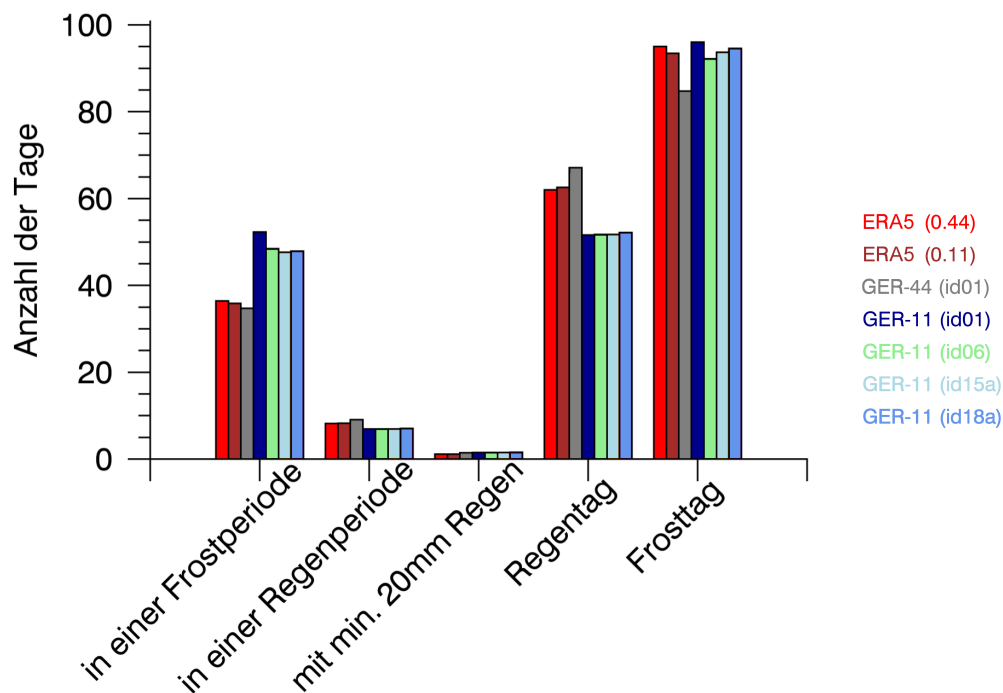


Abb. 87: Vergleich zwischen Validationsdaten und Modellläufen (Parametrisierungsunterschiede vgl. Kapitel 6.2) anhand verschiedener Kenntage in Deutschland für den Winter 2005/2006. (Zeitraum: November 2005 bis März 2006) (Die EUR-44-Daten sind auf das GER-11-Gebiet geschnitten, sodass sich alle dargestellten Zahlen auf das GER-11-Gebiet beziehen.)

In Abbildung 88 ist die mittlere Maximaltemperatur von November 2005 bis März 2006 sowie die jeweiligen Differenzen der fünf verschiedenen Modellläufe dargestellt. Analog dazu ist in Abbildung 89 die Anzahl der Frosttage, welche auf Basis der Minimumtemperatur berechnet

werden, im gleichen Zeitraum abgebildet. Deutlich zu sehen ist die Zunahme der Frosttage mit der Zunahme der Kontinentalität sowie der Einfluss der Topographie. In den Alpen und dem abgebildeten Teil der Karpaten wurden mehr als 135 der insgesamt 150 Tage des ausgewählten Zeitraums als Frosttag klassifiziert. Die Maximaltemperatur erreicht im Winter 2005/2006 einen Durchschnittswert von ca. 7–11 °C.

Es fällt auf, dass sich die Differenzmuster des gröber aufgelöste GER-44-Laufs (Abb. 88C und Abb. 89C) bei beiden Variablen deutlich von den jeweiligen Differenzen der GER-11-Läufe unterscheiden. Bei den Frosttagen lässt sich eine räumliche Korrelation zwischen den tatsächlichen Werten und den Differenzen erkennen. Eine kleine Anzahl an Frosttagen, die vor allem im südlichen und westlichen Teil des Untersuchungsgebiets zu finden ist, wird vom Modell deutlich unterschätzt. Mittlere Werte für Frosttage sind sehr gut abgebildet, eine höhere Anzahl an FD (z. B. Tschechien und Slowakei) wird aber überschätzt (Abb. 89C). Dieses Differenzmuster unterscheidet sich signifikant von den Mustern der GER-11-Versionen, bei denen eine annähernd flächendeckende Überschätzung erfolgt. Die maximale Temperatur wird vom GER-44-Lauf (Abb. 88C) in der Fläche im Vergleich zu den Validationsdaten größtenteils überschätzt. Besonders im nordwestlichen Teil Italiens und vereinzelt in der Region, wo sich die Grenzen von Österreich, Ungarn und Slowenien treffen, werden Abweichungen von 0.5–1.5 °C erreicht. Ein großflächiger markanter negativer Bias (–0.5 bis –1.5 °C) ergibt sich in Zentralpolen und der Slowakei sowie an vereinzelt Stellen in den Alpen. Für die GER-11-Modellversionen ist der Bias für die dargestellte Landmasse überwiegend negativ. Zu den beim GER-44-Lauf beschriebenen negativen Differenzmustern für Polen und der Slowakei, von denen das Muster von GER-11-id01-Lauf etwas nach Westen und Süden verschoben ist, kommen weitere großflächige Regionen wie Nordtschechien, Bereiche westlich von Berlin sowie Bereiche um München und Nürnberg dazu, welche ebenfalls einen negativen Bias von –0.5 bis –1 °C haben. Bei Betrachtung der verschiedenen GER-11-Versionen fällt auf, dass sich diese Flächen mit negativen Bias sukzessiv verkleinern und die neuen REMO-Läufe diese Bereiche besser modellieren, was an der generellen Erwärmung durch die Modelländerungen erfolgt.

Für die Minimumtemperatur scheint diese Erwärmung ebenfalls - jedoch in anderen Regionen - zu gelten, da die Anzahl der Frosttage, die vom bisherigen Modelllauf GER-11-id01 vor allem im Zentrum des Untersuchungsgebiets deutlich überschätzt wurde, in allen neuen Modellversionen etwas reduziert ist, sodass alle positiven Differenzen etwas besser modelliert werden. Bei den Flächen mit zuvor negativen Bias ist dagegen keine Verbesserung, aber auch keine Verschlechterung zu erkennen. In den nicht dargestellten Niederschlagsmustern sind, wie bei der Hitzesommer-Analyse bereits festgestellt, keine signifikanten räumlichen Änderungen zwischen den Modellläufen zu erkennen.



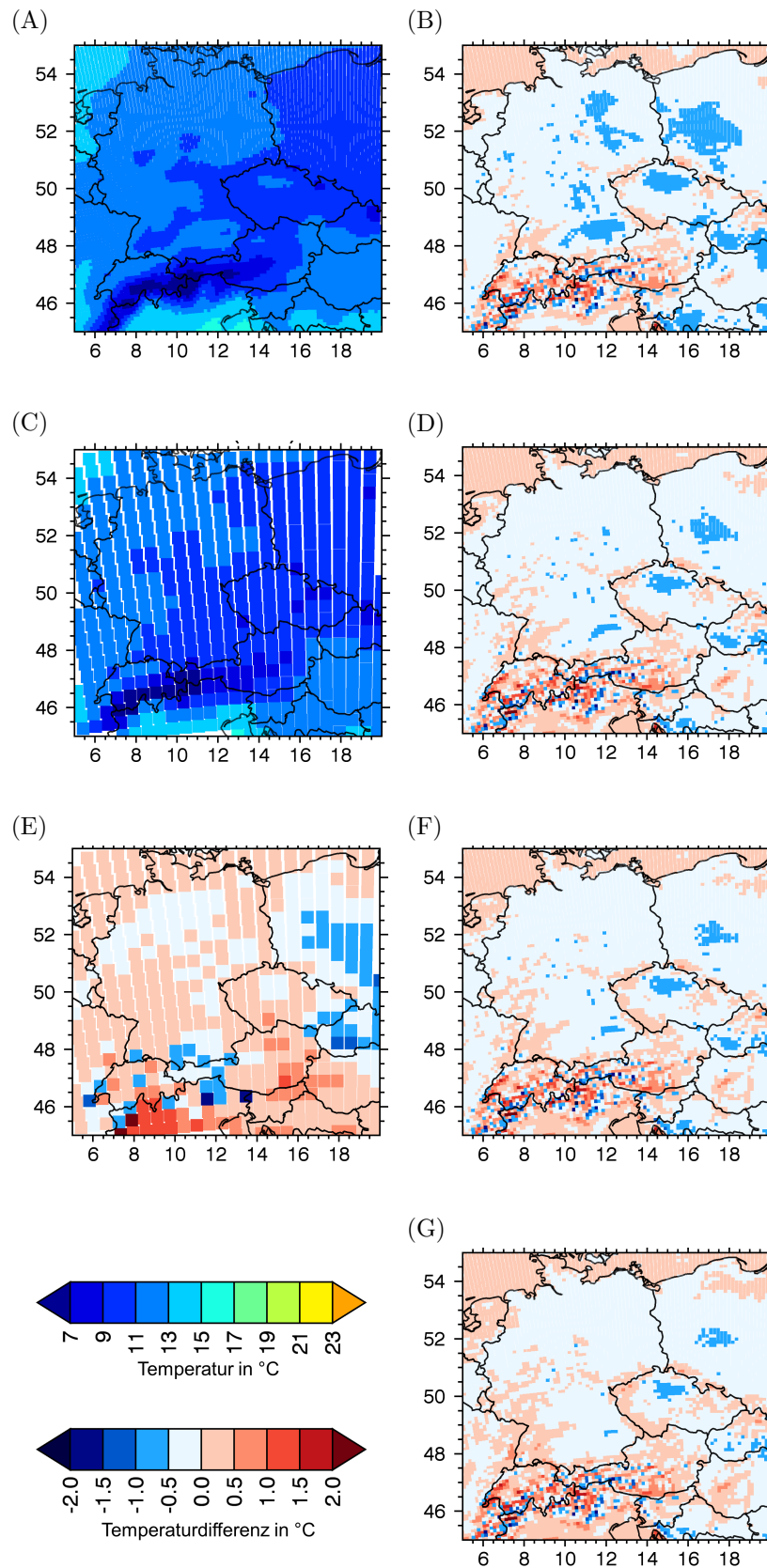


Abb. 88: Räumlicher Vergleich der mittleren Maximumtemperatur zwischen Validationsdaten und Modellläufen (Parametrisierungsunterschiede vgl. Kapitel 6.2) für November 2005 bis März 2006

(A) ERA5 (0.11) / (B) GER-11 (id01) - ERA5 (0.11) / (C) ERA5 (0.44) /  
 (D) GER-11 (id06) - ERA5 (0.11) / (E) EUR-44 (id01) - ERA5 (0.44) /  
 (F) GER-11 (id15a) - ERA5 (0.11) / (G) GER-11 (id18a) - ERA5 (0.11)

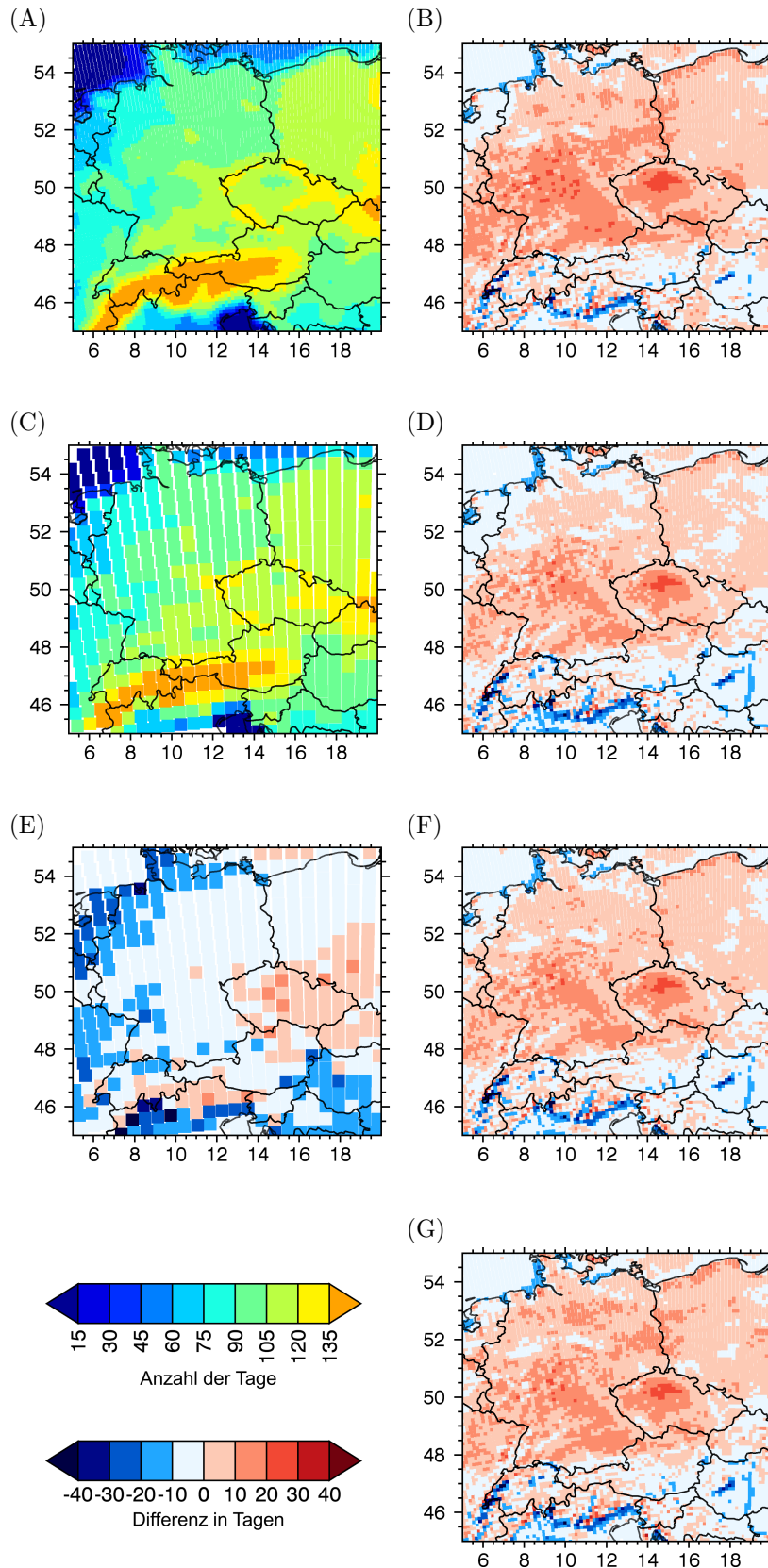


Abb. 89: Räumlicher Vergleich der Anzahl der Frosttage zwischen Validationsdaten und Modellläufen (Parametrisierungsunterschiede vgl. Kapitel 6.2) für November 2005 bis März 2006

(A) ERA5 (0.11) / (B) GER-11 (id01) - ERA5 (0.11) / (C) ERA5 (0.44) /  
 (D) GER-11 (id06) - ERA5 (0.11) / (E) EUR-44 (id01) - ERA5 (0.44) /  
 (F) GER-11 (id15a) - ERA5 (0.11) / (G) GER-11 (id18a) - ERA5 (0.11)

### **Zusammenfassung der Modellperformanz für die Extremereignisse**

In den dargestellten Beispielen zweier Extremereignisse (Hitzewelle 2003, kalter und langer Winter 2005/06) konnte gezeigt werden, dass sich die verschiedenen Änderungen in den Modellläufen nicht so stark auswirken, wie die räumlichen und betragsmäßigen Änderungen der Eingangsdaten und die neuen Parametrisierungen vermuten lassen würden, was mit den zuvor durchgeführten Analysen einhergeht. Die größte Verbesserung im Vergleich zu den Validationsdaten wird durch die Erhöhung der Auflösung von 0.44 auf 0.11° erreicht. Des Weiteren liefern die durchgeführten Analysen die Erkenntnis, dass Unterschiede zwischen verschiedenen Läufen teilweise durch räumliche oder zeitliche Mittelung der Ausgangsdaten überdeckt werden.

Es konnte festgestellt werden, dass in manchen Variablen und zu bestimmten Zeiträumen die Karte der bisherigen Bodendaten deutlich im räumlichen Differenzmuster nachweisbar ist (z. B. Minimumtemperatur im Hitzesommer 2003, Abb. 85B und Abb. 85E), was durch die neu eingeführten Bodendaten verbessert werden konnte. Zudem führt Soilgrid besonders in Kombination mit dem neuen Bodentemperaturschema zu einer flächendeckenden Abkühlung der Minimumtemperatur im Sommer sowie Erwärmung der Maximumtemperatur im Winter, was größere räumliche Verbesserungen hinsichtlich der jeweiligen Variablen mit sich bringt, da REMO bisher die Temperaturamplitude unterschätzt.

Dennoch geht aus den Analysen hervor, dass sich der bisherige Modelllauf id01 und der Modelllauf id18a, in dem alle in dieser Arbeit vorgenommenen Änderungen aktiv sind, sehr ähneln. Besonders deutlich ist dies anhand der Zeitreihenanalyse und im Vergleich zu den beiden Zwischenschrittläufen geworden, die in den allermeisten Fällen größere Differenzen zur Modellversion id01 hatten.

Zudem wirkten sich alle Änderungen hauptsächlich in den Temperaturvariablen und kaum im Niederschlag aus, obwohl einerseits die teilweise großen topographischen Differenzen und andererseits die Bodenwasser-Atmosphäre-Interaktion auch größere Änderungen im Niederschlagssignal vermuten hätten lassen. Insgesamt werden jedoch beide Extremereignisse besonders mit der höheren Auflösung sehr gut im Vergleich zu den Validationsdaten abgebildet.

# 7 Diskussion der Änderungen und Ergebnisse

## 7.1 Probleme beim Vergleich zwischen Validations- und Modelldaten

Bevor näher auf die Ergebnisse der neuen Daten und Prozesse für REMO eingegangen werden kann, müssen die aufgezeigten Unterschiede zwischen den verschiedenen Validationsdaten diskutiert werden. Durch den Fortschritt in Forschung und Technik können Klimamodelle in immer höheren Auflösungen betrieben werden. Gleichzeitig sind dafür jedoch immer weniger Messdaten zur Validation verfügbar. Die tatsächlichen Beobachtungsdaten, die an Messstationen erhoben werden, werden durch unterschiedliche physikalische und statistische Interpolationsmethoden in Rasterdaten umgewandelt. Diese stellen somit keine realen in situ Messungen mehr dar. Die dadurch entstehenden Fehler und Ungenauigkeiten der Rasterdaten sind einerseits von der statistischen Methode und andererseits von der räumlichen Verteilung und Dichte der Messungen abhängig. Viele Studien vergleichen deshalb verschiedene Validationsdatensätze mit Stationsdaten und unterschiedliche Datensätze untereinander, um systematische Fehler und Schwächen einzelner Datensätze aufzudecken.

Beispielsweise wies L. Gao (2013) bei seinem Vergleich zwischen den ERA-Interim- und den E-OBS-Daten für die Zentralalpen nach, dass der großräumige Fehler bei  $0.6\text{ }^{\circ}\text{C}$  liegt, besonders aber die Extremtemperaturen von ERA-Interim schlechter abgebildet werden. Zudem stellte er fest, dass zwar die Standardabweichung des Niederschlags bei beiden Datensätzen gut für die Region übereinstimmt, der tägliche mittlere absolute Fehler aber trotzdem mit  $4.5\text{--}9.5\text{ mm}$  durchaus kritisch zu betrachten ist. Ein weiteres Beispiel für die Wichtigkeit verschiedene Datengrundlagen für Analysen zu verwenden und Datensätze paarweise und auf unterschiedliche Weise zu vergleichen, liefern Dascalu et al. (2020). Sie fanden bei der Zeitreihendarstellung (1995-2017) der räumlichen Korrelationskoeffizienten für Europa, welche dem Vergleich von aus COSMO-REA6 und UERRA zu E-OBS abgeleiteten Hitzeindizes diente, ein plötzlichen Niveausprung von ca. 0.94 auf 0.86 im Jahr 2005, welcher auf die E-OBS-Daten zurückzuführen ist, da die Korrelationskurve der beiden anderen Datensätze keine Verringerung anzeigte. Bei anderen Variablen und weiteren saisonalen Analysen oder in der räumlichen Kartendarstellung fiel dieser Sprung nicht auf. Auch Kyselý & Plavcová (2010), welche E-OBS (V2.0) mit Stationsdaten in Tschechien verglichen, konnten zeigen, dass besonders die Minimumtemperatur von E-OBS einen deutlichen Bias aufweist. Sie fordern, dass die Stationsdichte und die Qualität der Daten, die zur Erzeugung der flächendeckenden Rasterdatensätze verwendet werden, verbessert werden muss, damit diese zur Validation von Klimamodellen eingesetzt werden können.

Speziell in Hinblick auf den Niederschlag sind die in dieser Arbeit nachgewiesenen großen Unterschiede zwischen Modell- und Validierungsdaten unter dem Gesichtspunkt zu betrachten, dass die verwendeten Validationsdaten fehlerhaft sein können. Q. Sun et al. (2018) stellen beispielsweise beim Vergleich verschiedener globaler Niederschlagsdatensätze fest, dass es Unterschiede von bis zu 300 mm im Jahresniederschlag gibt. Sie merken auch an, dass besonders Reanalysedaten

die größten Unsicherheiten erreichen, wobei für Europa (besonders im gebirgigen Süden) die ECMWF-Reanalysedaten, die teilweise in dieser Arbeit ebenfalls als Validierungsdaten eingesetzt werden, schlechter abschneiden als andere Niederschlagsdatensätze. Dies zeigt, dass bereits der Vergleich verschiedener Validationsdaten zu unterschiedlichen Ergebnissen führen kann. Durch diesen Umstand wird die Validation von Klimamodellen generell erschwert, da die Auswahl des Untersuchungsgebietes oder auch die Auswahl eines Validationsdatensatzes die Beurteilung von Modelländerungen beeinflussen und verfälschen kann.

Neben dem Gebrauch unterschiedlicher Datensätze muss auch der Einfluss der Aufbereitung der Beobachtungsdaten beachtet werden. Damit Modell- und Validationsdaten räumlich miteinander verglichen werden können, müssen sie mittels einer Interpolation auf das gleiche Koordinatengitter gebracht werden. Hierfür stehen eine Vielzahl unterschiedlicher Methoden zur Auswahl. Je kleiner der Unterschied zwischen der räumlichen Auflösung der Modell- und Validationsdaten ist, desto geringer fallen auch die Unsicherheiten, die durch die Interpolation entstehen, aus. Auch die Prozessierungsreihenfolge beeinflusst die Validationswerte. Besonders bei einer Weiterverwendung der interpolierten Daten zur Ableitung anderer Größen, wie es beispielsweise bei den Extremklimaindizes der Fall ist, die von festgelegten Grenzwerten abhängig sind, spielt die Reihenfolge eine wichtige Rolle. Diaconescu et al. (2015), welche die Auswirkungen der Reihenfolge von Interpolation und Kennwertbildung für Niederschlagsdaten in Afrika untersuchten, benennen die beiden möglichen Fälle als „First-Step-“ und „Last-Step-“Procedure. Während bei der First-Step-Methode zuerst die Eingangsdaten auf das Zielraster interpoliert werden und anschließend alle Kennwerte gebildet werden, werden bei der Last-Step-Methode erst die Kennwerte auf dem Originalraster berechnet und anschließend diese Werte auf das Zielraster interpoliert. Bei der Analyse von Diaconescu et al. (2015) kann dadurch zwischen den Methoden ein Unterschied von  $5.2 \frac{\text{mm}}{\text{d}}$  beim Niederschlag nachgewiesen werden. Dabei erreicht die First-Step-Methode höhere Differenzen und Standardabweichungen unabhängig von der verwendeten Interpolationsmethode. Die Autoren weisen außerdem darauf hin, dass es bisher kein einheitliches Vorgehen gibt und dass in manchen Studien die Prozessierungsreihenfolge nicht einmal erwähnt wird. In dieser Arbeit wurde die First-Step-Methode verwendet, da für die verwendeten Klimaindizes „ganze Tage“ als Einheit verwendet werden sollte, um hierdurch die kleinen Unterschiede zwischen den Modellversionen besser darzustellen. Die Auswirkung der durch diese Methode erzeugten Werte kann besonders gut anhand der Balkendiagramme in Kapitel 6.5.2.2 veranschaulicht werden. Durch die Interpolation werden die originalen ERA5-Daten verändert und die anschließende Klimaindexbildung führt nach einer räumlichen Mittelung für die beiden unterschiedlichen Auflösungen nicht mehr zum gleichen Ergebnis. Um die Entscheidung der veränderten Modellperformanz unabhängiger von einem ausgewählten Beobachtungsdatensatz zu machen, wurden in der Arbeit mehrere Validationsdatensätze zum Vergleich mit den Modelldaten berücksichtigt. Durch die Tatsache, dass die unterschiedlichen Validationsdaten auf Basis von unterschiedlichen Höhendaten aus Stationsdaten interpoliert wurden und dass auch für das Modell verschiedene Topographiekarten als Grundlage verwendet wurden, was die Analyseergebnisse der Temperaturvariablen stark beeinflusst, wurde eine Höhenkorrektur mit einer konstanten Lapse Rate durchgeführt (vgl. Kapitel 4.5.2). Obwohl dies ein gängiges Verfahren in der Handhabung von Temperaturdaten, die auf unterschiedlichen Topographien basieren, ist, weisen unter anderem L. Gao et al. (2012) bei ihrer Höhenkorrektur

der ERA-Interim-Daten darauf hin, dass die Verwendung von konstanten Korrekturfaktoren vor allem in gebirgigen Gebieten problematisch ist und zu größeren Fehlern führen kann. Zudem muss beachtet werden, dass die zur Verfügung gestellten Rasterdatensätze auf Grundlage von Stationsdaten interpoliert werden und die Stationsdichte in gebirgigen Regionen nicht so hoch ist. Daher lassen sich die betragsmäßig höheren Differenzen zwischen Validations- und Modelldaten besonders in diesen Gebieten mit einem komplexen topographischen Terrain nicht nur mit einer schlechteren Abbildung oder dem Fehlen von dort vorhandenen Prozessen, sondern auch mit einer schlechteren Repräsentativität der Validationsdaten erklären.

### **7.2 Einschätzung der Modellperformanz durch eingeführte Änderungen, Anregungen und weitere Erkenntnisse für die Modellentwicklung**

Wie in Kapitel 6 ausführlich beschrieben wurde, werden durch die Vielzahl an Änderungen in den Daten und Prozessen kaum Verbesserungen in den Qualitätsindizes wie dem Korrelationskoeffizienten oder dem RMSE erreicht. Dennoch ist anhand der Auswirkungen beim Einschalten des originalen 5Layer-Schemas zu erkennen, dass REMO durchaus sensibel auf Änderungen in der Parametrisierung reagiert und auch Verschlechterungen durch neue Prozesse auftreten können. Änderungen in eingesetzten Parametrisierungen können aufgrund der komplexen und hochgradig nicht linearen Prozessketten sowie der positiven und negativen Rückkopplungen, die zwischen den beiden Systemen Atmosphäre und Landoberfläche auftreten, großen Einfluss haben. In dieser Arbeit wurden mehrere Parametrisierungen vor allem in Hinblick auf Bodenprozesse überarbeitet. Daher ist es als Erfolg zu werten, dass alle eingebauten Änderungen keine Verschlechterung oder zumindest leichte Verbesserungen in den Qualitätsindizes anzeigen. Zudem wurde aufgrund der Menge der Änderungen hauptsächlich auf großskalige und saisonale Effekte eingegangen. Durch gegensätzliche Wechselwirkungen verschiedener Variablen und durch die Art der Aggregation bei der Analyse heben sich die Unterschiede der einzelnen Modellversionen teilweise auf.

Ein wichtiger Aspekt dieser Arbeit ist der Einbau des organischen Anteils und dessen Nutzung für die Bodenparametrisierung. Boden mit hohem organischen Anteil zeichnet sich durch eine geringere Wärmeleitfähigkeit, aber höhere Wärmekapazität aus. Außerdem führt das höhere Porenvolumen solcher Böden zu höheren Wasserkapazitäten, da mehr Wasser vom Boden aufgenommen werden kann. Durch diese Berücksichtigung des organischen Anteils werden in dieser Arbeit in beiden Untersuchungsgebieten durchschnittlich höhere Bodentemperaturen in allen Jahreszeiten erreicht. Dies widerspricht zunächst der Aussage von Guimberteau et al. (2018), welche den Einfluss der Berücksichtigung von organischem Material mittels des LSM ORCHIDEE testen und eine Reduktion der Sommertemperaturen nachweisen. Bei der genaueren räumlichen Analyse des organischen Einflusses auf die Modellergebnisse von REMO wird jedoch deutlich, dass in den höheren Breiten des EUR-44-Gebiets, für die auch Guimberteau et al. (2018) ihre Analyse durchgeführt haben und in denen durch den neuen Bodendatensatz flächendeckend höhere organische Anteile ausgewiesen werden, ebenfalls kühlere Bodentemperaturen modelliert

werden. Hier wird unter anderem der Einfluss der räumlichen und zeitlichen Aggregation der Daten auf die Analyseergebnisse deutlich. Guimberteau et al. (2018) weisen bei ihrer Studie zusätzlich darauf hin, dass für die gute Abbildung der Bodentemperaturen nicht nur ein gutes thermisches Schema, sondern auch eine gute Abbildung von Schnee im Winter und der Boden-Vegetation-Hydrologie-Prozesse im Sommer notwendig ist. Die thermalen Bodenparameter beeinflussen über den Bodenwärmestrom die im Modell verfügbare Energiemenge und damit die Interaktion zwischen den beiden Systemen Atmosphäre und Boden (vgl. van Looy et al., 2017). Daher ist es wichtig, dass die Bodenparameter vom Modell gut nachgebildet werden, weil es ansonsten zu großen Fehleinschätzungen der Energieflüsse kommt, was wiederum starke Auswirkungen auf die Boden- und Oberflächentemperatur hat (vgl. He et al., 2019).

Ein vermehrtes Interesse an diesen Bodenparametern und deren Weiterentwicklung vor allem in Hinblick auf PTFs, die für die Generierung von Parametern eingesetzt werden, welche für Landoberflächen- und Klimamodelle benötigt werden, lässt sich an der Häufung der Veröffentlichungen und Studien, die im Bearbeitungszeitraum dieser Arbeit entstanden sind, nachvollziehen. Die Einschätzung, das Ranking verschiedener Parametrisierungen und die Auswahl einer für ein Modell geeigneten Formel für die jeweiligen Bodenparameter ist jedoch schwierig, was der Vergleich der Literaturwerte in Kapitel 5.1.1 gezeigt hat. Auch Ergebnisse von zusätzlichen Vergleichsstudien sind immer von der jeweiligen Datengrundlage, deren Umfang an Informationen und Details sowie besonders vom jeweiligen Untersuchungsgebiet abhängig. Abdelbaki (2021) zeigt in einem Vergleich verschiedener PTFs, dass z. B. die in dieser Arbeit zunächst ausgewählte Parametrisierung nach Adhikary et al. (2008) für hydrologische Kennwerte hauptsächlich in feinen Böden und bei hoher Sättigung gute Ergebnisse liefern, in anderen Bodenzusammensetzungen und -zuständen jedoch andere PTFs überlegen sind. Die Kombination von verschiedenen Methoden oder die Verwendung einer Ensemble-Methode, wie es z. B. Y. Zhang et al. (2020) für PTFs in ESMs vorschlagen, birgt jedoch ein großes Risiko an Willkür, da durch die Auswahl und Anzahl der Ensemble-Mitglieder Modellergebnisse ebenso beeinflusst werden können. Für ein Klimamodell sollten aufgrund des ohnehin schon komplexen Zusammenspiels der verschiedenen Modellkomponenten möglichst einfache und in allen Regionen und allen Zuständen gut abschneidende Parametrisierungen verwendet werden oder eine differenzierte Anwendung sehr gut begründet und dokumentiert sein. Dass selbst in den jeweiligen Expertenrunden Uneinigkeit über die Performanz und das Ranking verschiedener Methoden herrscht, zeigt, dass hier noch reger Forschungsbedarf vorliegt. Die Ergebnisse der aktuellen und zukünftigen Forschungen in diesem Teilbereich sollten beachtet, mit den bestehenden Methoden verglichen und für den Einbau in REMO in Erwägung gezogen werden. Ferner sollten nicht nur die Werte der Parametrisierungen, sondern auch die Art der Parametrisierung beachtet werden. Kracher et al. (2009) untersuchten beispielsweise den Anteil des Residuums der Energiebilanzgleichung für drei verschiedene Modelle (abgeänderte Version von REMO, SEWAB und TERRA) und fanden heraus, dass der Bodenwärmestrom von REMO und TERRA stark überschätzt wird, weil dieser jeweils aus den anderen Komponenten der Energiebilanzgleichung berechnet wird.

Des Weiteren werden die in einem Klimamodell benötigten Bodenparameter und -variablen durch die Verfügbarkeit und Qualität der konstanten Eingangsdaten beeinflusst. Wie in den

in dieser Arbeit durchgeführten räumlichen Vergleichen der Karten verschiedener Datensätze nachgewiesen wurde, unterscheiden sich die Rasterdatensätze trotz gemeinsamer Ausgangsdaten deutlich. Ferner sind Artefakte, die auf eine nicht vollständige Homogenisierung von lokalen Datensätzen zurückzuführen sind, trotz einer starken Aggregation von 1 auf 50 km aufgefallen. Für ein besseres Verständnis bzw. eine Erklärung solcher Auffälligkeiten wäre eine gute und ausführliche Dokumentation der verwendeten Daten und deren Aufbereitungsmethoden für einzelne Datensätze wichtig. Ebenso wäre eine öffentlich zugängliche Sammlung aller weltweit durchgeführten Felddaten und Laboranalysen sinnvoll, um einerseits räumliche Datenlücken schließen und andererseits bestehende Studienergebnisse unter anderen Bedingungen testen zu können. Oftmals werden die erhobenen Bodenstichproben nur hinsichtlich einzelner Aspekte untersucht oder für spezielle Analysen verwendet. Durch die Veröffentlichung der Originaldaten könnten andere Studien und Forschungszweige ebenfalls profitieren. Der Zusammenschluss von hochaufgelösten Bodendaten in einer homogenisierten und standardisierten globalen Datenbank befindet sich im Aufbau (vgl. Arrouays et al., 2018; S. Chen et al., 2020). Trotzdem werden auch weiterhin unabhängige Datensätze, die nicht in den standardisierten Werten enthalten sind, zur Validation benötigt. Aber selbst die perfekte hochaufgelöste Bodendatenbank verbessert ein Modell nicht, wenn die Daten und deren Aufbereitung nicht für die Anwendung und das Modell geeignet sind. Die Aufbereitung der Daten ist nicht nur bei den variablen klimatischen Validationsdaten (vgl. Kapitel 7.1), sondern auch für konstante Eingangsdaten entscheidend. Die verwendete Modellauflösung und die zur Aggregation der Eingangsdaten benutzten Methoden beeinflussen das Modellergebnis. Ein Beispiel hierfür liefern Montzka et al. (2017), die anhand von Schaap et al. (2001) die van Genuchten Parameter für den SoilGrids-Datensatz auf unterschiedliche räumliche Skalen ableiten. Sie stellen fest, dass durch die Interpolation der 1 km Bodendaten auf 16 km ca. 10% und auf 128 km ca. 30% der Varianz der Bodendaten verloren gehen. Die Berücksichtigung der Anteile verschiedener Wertegruppen (z. B. Bodenarten oder Bodentexturen) bzw. der Variabilität innerhalb eines Pixels bei metrischen Variablen, wie es für die Landbedeckung bereits vielfach durchgeführt wird (zum Beispiel durch Plant-Functional-Types in REMO-iMOVE), könnte für die Zukunft auch für die Bodenparameter überdacht werden.

Besonders wichtig für eine möglichst realitätsnahe Modellierung der thermischen und hydrologischen Prozesse im Boden ist die Einteilung in verschiedene Bodenschichten. Wie in Kapitel 5.2 erläutert, hängt beispielsweise die Qualität der Lösung der partiellen Differentialgleichung zur Berechnung der Bodentemperaturen nicht nur vom gewählten Modellzeitschritt, sondern auch von der jeweiligen Schichtdicke ab. In REMO sind bisher fünf Bodenschichten, deren Tiefen aus dem Europa Modell übernommen wurden, für die Bodentemperaturberechnung enthalten. Auch für die Reanalysedaten ERA5 werden nur vier Schichten verwendet, was durch die Studien von Deardorff (1978) und Warrilow et al. (1986) definiert wird, weil diese in Kombination mit den verwendeten Werten für die Bodenparameter (konstante volumetrische Wärmekapazität und variabler Wärmeleitfähigkeit nach Peters-Lidard et al. (1998)) die unterschiedlichen Amplituden und Phasen der Bodenvariablen auf allen Zeitskalen (täglich bis jährlich) gut abbilden können (ECMWF, 2016). Andere Klimamodelle, wie beispielsweise das CLM5 mit standardmäßig 20 Boden- und fünf Grundgesteinsschichten (bis 49.556 m Tiefe) (Lawrence et al., 2019), und vor allem hydrologische Modelle, wie Hydrus-1D (Simunek et al., 2005) oder WaSiM (Schulla, 2021)



mit über 100 möglichen Bodenschichten, haben deutlich mehr und eine variable Anzahl an Bodenschichten, die für die jeweiligen Modellläufe durch den Benutzer angepasst werden können. Hier könnten weitere Sensitivitätsstudien für REMO ansetzen, um Effekte unterschiedlicher Anzahl und Tiefe der Bodenschichten auch in Hinblick auf eine variable horizontale Modellauflösung zu untersuchen. Die Anpassung in REMO im Bezug auf die hydrologischen Bodenprozesse von einem Bucket auf das 5Layer-Schema (Modellversion id11) zeigte bei den Differenzanalysen aller Modelle die größten Auswirkungen auf die Modellergebnisse. Da sich die physikalischen Eigenschaften von Wasser stark von denen von trockenen Boden unterscheiden, hat der Wassergehalt den größten Einfluss auf die Bodentemperatur und damit auch auf die gesamten Wärmeflüsse im Modell. Gleichzeitig stellt die Verfeinerung der vertikalen Auflösung der Bodenschichten eine Herausforderung an die ohnehin schlechte Datenlage hinsichtlich Validierungsdaten für die Bodentemperatur und den Bodenwassergehalt dar. Außerdem wird in den meisten Modellen weder der Aufbau noch die Struktur des Bodens berücksichtigt, weil hierfür noch weniger Daten vorliegen und der Effekt für die bisherigen Auflösungen minimal war. Bei höheren Modellauflösungen werden jedoch der Aufbau und das Gefüge des Bodens, wie beispielsweise wasserundurchlässige Tonschichten, die Wassersperrschichten bilden, oder die Berücksichtigung von Klüften, Wurzeln und Ähnlichem wichtig, um die Bodenprozesse möglichst gut abbilden zu können. Ansätze und deren Effekte, auch die Bodenstruktur, die hauptsächlich die hydraulischen Leitfähigkeiten im Boden beeinflusst, durch Parametrisierungen sowohl lokal als auch global in ein ESM zu integrieren, sind bei Faticchi et al. (2020)) zu finden. Dass für unterschiedliche räumliche Anordnung verschiedener Materialien unterschiedliche Berechnungsmethoden zur Verfügung stehen, zeigen Dong et al. (2015) anhand der Ableitung der Wärmeleitfähigkeit für gesättigte und ungesättigte Böden. Zudem werden in LSMs und in Klimamodellen aufgrund der räumlichen Parallelisierung der Rechenoperationen (gleichzeitiges und daher wesentlich schnelleres Ausführen von Rechnungen auf mehreren CPU-Kernen durch Unterteilung des Untersuchungsgebietes), die Wärme- und Wasserflüsse nur vertikal als eindimensionale Diffusionsgleichungen betrachtet (Niu & Zeng, 2012). In der Realität finden jedoch auch horizontale Flüsse statt. Für die bisherigen Modellauflösungen waren diese aufgrund des Einflusses im Vergleich zu den vertikalen Flüssen vernachlässigbar. Je höher die Modellauflösung wird, desto mehr spielen aber auch diese Prozesse eine Rolle, sodass entweder eine Parametrisierung oder die Lösung der dreidimensionalen thermischen und hydrologischen Diffusionsgleichungen ins Modell integriert werden müssen. Ab welcher Modellskala diese Prozesse einen entscheidenden Effekt auf die Modellergebnisse haben, könnte ebenfalls Gegenstand weiterer Studien sein. Weitere Prozesse wie beispielsweise verschiedene biochemische Reaktionen (van Looy et al., 2017), eine temperaturabhängige Oberflächenalbedo (S. K. Saha, 2006) oder die Veränderung der Bodenalbedo durch den darin befindlichen Wassergehalt, was im Klimamodell aber nur bei einer unbewachsenen Landoberfläche zutragen kommt, welche jedoch starke Wertunterschiede auch für verschiedene Texturklassen hat, wie F. J. Braun (2002) nachweist, sind in der aktuellen Standardversion von REMO und auch in anderen Modellen noch nicht berücksichtigt, werden aber für die höheren Modellauflösungen wichtig.

Auch für die Topographie-Daten, die in dieser Arbeit thematisiert und analysiert wurden, spielt die Modell(sub)skala eine wichtige Rolle. Die Varianz der Topographie wird in REMO bisher indirekt zur Berechnung des Anteils des Oberflächenabflusses verwendet. Die in REMO dafür

verwendete Formel (vgl. Formel 3.4) ist jedoch an die bisherigen GTOPO-Daten und deren Originalauflösung von 1 km angepasst (vgl. höhere Werte für  $\sigma_0$  und  $\sigma_{max}$  in der Version von Dümenil & Todini (1992) für gröbere Modellauflösungen). Diese empirischen Parameter müssen künftig für noch höher aufgelöste Topographiedaten angepasst werden, was im Rahmen dieser Arbeit noch nicht geschehen ist. Diese Nichtanpassung sowie die Verwendung von fest definierten Grenzwerten in der Formel für  $b$  sind mit Gründe dafür, dass trotz der teilweise großen Wertunterschiede der topographischen Varianz kaum Änderungen in den davon abhängigen hydrologischen Variablen nachgewiesen werden konnten. Zudem wird in REMO zwar die Steilheit der Modelloberfläche durch die Varianz in der Kurvenparameterformel indirekt für den Oberflächenabfluss berücksichtigt, jedoch werden keinerlei Abflussnetze generiert oder ein zusätzliches nachträgliches Versickern von Wasser, welches sich in Senken sammelt, wie es in hydrologischen Modellen passiert, angenommen. Diese Prozessabstriche können auf den bisherigen Modellskalen von 12–50 km, auf denen REMO betrieben wurde, vernachlässigt werden, für höhere Modellauflösungen und bei der Entwicklung hin zu einem regionalen Erdsystemmodell sollte dies jedoch berücksichtigt werden. Des Weiteren könnten die Hangneigung und -orientierung, die aufgrund von Abschattungseffekten besonders in Gebieten mit heterogener Topographie einen großen Anteil an den Werten der Strahlungsvariablen haben, welche wiederum die Hauptquelle und -senke der Oberflächenenergie darstellen (Oliphant et al., 2003; M. D. Müller & Scherer, 2005), in REMO berücksichtigt werden. Besonders in Hinblick auf die zukünftig höhere Zielauflösung des Modells wird die differenzierte Betrachtung der direkten und indirekten Strahlung von beschatteten Tälern oder nicht beschatteten Hängen (vgl. Kotlarski, 2007) nötig werden. Oliphant et al. (2003) weisen bei ihrer Studie zur räumlichen Variabilität der Strahlungsflüsse in komplexen topographischen Gebieten darauf hin, dass aufgrund der nicht linearen Zusammenhänge von Strahlung und der Umgebung eine Betrachtung der Werte auf der Subskala nötig ist. Als Grundlage für diese Weiterentwicklung sowohl hinsichtlich der hydrologischen Komponenten als auch hinsichtlich der Strahlungsvariablen können dabei die durch diese Arbeit neu für REMO verfügbaren 30 m aufgelösten ASTER-Daten dienen.

Die unter anderem aus der Topographie und deren Variation abgeleitete Rauigkeitslänge hat ebenfalls nicht nur auf die Windvariablen, sondern auch auf den Energieaustausch zwischen der Modelloberfläche und der Atmosphäre einen Einfluss. Sie stellt aufgrund dieser Wechselwirkungen eine wichtige Variable im Klimasystem dar, für die jedoch verschiedene Berechnungsformeln vorliegen. REMO verwendet beispielsweise für den topographischen Teil des Rauigkeitswertes mit der in Anhang D dargestellten Formel D.4 eine ähnliche, aber nicht ganz identische Formel wie das nicht-hydrostatische numerische Wettervorhersagemodell COSMO, welche bei Buzzi (2008) beschrieben ist und welche ebenfalls auf 1 km aufgelöste Topographiedaten angepasst ist. Durch die neuen höher aufgelösten Topographiedaten können und müssen angesichts einer Erhöhung der Modellauflösung die bestehenden Parametrisierungen optimiert und beispielsweise Rauigkeitslängen auf der Subskala bereitgestellt werden, was bereits bei Buzzi (2008) in Kombination mit dem COSMO-Modell zu Verbesserungen der 2 m-Temperatur in den Alpen geführt hat. Auch Rontu (2006) konnte im HIRLAM mit der Anpassung der Parametrisierungen, die mit der Topographie zusammenhängen, besonders in Regionen mit komplexen Terrain Verbesserungen in den Windvariablen erzielen.

Insgesamt sind zur Einschätzung, ob die neuen Datensätze und Parametrisierungen zur Verbesserung von REMO beitragen und zukünftig standardmäßig verwendet werden sollten, weitere umfassende Analysen notwendig. Die Tatsache, dass bereits leichte Differenzen einer Variable zwischen zwei Modellversionen aufgrund der komplexen Wechselwirkungen große Auswirkungen haben können oder dass große Differenzen in Variablen kaum Effekte auf andere Variablen haben, erschwert das Nachvollziehen der Zusammenhänge und die Lokalisation von Modelloptimierungspunkten. Um mögliche systematische Fehler im Modell zu minimieren, muss zunächst die Ursache für diese gefunden werden. Beispielsweise fiel bei der Modellierung im Zuge der tendenziell leichten Temperaturerhöhung aufgrund der neuen Bodendaten in manchen höhergelegenen Regionen Regen statt Schnee. Dadurch entwickelte sich keine Schneedecke und aufgrund der damit verbundenen niedrigeren Albedo (Schneeealbedo ist mit 0.6–0.9 sehr hoch) hat sich auch die Strahlungsbilanz geändert, wodurch wiederum eine Erwärmung der Oberfläche ausgelöst wird. Dies ist jedoch nur bei der Betrachtung von einzelnen Pixeln und Tagesdaten aufgefallen, da bei der monatlichen oder saisonalen Betrachtung andere Effekte dieses Phänomen überlagerten. Andere Studien wie beispielsweise Kothe et al. (2011), welche die Strahlungskomponenten auch unter Berücksichtigung von Wolkenbedeckung, Oberflächenalbedo und Oberflächentemperatur hauptsächlich für das CLM-Modell, aber auch für REMO im Vergleich zu verschiedenen Referenzdaten untersuchen, oder Kracher et al. (2009), welche die Strahlungsflüsse für unterschiedliche Modelle miteinander vergleichen, gehen sogar noch eine Zeitebene tiefer und unterscheiden bei ihren Analysen zwischen Tages- und Nachtwerten. Für REMO, welches bei Kracher et al. (2009) in einer veränderten Version verwendet wurde, konnte so nachgewiesen werden, dass der Bodenwärmestrom, der sensible Wärmefluss und die Nettostrahlung am Tag überschätzt und in der Nacht dagegen unterschätzt werden. Dies zeigt, dass zusätzliche Analysen sowohl hinsichtlich einer räumlich als auch zeitlich differenzierten Betrachtung notwendig sind, um Prozessketten innerhalb des Modells besser verstehen zu können. Zudem könnten weitere Analysemethoden, wie Kompositanalysen, verschiedene Frequenzanalysen oder die separate Betrachtung verschiedener Zeitreihenanteile und deren Veränderung, größere Unterschiede zwischen den in dieser Arbeit getesteten Modellversionen aufzeigen als die durchgeführten Standardanalysen wie Mittelwert-, RMSE- oder Korrelationsvergleich.

Durch die in der vorliegenden Arbeit eingeführten variablen Anfangsbedingungen für Topographie und Boden sowie die Möglichkeit aus unterschiedlichen Parametrisierungen auszuwählen, kann REMO besser mit anderen regionalen Klimamodellen verglichen oder kombiniert werden, da der Aspekt der verschiedenen Eingangsdaten oder Parametrisierungen für die Ergebnisse der anderen Klimamodellläufe nicht mehr gegeben ist. Unterschiede könnten daher direkt an dem Vorhandensein, der Anzahl oder Art der simulierten Prozesse festgemacht werden. Das erleichtert die Ursachenforschung für systematische Modellunterschiede, die Auswahl an geeigneten Methoden und Parametrisierungen für bestimmte Prozesse und damit die Entwicklung der regionalen Klimamodelle hin zu regionalen Erdsystemmodellen.

Außerdem hat die strukturierte Analyse dieser Arbeit ergeben, dass der größte Qualitätssprung bei der Erhöhung der Auflösung von  $0.44^\circ$  auf  $0.11^\circ$  auftritt. Diese starke Verbesserung der Modellergebnisse durch die feinere Modellauflösung wird ebenfalls in den Studien von EURO-CORDEX

(Jacob et al., 2020), in deren Ensembleläufe auch REMO verwendet wird, hervorgehoben. Daher ist es essentiell, dass für den nächsten Schritt hin zu Modellauflösungen in wenigen Kilometerbereich geeignete Ausgangsdaten vorliegen. Dies konnte durch die neu eingebauten Topographie- und Bodendaten sowie die neuen Parametrisierungen erreicht werden.

## 8 Fazit und Ausblick

Wie zu Beginn dieser Arbeit erläutert, handelt es sich bei Klimamodellen um vereinfachte Abbildungen des komplexen Klimasystems der Erde. Erkenntnisse, wie sich das Klima in Zukunft entwickelt, und die Klimafolgenforschung nehmen in der Literatur und Forschung ein immer wichtigeres Themengebiet ein, dennoch sollte auch weiterhin Grundlagenforschung betrieben werden. Klimamodelle liefern die Daten für diese Arbeiten und je besser die Klimamodelle die Wirklichkeit abbilden, desto weniger aufwendig ist das Postprozessieren der verwendeten Modell-daten wie etwa Downscaling oder Bias-Korrekturen. Beispielsweise hoffen Navarro-Racines et al. (2020) darauf, dass durch den Fortschritt in der Klimamodellierung, diese Postprozessierungen in Zukunft nicht mehr nötig sein werden. Allerdings sind viele Zusammenhänge im Klimasystem, vor allem auf kleinen räumlichen Skalen, noch nicht hinreichend erforscht und bekannt. Hierfür sind die regionalen Klimamodelle das Werkzeug zur Gewinnung von Erkenntnissen, indem die Qualität verschiedener Parametrisierungen verglichen und bestehende Ansätze immer in Hinblick auf vorkommende Modellfehler hinterfragt werden.

Insgesamt konzentrierte sich diese Arbeit auf den methodischen und technischen Aspekt. Alle Änderungen wurden mit dem Ziel auf eine Weiterverwendung, die kompatibel mit dem bisherigen Modell ist, vorgenommen. Es wurde darauf geachtet, dass die Änderungen zwar eine inhaltliche Verbesserung und die verwendeten Parametrisierungen eine Spezialisierung der bereits vorhandenen Parameter und Prozesse darstellen, dass aber dennoch so wenig zusätzliche Eingangsdaten wie möglich benötigt werden und dass REMO als Regionalmodell weiterhin global einsetzbar bleibt.

In dieser Arbeit hat sich herausgestellt, dass bereits für die Anfangsbedingungen, die zum Betreiben eines Klimamodells notwendig sind, viele unterschiedliche Datensätze zur Verfügung stehen, welche jeweils verschiedene Stärken und Schwächen aufweisen. Durch den technischen Fortschritt und durch die neuen Erkenntnisse in der Forschung können immer höher aufgelöste Daten produziert werden, die zum Einbau in ein Klimamodell benutzt werden können. Wie durch die durchgeführten Analysen gezeigt wurde, sollten trotzdem alle Daten geprüft werden, ob sie sich für das jeweilige Modell eignen und wie die Daten jeweils angepasst werden müssen.

Die Eingangsdaten für ein Klimamodell spielen trotz ihrer zeitlich konstanten Werte eine wichtige Rolle, die jedoch erst durch die Kombination mit anderen Variablen und mit den modellierten Prozessen deutlich wird. Wegen den vielen Wechselwirkungen zwischen der Atmosphäre und dem Boden können durch eine verbesserte Abbildungen der Boden- und Vegetationsprozesse Klimamodellergebnisse nachhaltig optimiert werden (Prein et al., 2015). Dass trotz größerer Änderungen der Rahmenbedingungen in dieser Arbeit der Unterschied zu den verwendeten Validationsdaten wesentlich größer ist als die Variabilität zwischen den verschiedenen Modellläufen, zeigt, dass im Klimamodell noch nicht alle Prozesse ausreichend abgebildet sind. Es konnte beispielsweise nachgewiesen werden, dass das alleinige Verändern der thermischen Bodenparameter kaum Änderungen in den thermischen Flüssen und resultierenden Temperaturen bewirken. Durch

die Kombination mit einem veränderten Wassergehalt (Modellversion id11) treten dagegen größere Veränderungen auf. Durch weitere Studien in diesem Modellbereich, die teilweise bereits parallel in einer anderen Dissertation stattfinden und sich auf die verbesserte Darstellung der Bodenhydrologie im Modell konzentrieren, können die vorgestellten Ergebnisse dieser Arbeit weiter ergänzt werden. Besonders hilfreich für diese erfolgreiche Weiterentwicklung der Modelle sind Validationsdaten, die sich nicht nur auf die Zielvariablen wie Lufttemperatur und Niederschlag, sondern auch auf die Hilfs- und Zwischenschrittvariablen konzentrieren. Vor allem für die in dieser Arbeit betrachteten Bodenvariablen liegen jedoch nur wenige Daten vor, welche sich zudem deutlich unterscheiden.

Zukünftig sind weitere Sensitivitätsstudien auf noch höheren Auflösungen, die allerdings die nicht-hydrostatische Version von REMO voraussetzen, notwendig, um das ganze Potential der neuen Daten und Parametrisierungen testen zu können. Des Weiteren zeigt die vorliegende Arbeit einige Teile in REMO mit Verbesserungspotential auf. Besonders die Prozesse und Parametrisierungen im Modellboden sind noch nicht ausreichend erforscht. Die in dieser Arbeit weiterentwickelte REMO-Version und die neue Bodenbibliothek-Version bilden eine Grundlage für weitere Studien in diesen Teilen von REMO.

# Literaturverzeichnis

- Abdelbaki, A. M. (2021). „Assessing the best performing pedotransfer functions for predicting the soil–water characteristic curve according to soil texture classes and matric potentials“. In: *European Journal of Soil Science* 72.1, S. 154–173. DOI: 10.1111/ejss.12959.
- Abel, D., F. Pollinger, K. Ziegler & H. Paeth (2020). „Extension of the regional climate model REMO by a 5-layer soil scheme“. In: *General Assembly 2020*. Hrsg. von EGU. DOI: 10.5194/egusphere-egu2020-2964.
- Abrams, M., R. Crippen & H. Fujisada (2020). „ASTER Global Digital Elevation Model (GDEM) and ASTER Global Water Body Dataset (ASTWBD)“. In: *Remote Sensing* 12.7, S. 1156. DOI: 10.3390/rs12071156.
- Abu-Hamdeh, N. H. (2003). „Thermal Properties of Soils as affected by Density and Water Content“. In: *Biosystems Engineering* 86.1, S. 97–102. DOI: 10.1016/S1537-5110(03)00112-0.
- Acutis, M. & M. Donatelli (2003). „SOILPAR 2.00: software to estimate soil hydrological parameters and functions“. In: *European Journal of Agronomy* 18.3-4, S. 373–377.
- Adhikary, P. P., D. Chakraborty, N. Kalra, C. B. Sachdev, A. K. Patra, S. Kumar, R. K. Tomar, P. Chandna, D. Raghav & K. Agrawal (2008). „Pedotransfer functions for predicting the hydraulic properties of Indian soils“. In: *Soil Research* 46.5, S. 476–484.
- Aina, P. O. & S. P. Periaswamy (1985). „ESTIMATING AVAILABLE WATER-HOLDING CAPACITY OF WESTERN NIGERIAN SOILS FROM SOIL TEXTURE AND BULK DENSITY, USING CORE AND SIEVED SAMPLES“. In: *Soil Science* 140.1, S. 55–58. DOI: 10.1097/00010694-198507000-00007.
- Al Majou, H., B. Hassani & A. Bruand (2018). „Transferability of continuous- and class-pedotransfer functions to predict water retention properties of semiarid Syrian soils“. In: *Soil Use and Management* 34.3, S. 354–369. DOI: 10.1111/sum.12424.
- Alaya, I., M. M. Masmoudi, P. Lagacherie, G. Coulouma, F. Jacob & N. Ben Mechlia (2017). „Performance of Saxton and Rawls Pedotransfer Functions for Estimating Soil Water Properties in the Cap Bon Region-Northern Tunisia“. In: *Water and Land Security in Drylands*. Hrsg. von M. Ouessar, D. Gabriels, A. Tsunekawa & S. Evett. Cham: Springer International Publishing, S. 77–85. DOI: 10.1007/978-3-319-54021-4\_8.
- Aldrian, E., D. Sein, D. Jacob, L. D. Gates & R. Podzun (2005). „Modelling Indonesian rainfall with a coupled regional model“. In: *Climate Dynamics* 25.1, S. 1–17. DOI: 10.1007/s00382-004-0483-0.
- Alexander, L. V. et al. (2006). „Global observed changes in daily climate extremes of temperature and precipitation“. In: *Journal of Geophysical Research* 111.D5. DOI: 10.1029/2005JD006290.
- Amante, C. (2009). *ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis*. DOI: 10.7289/V5C8276M.
- Amelung, W., H.-P. Blume, H. Fleige, R. Horn, E. Kandeler, I. Kögel-Knabner, R. Kretschmar, K. Stahr & B.-M. Wilke (2018). *Scheffer/Schachtschabel Lehrbuch der Bodenkunde*. 17., überarbeitete und ergänzte Auflage. Lehrbuch. Berlin: Springer Spektrum.
- An, K., W. Wang, Y. Zhao, W. Huang, L. Chen, Z. Zhang, Q. Wang & W. Li (2016). „Estimation from Soil Temperature of Soil Thermal Diffusivity and Heat Flux in Sub-surface Layers“. In: *Boundary-Layer Meteorology* 158.3, S. 473–488. DOI: 10.1007/s10546-015-0096-7.
- Arakawa, A. & V. R. Lamb (1977). „Computational Design of the Basic Dynamical Processes of the UCLA General Circulation Model“. In: *General Circulation Models of the Atmosphere*. Bd. 17. Methods in Computational Physics: Advances in Research and Applications. Elsevier, S. 173–265. DOI: 10.1016/B978-0-12-460817-7.50009-4.

- Arrouays, D. et al. (2018). „THE GLOBALSOILMAP PROJECT: PAST, PRESENT, FUTURE, AND NATIONAL EXAMPLES FROM FRANCE“. In: *Dokuchaev Soil Bulletin* 95, S. 3–23. DOI: 10.19047/0136-1694-2018-95-3-23.
- Arrouays, D. et al. (2017). „Soil legacy data rescue via GlobalSoilMap and other international and national initiatives“. In: *GeoResJ* 14.9, S. 1–19. DOI: 10.1016/j.grj.2017.06.001.
- Arruda, F. B., J. Zullo Junior & J. B. de Oliveira (1987). „Soil parameters for calculating available water based on soil texture“. In: *Revista Brasileira de Ciencia do Solo (Brazil)*.
- Arya, S. P. (2008). *Introduction to micrometeorology*. 2. ed., [Nachdr.] Bd. 79. International geophysics series. San Diego, Calif.: Acad. Press.
- Asselin, R. (1972). „Frequency Filter for Time Integrations“. In: *Monthly Weather Review* 100.6, S. 487–490. DOI: 10.1175/1520-0493(1972)100<0487:FFFTI>2.3.CO;2.
- Bachmann, J. (2014). „Thermisches Verhalten der Böden“. In: *Handbuch der Bodenkunde*. Hrsg. von H.-P. Blume, K. Stahr, W. Fischer, G. Guggenberger, R. Horn, H.-G. Frede & P. Felix-Henningsen. Wiley, S. 1–34. DOI: 10.1002/9783527678495.hbbk2005006.
- Ballabio, C., P. Panagos & L. Monatanarella (2016). „Mapping topsoil physical properties at European scale using the LUCAS database“. In: *Geoderma* 261, S. 110–123. DOI: 10.1016/j.geoderma.2015.07.006.
- Balland, V. & P. A. Arp (2005). „Modeling soil thermal conductivities over a wide range of conditions“. In: *Journal of Environmental Engineering and Science* 4.6, S. 549–558. DOI: 10.1139/s05--007.
- Balsamo, G., C. Albergel, A. Beljaars, S. Boussetta, E. Brun, H. Cloke, D. P. Dee, E. Dutra, J. Muñoz-Sabater, F. Pappenberger, P. de Rosnay, T. Stockdale & F. Vitart (2015). „ERA-Interim/Land: a global land surface reanalysis data set“. In: *Hydrology and Earth System Sciences* 19.1, S. 389–407. DOI: 10.5194/hess-19-389-2015.
- Balsamo, G., A. Beljaars, K. Scipal, P. Viterbo, B. van den Hurk, M. Hirschi & A. K. Betts (2009). „A Revised Hydrology for the ECMWF Model: Verification from Field Site to Terrestrial Water Storage and Impact in the Integrated Forecast System“. In: *Journal of Hydrometeorology* 10.3, S. 623–643. DOI: 10.1175/2008JHM1068.1.
- Bamber, J. L., S. Ekholm & W. B. Krabill (2001). „A new, high-resolution digital elevation model of Greenland fully validated with airborne laser altimeter data“. In: *Journal of Geophysical Research* 106.B4, S. 6733–6745. DOI: 10.1029/2000JB900365.
- Batjes, N. H. (1996). „Development of a world data set of soil water retention properties using pedotransfer rules“. In: *Geoderma* 71.1-2, S. 31–52. DOI: 10.1016/0016-7061(95)00089-5.
- Batjes, N. H., G. Fischer, F. Nachtergaele, V. S. Stolbovoy & H. van Velthuizen (1997). *Soil Data Derived from WISE for Use in Global and Regional AEZ Studies (Version 1.0): Interim Report*.
- Benoit, R. (1976). „A Comprehensive Parameterization of the Atmospheric Boundary Layer for General Circulation Models“. Dissertation. McGill University Canada.
- Beringer, J., A. H. Lynch, F. S. Chapin, M. Mack & G. B. Bonan (2001). „The Representation of Arctic Soils in the Land Surface Model: The Importance of Mosses“. In: *Journal of Climate* 14.15, S. 3324–3335. DOI: 10.1175/1520-0442(2001)014<3324:TROASI>2.0.CO;2.
- Berrisford, P., D. P. Dee, R. Brugge, K. Fielding, M. Fuentes, P. W. Kallberg, Kobayashi, Uppala, S. M. & A. J. Simmons (2011). „The ERA-Interim Archive: Version 2.0“. In: *ERA Report Series 1, ECMWF*.
- Berrisford, P., D. P. Dee, K. Fielding, M. Fuentes, P. W. Kallberg, S. Kobayashi & S. M. Uppala (2009). „The ERA-interim archive: (version 1.0)“. In: *ERA Report Series: European Centre for Medium Range Weather Forecasts*.
- Best, M. J., M. Pryor, D. B. Clark, G. G. Rooney, R. L. H. Essery, C. B. Ménard, J. M. Edwards, M. A. Hendry, A. Porson & N. GEDNEY (2011). „The Joint UK Land Environment Simulator (JULES), model description–Part 1: energy and water fluxes“. In: *Geoscientific Model Development* 4.1, S. 677–699.



- Bildirici, İ. Ö. & R. A. Abbak (2017). „COMPARISON OF ASTER AND SRTM DIGITAL ELEVATION MODELS AT ONE-ARC-SECOND RESOLUTION OVER TURKEY“. In: *Selcuk University Journal of Engineering, Science and Technology* 5.1, S. 16–25. DOI: 10.15317/Scitech.2017.66.
- Blanc & Milton L. (1957). „Measuring Soil Moisture“. In: *Weekly Weather and Crop Bulletin* 44.18.
- Block, A. (2007). „Unsicherheiten in Oberflächen- und Bodenparametern und ihre Auswirkungen auf die Ergebnisse regionaler Klimasimulationen“. Dissertation. Brandenburgische Technische Universität Cottbus.
- Blome, T. (2014). „Influence of different permafrost processes on the large-scale energy and water cycles over Siberia“. Diss.
- Blott, S. J. & K. Pye (2012). „Particle size scales and classification of sediment types based on particle size distributions: Review and recommended procedures“. In: *Sedimentology* 59.7, S. 2071–2096. DOI: 10.1111/j.1365-3091.2012.01335.x.
- Bonan, G. B. (2019). *Climate change and terrestrial ecosystem modeling*. Cambridge, United Kingdom: Cambridge University Press. DOI: 10.1017/9781107339217.
- Bonan, G. B., K. W. Oleson, M. Vertenstein, S. Levis, X. Zeng, Y. Dai, R. E. Dickinson & Z.-L. Yang (2002). „The Land Surface Climatology of the Community Land Model Coupled to the NCAR Community Climate Model“. In: *Journal of Climate* 15.22, S. 3123–3149. DOI: 10.1175/1520-0442(2002)015<3123:TLSCOT>2.0.CO;2.
- Bonekamp, P. N. J., E. Collier & W. W. Immerzeel (2018). „The Impact of Spatial Resolution, Land Use, and Spinup Time on Resolving Spatial Precipitation Patterns in the Himalayas“. In: *Journal of Hydrometeorology* 19.10, S. 1565–1581. DOI: 10.1175/JHM-D-17-0212.1.
- Bortz, J. & C. Schuster (2010). *Statistik für Human- und Sozialwissenschaftler*. Berlin, Heidelberg: Springer Berlin Heidelberg. DOI: 10.1007/978-3-642-12770-0.
- Botula, Y.-D. (2013). „Indirect methods to predict hydrophysical properties of soils of Lower Congo“. Dissertation. Ghent University.
- Botula, Y.-D., W. M. Cornelis, G. Baert & E. van Ranst (2012). „Evaluation of pedotransfer functions for predicting water retention of soils in Lower Congo (DR Congo)“. In: *Agricultural Water Management* 111, S. 1–10.
- Bouma, J. (1989). „Using soil survey data for quantitative land evaluation“. In: *Advances in soil science* 9, S. 177–213. DOI: 10.1007/978-1-4612-3532-3\_4.
- Bowen, I. S. (1926). „The Ratio of Heat Losses by Conduction and by Evaporation from any Water Surface“. In: *Physical Review* 27.6, S. 779–787. DOI: 10.1103/PhysRev.27.779.
- Braden, H. (1995). *The model AMBETI: A detailed description of a soil-plant-atmosphere model*. Bd. 195. Berichte des Deutschen Wetterdienstes. Offenbach am Main: Selbstverl. des Dt. Wetterdienstes.
- Braun, B., W. Helms & E. Makowski (1985). „Berechnung der Frostausbreitung im Bergbau“. In: *Handbuch des Gefrierschachtbaus im Bergbau*. Hrsg. von J. Klein. Glückauf-Betriebsbücher. Essen: Verlag Glückauf, S. 178–198.
- Braun, F. J. (2002). „Mesoskalige Modellierung der Bodenhydrologie“. Dissertation. DOI: 10.5445/IR/5432002.
- Braun, F. J. & G. Schädler (2005). „Comparison of Soil Hydraulic Parameterizations for Mesoscale Meteorological Models“. In: *Journal of Applied Meteorology* 44.7, S. 1116–1132. DOI: 10.1175/JAM2259.1.
- Briggs, L. J. & H. L. Shantz (1912). „The Wilting Coefficient and Its Indirect Determination“. In: *Botanical Gazette* 53.1, S. 20–37. DOI: 10.1086/330708.
- Brisson, E., M. Demuzere & N. P. van Lipzig (2016). „Modelling strategies for performing convection-permitting climate simulations“. In: *Meteorologische Zeitschrift* 25.2, S. 149–163. DOI: 10.1127/metz/2015/0598.
- Brooks, R. H. & A. T. Corey (1964). „Hydraulic properties of porous media“. In: *Hydrology Papers, Colorado State University, Fort Collins, Colorado* 3.

- Bruand, A., D. Baize & M. Hardy (1994). „Prediction of water retention properties of clayey soils: validity of relationships using a single soil characteristic“. In: *Soil Use and Management* 10.3, S. 99–103.
- Buntemeyer, L., N. Koldunov, K. Sieck, C. Teichmann, S. Kotlarski & A. R. Remedio (2015). „REMO 2015 MPI: User Guide“. In:
- Buzzi, M. (2008). „Challenges in operational numerical weather prediction at high resolution in complex terrain“. Dissertation. ETH Zurich. DOI: 10.3929/ethz-a-005698833.
- Cabos, W., D. V. Sein, A. Durán-Quesada, G. Liguori, N. V. Koldunov, B. Martínez-López, F. Alvarez, K. Sieck, N. Limareva & J. G. Pinto (2019). „Dynamical downscaling of historical climate over CORDEX Central America domain with a regionally coupled atmosphere–ocean model“. In: *Climate Dynamics* 52.7-8, S. 4305–4328. DOI: 10.1007/s00382-018-4381-2.
- Caglar, B., K. Becek, C. Mekik & M. Ozendi (2018). „On the vertical accuracy of the ALOS world 3D-30m digital elevation model“. In: *Remote Sensing Letters* 9.6, S. 607–615. DOI: 10.1080/2150704X.2018.1453174.
- Campbell, G. S. (1974). „A SIMPLE METHOD FOR DETERMINING UNSATURATED CONDUCTIVITY FROM MOISTURE RETENTION DATA“. In: *Soil Science* 117.6, S. 311–314. DOI: 10.1097/00010694-197406000-00001.
- Campbell, G. S. (1985). *Soil physics with BASIC: transport models for soil-plant systems*. Developments in soil science. Amsterdam: Elsevier.
- Carabajal, C. C. & J.-P. Boy (2016). „EVALUATION OF ASTER GDEM V3 USING ICESAT LASER ALTIMETRY“. In: *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLI-B4, S. 117–124. DOI: 10.5194/isprs-archives-XLI-B4-117-2016.
- Carsel, R. F. & R. S. Parrish (1988). „Developing joint probability distributions of soil water retention characteristics“. In: *Water Resources Research* 24.5, S. 755–769. DOI: 10.1029/WR024I005P00755.
- Chakraborty, D., S. Mazumdar, R. Garg, S. Banerjee, P. Santra, R. Singh & R. Tomar (2011). „Pedotransfer functions for predicting points on the moisture retention curve of Indian soils“. In: *Indian J Agric Sci* 81.
- Champeaux, J. L., V. Masson & F. Chauvin (2005). „ECOCLIMAP: A global database of land surface parameters at 1 km resolution“. In: *Meteorological Applications* 12.1, S. 29–32. DOI: 10.1017/S1350482705001519.
- Chen, F. & J. Dudhia (2001). „Coupling an Advanced Land Surface–Hydrology Model with the Penn State–NCAR MM5 Modeling System. Part I: Model Implementation and Sensitivity“. In: *Monthly Weather Review* 129.4, S. 569–585. DOI: 10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2.
- Chen, J., J. Wen, H. Tian, T. Zhang, X. Yang, D. Jia & X. Lai (2018). „A Study of Soil Thermal and Hydraulic Properties and Parameterizations for CLM in the SRYR“. In: *Journal of Geophysical Research: Atmospheres* 123.16, S. 8487–8499. DOI: 10.1029/2017jd028034.
- Chen, S. X. (2008). „Thermal conductivity of sands“. In: *Heat and Mass Transfer* 44.10, S. 1241–1246. DOI: 10.1007/s00231-007-0357-1.
- Chen, S., V. L. Mulder, L. Poggio, P. Roudier, Z. Libohova, B. Minasny, Z. Shi, J. Hannam & D. Arrouays (2020). „Digital mapping of soil information at a broad-scale: A review on GlobalSoilMap“. In: *General Assembly 2020*. Hrsg. von EGU. DOI: 10.5194/egusphere-egu2020-8426.
- Clapp, R. B. & G. M. Hornberger (1978). „Empirical Equations for Some Soil Hydraulic Properties“. In: *Water Resources Research* 14.4.
- Clemens, M. (1998). „Sensitivitätsstudien der im REMO implementierten Parametrisierungen des DM/DM und des ECHAM“. Diplomarbeit. Kiel: Christian-Albrechts-Universität Kiel.
- Cohen, B. H. (2002). „Calculating a factorial ANOVA from means and standard deviations“. In: *Understanding Statistics: Statistical Issues in Psychology, Education, and the Social Sciences* 1.3, S. 191–203.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. 2. ed. Hillsdale, NJ: Erlbaum.

- Conover, W. J. (1980). *Practical nonparametric statistics (2nd ed.)* New York, NY.
- Cornes, R. C., G. van der Schrier, E. J. M. van den Besselaar & P. D. Jones (2018). „An Ensemble Version of the E-OBS Temperature and Precipitation Data Sets“. In: *Journal of Geophysical Research: Atmospheres* 123.17, S. 9391–9409. DOI: 10.1029/2017JD028200.
- Cosby, B. J., G. M. Hornberger, R. B. Clapp & T. Ginn (1984). „A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils“. In: *Water Resources Research* 20.6, S. 682–690.
- Courty, L., J. C. Soriano-Monzalvo & A. Pedrozo-Acuña (2017). *Evaluation of open-access global digital elevation models (AW3D30, SRTM and ASTER) for flood modelling purposes*. DOI: 10.5281/zenodo.885571.
- Cox, P. M., R. A. Betts, C. B. Bunton, R. L. H. Essery, P. R. Rowntree & J. Smith (1999). „The impact of new land surface physics on the GCM simulation of climate and climate sensitivity“. In: *Climate Dynamics* 15, S. 183–203. DOI: 10.1007/s003820050276.
- Cressman, G. P. (1959). „An operational objective analysis system“. In: *Mon. Wea. Rev* 87.10, S. 367–374.
- Cui, X., G. Huang & W. Chen (2008). „Notes of numerical simulation of summer rainfall in China with a regional climate model REMO“. In: *Advances in Atmospheric Sciences* 25.6, S. 999–1008. DOI: 10.1007/s00376-008-0999-z.
- Dai, Y., W. Shangguan, Q. Duan, B. Liu, S. Fu & G. Niu (2013). „Development of a China Dataset of Soil Hydraulic Parameters Using Pedotransfer Functions for Land Surface Modeling“. In: *Journal of Hydrometeorology* 14.3, S. 869–887. DOI: 10.1175/JHM-D-12-0149.1.
- Dai, Y., W. Shangguan, N. Wei, Q. Xin, H. Yuan, S. Zhang, S. Liu, X. Lu, D. Wang & F. Yan (2019). „A review of the global soil property maps for Earth system models“. In: *SOIL* 5.2, S. 137–158. DOI: 10.5194/soil-5-137-2019.
- Dai, Y., N. Wei, H. Yuan, S. Zhang, W. Shangguan, S. Liu, X. Lu & Y. Xin (2019). „Evaluation of Soil Thermal Conductivity Schemes for Use in Land Surface Modeling“. In: *Journal of Advances in Modeling Earth Systems* 11.11, S. 3454–3473. DOI: 10.1029/2019MS001723.
- Dai, Y., Q. Xin, N. Wei, Y. Zhang, W. Shangguan, H. Yuan, S. Zhang, S. Liu & X. Lu (2019). „A Global High-Resolution Data Set of Soil Hydraulic and Thermal Properties for Land Surface Modeling“. In: *Journal of Advances in Modeling Earth Systems* 11.9, S. 2996–3023. DOI: 10.1029/2019MS001784.
- Dascalu, S. I., M. Gothard & A. Dumitrescu (2020). *European Reanalysis ERA5, UERRA MESCAN-SURFEX, COSMO\_REA6 compared to station-based data: Work Package 6, Deliverable 6.2: Report on the datasets inter-comparisons with regard to selected ECVs and INDECIS-ISD*. URL: <http://www.indecis.eu/docs/Deliverables/62summary.pdf>.
- Dashtaki, S. G. & M. Homae (2004). „Using geometric mean particle diameter to derive point and continuous pedotransfer functions“. In: *EuroSoil*, S. 1–10.
- Davies, H. C. (1976). „A lateral boundary formulation for multi-level prediction models“. In: *Quarterly Journal of the Royal Meteorological Society* 102.432, S. 405–418. DOI: 10.1256/smsqj.43209.
- Davin, E. L., D. Rechid, M. Breil, R. M. Cardoso, E. Coppola, P. Hoffmann, L. L. Jach, E. Katragkou, N. de Noblet-Ducoudré, K. Radtke, M. Raffa, P. M. M. Soares, G. Sofiadis, S. Strada, G. Strandberg, M. H. Tölle, K. Warrach-Sagi & V. Wulfmeyer (2019). „Biogeophysical impacts of forestation in Europe: First results from the LUCAS Regional Climate Model intercomparison“. In: *Earth System Dynamics Discussions* 11.1, S. 1–31. DOI: 10.5194/esd-2019-4.
- Davin, E. L., R. Stöckli, E. B. Jaeger, S. Levis & S. I. Seneviratne (2011). „COSMO-CLM2: A new version of the COSMO-CLM model coupled to the Community Land Model“. In: *Climate Dynamics* 37.9-10, S. 1889–1907. DOI: 10.1007/s00382-011-1019-z.
- de Boer, F. (2016). „HiHydroSoil: A high resolution soil map of hydraulic properties: Version 1.2“. In: *Future Water Report* 134.

- de Vries, D. A. (1952). „The Thermal Conductivity of Soil“. In: *Mededelingen van de Landbouwhogeschool te Wageningen* 52.1, S. 1–73.
- de Vries, D. A. (1963). „Thermal Properties of Soils“. In: *Physics of Plant Environment*. Hrsg. von W. R. van Wijk. New York: Wiley, S. 210–235.
- de Vries, D. A. (1975). „Heat Transferr in Soils“. In: *Transfer processes in the plant environment*. Hrsg. von D. A. de Vries & N. H. Afgan. Advances in thermal engineering. Washington, D.C.: Scripta Bk. Co, S. 5–28.
- de Vries, D. A. & N. H. Afgan, Hrsg. (1975). *Transfer processes in the plant environment*. Advances in thermal engineering. Washington, D.C.: Scripta Bk. Co.
- Deardorff, J. W. (1978). „Efficient prediction of ground surface temperature and moisture, with inclusion of a layer of vegetation“. In: *Journal of Geophysical Research* 83.C4, S. 1889. DOI: 10.1029/JC083IC04P01889.
- Decharme, B., A. Boone, C. Delire & J. Noilhan (2011). „Local evaluation of the Interaction between Soil Biosphere Atmosphere soil multilayer diffusion scheme using four pedotransfer functions“. In: *Journal of Geophysical Research* 116.D20.
- Dee, D. P., S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo & d. P. Bauer (2011). „The ERA–Interim reanalysis: Configuration and performance of the data assimilation system“. In: *Quarterly Journal of the Royal Meteorological Society* 137.656, S. 553–597.
- Despotovic, M., V. Nedic, D. Despotovic & S. Cvetanovic (2016). „Evaluation of empirical models for predicting monthly mean horizontal diffuse solar radiation“. In: *Renewable and Sustainable Energy Reviews* 56, S. 246–260. DOI: 10.1016/j.rser.2015.11.058.
- Deutsches Klimarechenzentrum (1993). *Report No. 06: The ECHAM3 Atmospheric General Circulation Model*.
- Deutschländer, T. & B. Wichura (2006). „Das Münsterländer Schneechaos am 1. Adventswochenende 2005“. In: *Klimawandel in Deutschland - Hurrikane im Nordatlantik - Aktuelle Ergebnisse des Klimamonitorings*. Hrsg. von DWD. Klimastatusbericht. Offenbach am Main: Deutscher Wetterdienst.
- Dharumarajan, S., R. Hegde, M. Lalitha, B. Kalaiselvi & S. K. Singh (2019). „Pedotransfer Functions for Predicting Soil Hydraulic Properties in Semi-Arid Regions of Karnataka Plateau, India“. In: *Current Science* 116.7, S. 1237. DOI: 10.18520/cs/v116/i7/1237-1246.
- Di Giuseppe, F., D. Cesari & G. Bonafé (2011). „Soil Initialization Strategy for Use in Limited-Area Weather Prediction Systems“. In: *Monthly Weather Review* 139.6, S. 1844–1860. DOI: 10.1175/2011MWR3279.1.
- Di Gregorio, A. (2005). *Land cover classification system software version 2: Based on the orig. software version 1*. rev. Bd. 8. Environment and natural resources series Geo-spatial data and information. Rome.
- Di Luca, A., R. de Elía & R. Laprise (2015). „Challenges in the Quest for Added Value of Regional Climate Dynamical Downscaling“. In: *Current Climate Change Reports* 1.1, S. 10–21. DOI: 10.1007/s40641-015-0003-9.
- Diaconescu, E. P., P. Gachon & R. Laprise (2015). „On the Remapping Procedure of Daily Precipitation Statistics and Indices Used in Regional Climate Model Evaluation“. In: *Journal of Hydrometeorology* 16.6, S. 2301–2310. DOI: 10.1175/JHM-D-15-0025.1.
- Dickinson, R. E. (1995). „Land processes in climate models“. In: *Remote Sensing of Environment* 51.1, S. 27–38. DOI: 10.1016/0034-4257(94)00062-R.
- Dickinson, R. E., A. Henderson-Sellers, P. Kennedy & M. Wilson (1986). *Biosphere-atmosphere Transfer Scheme (BATS) for the NCAR Community Climate Model*. DOI: 10.5065/D6668B58.
- Dijkerman, J. C. (1988). „An Ustult-Aquult-Tropept catena in Sierra Leone, West Africa, II. Land qualities and land evaluation“. In: *Geoderma* 42.1, S. 29–49. DOI: 10.1016/0016-7061(88)90021-3.
- Dobler, A., H. Feldmann, U. Ulbrich, D. Jacob & C. Kottmeier (2017). „Grenzen und Herausforderungen der regionalen Klimamodellierung“. In: *Klimawandel in Deutschland*. Hrsg. von G. Brasseur, D. Jacob & S. Schuck-Zöller. Berlin: Springer Spektrum, S. 37–44. DOI: 10.1007/978-3-662-50397-3\_5.

- Doms, G., J. Förstner, E. Heise, H.-J. Herzog, D. Mironov, M. Raschendorfer, T. Reinhardt, B. Ritter, R. Schrodin, J.-P. Schulz & G. Vogel (2011). „A Description of the Nonhydrostatic Regional COSMO Model: Part II: Physical Parameterization“. In: *Consortium for Small-Scale Modelling (COSMO)*.
- Dong, Y., J. S. McCartney & N. Lu (2015). „Critical Review of Thermal Conductivity Models for Unsaturated Soils“. In: *Geotechnical and Geological Engineering* 33.2, S. 207–221. DOI: 10.1007/s10706-015-9843-2.
- Douglas, D. H. (1974). *COLLECTED ALGORITHMS, Cambridge MA: Harvard Laboratory for Computer Graphics*.
- Droogers, P. (2011). *Global Soil Map of Hydraulic Properties: Quantitative Global Soil Map*.
- Duan, S.-B., Z.-L. Li, B.-H. Tang, H. Wu, L. Ma, E. Zhao & C. Li (2013). „Land surface reflectance retrieval from hyperspectral data collected by an unmanned aerial vehicle over the Baotou test site“. In: *PloS one* 8.6, e66972. DOI: 10.1371/journal.pone.0066972.
- Dufresne, J.-L. et al. (2013). „Climate change projections using the IPSL-CM5 Earth System Model: From CMIP3 to CMIP5“. In: *Climate Dynamics* 40.9-10, S. 2123–2165. DOI: 10.1007/s00382-012-1636-1.
- Dümenil, L. & E. Todini (1992). „A rainfall-runoff scheme for use in the Hamburg climate model“. In: *Advances in Theoretical Hydrology: A tribute to James Dooge*. Elsevier Science Publishers BV, S. 129–157. DOI: 10.1016/B978-0-444-89831-9.50016-8.
- Durner, W. & S. C. Iden (2015). „Skript Bodenphysikalische Versuche, Institut für Geoökologie, Abteilung Bodenkunde und Bodenphysik“. Skript. TU Braunschweig.
- DWD, Hrsg. (2004). *Der Hitzesommer 2003 - Vier Milliarden Jahre Klimageschichte im Überblick - Aktuelle Ergebnisse des Klimamonitorings*. Bd. 2003. Klimastatusbericht. Offenbach: Dt. Wetterdienst.
- DWD (2006). *Deutschlandwetter im Winter 2005/2006: Winter war knackig kalt, sehr trocken und sonnig*. URL: [https://www.dwd.de/DE/presse/pressemitteilungen/DE/2006/20060228\\_deutschlandwetter\\_winter.pdf?\\_\\_blob=publicationFile&v=2](https://www.dwd.de/DE/presse/pressemitteilungen/DE/2006/20060228_deutschlandwetter_winter.pdf?__blob=publicationFile&v=2).
- DWD, Hrsg. (2007). *Aktuelle Ergebnisse des Klimamonitorings*. Bd. 2006. Klimastatusbericht. Offenbach am Main: Deutscher Wetterdienst.
- DWD (2018). *DATENSATZBESCHREIBUNG: Tägliche Raster der mittleren Bodentemperatur in 5 cm Tiefe bei unbewachsenem Boden*. URL: [https://opendata.dwd.de/climate\\_environment/CDC/grids\\_germany/multi\\_annual/soil\\_temperature\\_5cm/BESCHREIBUNG\\_gridsgermany\\_multi\\_annual\\_soil\\_temperature\\_5cm\\_de.pdf](https://opendata.dwd.de/climate_environment/CDC/grids_germany/multi_annual/soil_temperature_5cm/BESCHREIBUNG_gridsgermany_multi_annual_soil_temperature_5cm_de.pdf).
- Earth Resources Observation And Science (EROS) Center (1994). *Global Land Cover Characterization Dataset (GLCCD)*. DOI: 10.5066/F7GB230D.
- Earth Resources Observation And Science (EROS) Center (1999). *Global 30 Arc-Second Elevation (GTOPO30)*. DOI: 10.5066/F7DF6PQS.
- Earth Resources Observation And Science (EROS) Center (2017). *Shuttle Radar Topography Mission (SRTM) Void Filled*. DOI: 10.5066/f7f76b1x.
- Ebita, A., S. Kobayashi, Y. Ota, M. Moriya, R. Kumabe, K. Onogi, Y. Harada, S. Yasui, K. Miyaoka, K. Takahashi, H. Kamahori, C. Kobayashi, H. Endo, M. Soma, Y. Oikawa & T. Ishimizu (2011). „The Japanese 55-year Reanalysis “JRA-55”: An Interim Report“. In: *SOLA* 7, S. 149–152. DOI: 10.2151/sola.2011-038.
- ECMWF (2016). *IFS Documentation CY41R2 - Part IV: Physical Processes*. DOI: 10.21957/tr5rv27xu.
- Edwards, P. N. (2011). „History of climate modeling“. In: *Wiley Interdisciplinary Reviews: Climate Change* 2.1, S. 128–139. DOI: 10.1002/wcc.95.
- Eggert, B. (2011). „Auswirkungen der Oberflächeneigenschaften in REMO auf die Simulation der unteren Atmosphäre“. Diplomarbeit. Universität Hamburg.
- Eidenshink, J. C. & J. L. Faundeen (1994). „The 1 km AVHRR global land data set: first stages in implementation“. In: *International Journal of Remote Sensing* 15.17, S. 3443–3462. DOI: 10.1080/01431169408954339.

- Ek, M. B., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno & J. D. Tarpley (2003). „Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model“. In: *Journal of Geophysical Research* 108.D22.
- Enßle, F., J. Heinzel & B. Koch (2012). *Evaluating Height Differences Between Global Digital Surface Models and ICESAT Heights at Footprint Geolocation*. January 23. – 25., 2012, Ostrava.
- FAO (1988). „FAO/UNESCO soil map of the world: Revised Legend“. In: *World Soil Resources Rep. 60, Rome, Italy*.
- FAO (1995). *Digital Soil Map of the World, Version 3.5*.
- FAO (2012). *Harmonized World Soil Database (version 1.2). FAO/IIASA/ISRIC/ISSCAS/JRC*.
- FAO-UNESCO (1971-1981). *Soil Map of the World, 1:5,000,000, Volumes I-X. UNESCO, Paris*.
- Farouki, O. T. (1981a). „The thermal properties of soils in cold regions“. In: *Cold Regions Science and Technology* 5.1, S. 67–75. DOI: 10.1016/0165-232X(81)90041-0.
- Farouki, O. T. (1981b). „Thermal properties of soils“. In: *CRREL monograph (Cold Regions Research and Engineering Laboratory)* 81.1.
- Farr, T. G., P. A. Rosen, E. Caro, R. Crippen, R. Duren, S. Hensley, M. Kobrick, M. Paller, E. Rodriguez, L. Roth, D. Seal, S. Shaffer, J. Shimada, J. Umland, M. Werner, M. Oskin, D. Burbank & D. Alsdorf (2007). „The Shuttle Radar Topography Mission“. In: *Reviews of Geophysics* 45.2, S. 1485. DOI: 10.1029/2005RG000183.
- Farrugia, P. S. & A. Micallef (2006). „Comparative analysis of estimators for wind direction standard deviation“. In: *Meteorological Applications* 13.01, S. 29. DOI: 10.1017/S1350482705001982.
- Fatichi, S., D. Or, R. Walko, H. Vereecken, M. H. Young, T. A. Ghezzehei, T. Hengl, S. Kollet, N. Agam & R. Avissar (2020). „Soil structure is an important omission in Earth System Models“. In: *Nature communications* 11.1, S. 522. DOI: 10.1038/s41467-020-14411-z.
- Finnern, H., Hrsg. (1994). *Bodenkundliche Kartieranleitung: Mit 91 Tabellen*. 4., verb. und erw. Aufl. Stuttgart: Schweizerbart.
- Fonseca, J. (2008-2017). <https://github.com/jrfonseca/gprof2dot> (2020): *Generate a dot graph from the output of several profilers*. URL: <https://github.com/jrfonseca/gprof2dot>.
- Forkuor, G. & B. Maathuis (2012). „Comparison of SRTM and ASTER Derived Digital Elevation Models over Two Regions in Ghana - Implications for Hydrological and Environmental Modeling“. In: *Studies on Environmental and Applied Geomorphology*. Hrsg. von T. Piacentini. InTech, S. 219–240. DOI: 10.5772/28951.
- Fortak, H. (1982). *Meteorologie*. 2. Aufl. Berlin: Reime.
- Fotso-Nguemo, T. C., D. A. Vondou, W. M. Pokam, Z. Y. Djomou, I. Diallo, A. Haensler, L. A. D. Tchotchou, P. H. Kamsu-Tamo, A. T. Gaye & C. Tchawoua (2017). „On the added value of the regional climate model REMO in the assessment of climate change signal over Central Africa“. In: *Climate Dynamics* 49.11-12, S. 3813–3838. DOI: 10.1007/s00382-017-3547-7.
- Fotso-Nguemo, T. C., D. A. Vondou, C. Tchawoua & A. Haensler (2017). „Assessment of simulated rainfall and temperature from the regional climate model REMO and future changes over Central Africa“. In: *Climate Dynamics* 48.11-12, S. 3685–3705. DOI: 10.1007/s00382-016-3294-1.
- Franklin, W. R. (2018). *PNPOLY - Point Inclusion in Polygon Test W. Randolph Franklin (WRF)*. URL: [https://wrf.ecse.rpi.edu/Research/Short\\_Notes/pnpoly.html](https://wrf.ecse.rpi.edu/Research/Short_Notes/pnpoly.html).
- Fröb, K. (2011). „Measuring and modeling of soil thermal properties and ground heat flux at two different sites at Lena Delta, Siberia“. Diplom Thesis. Leipzig: Universität Leipzig.
- Fu, W., J. Ma, P. Chen & F. Chen (2020). „Remote Sensing Satellites for Digital Earth“. In: *Manual of Digital Earth*. Hrsg. von H. Guo, M. F. Goodchild & A. Annoni. Singapore: Springer Singapore, S. 55–123. DOI: 10.1007/978-981-32-9915-3\_3.
- Gálos, B. (2010). „Analysis of forest-climate interactions, applying the regional climate model REMO“. Dissertation. University of West Hungary.

- Gao, F., C. B. Schaaf, A. H. Strahler, A. Roesch, W. Lucht & R. Dickinson (2005). „MODIS bidirectional reflectance distribution function and albedo Climate Modeling Grid products and the variability of albedo for major global vegetation types“. In: *Journal of Geophysical Research* 110.D1. DOI: 10.1029/2004JD005190.
- Gao, L., M. Bernhardt & K. Schulz (2012). „Elevation correction of ERA-Interim temperature data in complex terrain“. In: *Hydrology and Earth System Sciences* 16.12, S. 4661–4673. DOI: 10.5194/hess-16-4661-2012.
- Gao, L. (2013). „Validation and statistical downscaling of ERA-Interim reanalysis data for integrated applications“. Dissertation. Ludwig–Maximilians–Universität München.
- Gao, Y., S. Weiher, T. Markkanen, J.-P. Pietikäinen, H. Gregow, H. M. Henttonen, D. Jacob & A. Laaksonen (2015). „Implementation of the CORINE land use classification in the regional climate model remo“. In: *Boreal Environment Research* 20, S. 261–282.
- García-Herrera, R., J. Díaz, R. M. Trigo, J. Luterbacher & E. M. Fischer (2010). „A Review of the European Summer Heat Wave of 2003“. In: *Critical Reviews in Environmental Science and Technology* 40.4, S. 267–306. DOI: 10.1080/10643380802238137.
- Garratt, J. R. (1992). *The atmospheric boundary layer*.
- Gee, G. W. & D. Or (2002). „2.4 Particle-Size Analysis“. In: *Methods of Soil Analysis*. Hrsg. von J. H. Dane & G. Clarke Topp. SSSA Book Series. Madison, WI, USA: Soil Science Society of America, S. 255–293. DOI: 10.2136/sssabookser5.4.c12.
- Gesch, D. B., K. L. Verdin & S. K. Greenlee (1999). „New land surface digital elevation model covers the Earth“. In: *Eos, Transactions American Geophysical Union* 80.6, S. 69. DOI: 10.1029/99EO00050.
- Ghanbarian, B. & H. Millán (2010). „Point pedotransfer functions for estimating soil water retention curve“. In: *International Agrophysics* 24, S. 243–251.
- Giorgi, F. (1990). „Simulation of regional climate using a limited area model nested in a general circulation model“. In: *Journal of Climate* 3.9, S. 941–963.
- Giorgi, F. (2006). „Regional climate modeling: Status and perspectives“. In: *Journal de Physique IV (Proceedings)* 139.1, S. 101–118. DOI: 10.1051/jp4:2006139008.
- Giorgi, F. (2019). „Thirty Years of Regional Climate Modeling: Where Are We and Where Are We Going next?“ In: *Journal of Geophysical Research: Atmospheres*. DOI: 10.1029/2018JD030094.
- Giorgi, F. & L. O. Mearns (1999). „Introduction to special section: Regional Climate Modeling Revisited“. In: *Journal of Geophysical Research* 104.D6, S. 6335–6352. DOI: 10.1029/98JD02072.
- Global Soil Data Task Group (2000). *Global Soil Data Products CD-ROM (IGBP-DIS)*. CD-ROM. International Geosphere-Biosphere Programme, Data and Information System, Potsdam, Germany.
- Gogoi, P. P., V. Vinoj, D. Swain, G. Roberts, J. Dash & S. Tripathy (2019). „Land use and land cover change effect on surface temperature over Eastern India“. In: *Scientific reports* 9.1, S. 8859. DOI: 10.1038/s41598-019-45213-z.
- Golmohammadi, G., S. Prasher, A. Madani & R. Rudra (2014). „Evaluating Three Hydrological Distributed Watershed Models: MIKE-SHE, APEX, SWAT“. In: *Hydrology* 1.1, S. 20–39. DOI: 10.3390/hydrology1010020.
- Göttel, H. (2009). „Einfluss der nichthydrostatischen Modellierung und der Niederschlagsverdriftung auf die Ergebnisse regionaler Klimamodellierung“. In: *Berichte zur Erdsystemforschung* 60. DOI: 10.17617/2.994076.
- Grohmann, C. H. (2018). „Evaluation of TanDEM-X DEMs on selected Brazilian sites: Comparison with SRTM, ASTER GDEM and ALOS AW3D30“. In: *Remote Sensing of Environment* 212, S. 121–133. DOI: 10.1016/j.rse.2018.04.043.
- Guimberteau, M. et al. (2018). „ORCHIDEE-MICT (v8.4.1), a land surface model for the high latitudes: model description and validation“. In: *Geoscientific Model Development* 11.1, S. 121–163. DOI: 10.5194/gmd-11-121-2018.

- Hagemann, S. (2002). *An Improved Land Surface Parameter Dataset for Global and Regional Climate Models*.
- Hagemann, S., M. Botzet, L. Dümenil & B. Machenhauer (1999). *DERIVATION OF GLOBAL GCM BOUNDARY CONDITIONS FROM 1 KM LAND USE SATELLITE DATA*.
- Hagemann, S. & L. D. Gates (2003). „Improving a subgrid runoff parameterization scheme for climate models by the use of high resolution data derived from satellite observations“. In: *Climate Dynamics* 21.3-4, S. 349–359. DOI: 10.1007/s00382-003-0349-x.
- Hagemann, S. & T. Stacke (2015). „Impact of the soil hydrology scheme on simulated soil moisture memory“. In: *Climate Dynamics* 44.7-8, S. 1731–1750. DOI: 10.1007/s00382-014-2221-6.
- Haigh, S. K. (2012). „Thermal conductivity of sands“. In: *Géotechnique* 62.7, S. 617–625. DOI: 10.1680/geot.11.P.043.
- Hänsler, A., N. Koldunov, D. Sein, W. Sauf & D. Jacob (2016). „Very high resolution simulations of African climate with the Regional Climate Model REMO“. In: *NIC Symposium Proceedings*. Hrsg. von K. Binder, M. Müller, M. Kremer & A. Schnurpfeil. NIC Series. John von Neumann-Institut für Computing, S. 291–298.
- Hänsler, A. (2011). „Impact of climate change on the Coastal Climate of South-Western Africa“. Dissertation. University of Hamburg.
- Harris, I. C., P. D. Jones, T. J. Osborn & D. H. Lister (2014). „Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset“. In: *International Journal of Climatology* 34.3, S. 623–642. DOI: 10.1002/joc.3711.
- Harris, I. C., T. J. Osborn, P. Jones & D. Lister (2020). „Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset“. In: *Scientific data* 7.1, S. 109. DOI: 10.1038/s41597-020-0453-3.
- Haylock, M. R., N. Hofstra, A. M. G. Klein Tank, E. J. Klok, P. D. Jones & M. New (2008). „A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006“. In: *Journal of Geophysical Research* 113.D20. DOI: 10.1029/2008jd010201.
- He, H., K. Noborio, Ø. Johansen, M. Dyck & J. Lv (2019). „Normalized concept for modelling effective soil thermal conductivity from dryness to saturation“. In: *European Journal of Soil Science*. DOI: 10.1111/ejss.12820.
- He, H., D. He, J. Jin, K. M. Smits, M. Dyck, Q. Wu, B. Si & J. Lv (2020). „Room for improvement: A review and evaluation of 24 soil thermal conductivity parameterization schemes commonly used in land-surface, hydrological, and soil-vegetation-atmosphere transfer models“. In: *Earth-Science Reviews* 211, S. 103419. DOI: 10.1016/j.earscirev.2020.103419.
- Heck, B. & K. Seitz (2016). „Molodenski – quo vadis?“ In: *Handbuch der Geodäsie*. Hrsg. von W. Freeden & R. Rummel. Berlin, Heidelberg: Springer Berlin Heidelberg, S. 1–32. DOI: 10.1007/978-3-662-46900-2\_14-1.
- Hedderich, J. & L. Sachs (2018). *Angewandte Statistik*. Berlin, Heidelberg: Springer Berlin Heidelberg. DOI: 10.1007/978-3-662-56657-2.
- Henderson-Sellers, A. & R. E. Dickinson (1993). „Atmospheric-land surface fluxes“. In: *Modelling change in environmental systems*. Hrsg. von A. J. Jakeman, M. B. Beck & M. J. McAleer. John Wiley and Sons, S. 387–405.
- Hengl, T., J. M. de Jesus, R. A. MacMillan, N. H. Batjes, G. B. M. Heuvelink, E. Ribeiro, A. Samuel-Rosa, B. Kempen, J. G. B. Leenaars, M. G. Walsh & M. R. Gonzalez (2014). „SoilGrids1km—global soil information based on automated mapping“. In: *PloS one* 9.8, e105992. DOI: 10.1371/journal.pone.0105992.
- Hengl, T., J. Mendes de Jesus, G. B. M. Heuvelink, M. Ruiperez Gonzalez, M. Kilibarda, A. Blagotić, W. Shanguan, M. N. Wright, X. Geng, B. Bauer-Marschallinger, M. A. Guevara, R. Vargas, R. A. MacMillan, N. H. Batjes, J. G. B. Leenaars, E. Ribeiro, I. Wheeler, S. Mantel & B. Kempen (2017). „SoilGrids250m: Global gridded soil information based on machine learning“. In: *PloS one* 12.2, e0169748. DOI: 10.1371/journal.pone.0169748.



- Herrmann, F., L. Hübsch, J. Elbracht, N. Engel, L. Keller, R. Kunkel, U. Müller, H. Röhm, H. Vereecken & F. Wendland (2017). „Mögliche Auswirkungen von Klimaänderungen auf die Grundwasserneubildung in Niedersachsen“. In: *Hydrologie und Wasserbewirtschaftung* 61.4, S. 244–260.
- Hersbach, H. et al. (2018). *Operational global reanalysis: progress, future directions and synergies with NWP*. DOI: 10.21957/tkic6g3wm.
- Hersbach, H. et al. (2020). „The ERA5 global reanalysis“. In: *Quarterly Journal of the Royal Meteorological Society* 146.730, S. 1999–2049. DOI: 10.1002/qj.3803.
- Hillel, D. (1980). *Applications of Soil Physics*.
- Hillel, D. (2003). *Introduction to Environmental Soil Physics*. 1. Aufl. s.l.: Elsevier professional.
- Hirt, C., M. S. Filmer & W. E. Featherstone (2010). „Comparison and validation of the recent freely available ASTER-GDEM ver1, SRTM ver4.1 and GEODATA DEM-9S ver3 digital elevation models over Australia“. In: *Australian Journal of Earth Sciences* 57.3, S. 337–347. DOI: 10.1080/08120091003677553.
- Hofstra, N., M. Haylock, M. New & P. D. Jones (2009). „Testing E-OBS European high-resolution gridded data set of daily precipitation and surface temperature“. In: *Journal of Geophysical Research* 114.D21, S. 1032. DOI: 10.1029/2009JD011799.
- Howat, I. M., A. Negrete & B. E. Smith (2014). „The Greenland Ice Mapping Project (GIMP) land classification and surface elevation data sets“. In: *The Cryosphere* 8.4, S. 1509–1518. DOI: 10.5194/tc-8-1509-2014.
- Hubrechts, L. & J. Feyen (1996). „PEDOTRANSFER FUNCTIONS FOR THERMAL SOIL PROPERTIES“. Diss. Katholieke University Leuven.
- Hvidegaard, S. M., L. Sandberg Sørensen & R. Forsberg (2012). „ASTER GDEM validation using LiDAR data over coastal regions of Greenland“. In: *Remote Sensing Letters* 3.1, S. 85–91. DOI: 10.1080/01431161.2010.527389.
- IGBP (2000). *Global Soil Data Task (IGBP-DIS, ISO-image of CD): International Geosphere-Biosphere Program*. PANGAEA - Data Publisher for Earth & Environmental Science. DOI: 10.1594/PANGAEA.869912.
- Jacob, D. (2001). „A note to the simulation of the annual and inter-annual variability of the water budget over the Baltic Sea drainage basin“. In: *Meteorology and Atmospheric Physics* 77.1-4, S. 61–73. DOI: 10.1007/s007030170017.
- Jacob, D., L. Bärring, O. B. Christensen, J. H. Christensen, M. de Castro, M. Déqué, F. Giorgi, S. Hagemann, M. Hirschi, R. Jones, E. Kjellström, G. Lenderink, B. Rockel, E. Sánchez, C. Schär, S. I. Seneviratne, S. Somot, A. van Ulden & B. van den Hurk (2007). „An inter-comparison of regional climate models for Europe: Model performance in present-day climate“. In: *Climatic Change* 81.S1, S. 31–52. DOI: 10.1007/s10584-006-9213-4.
- Jacob, D., A. Elizalde, A. Haensler, S. Hagemann, P. Kumar, R. Podzun, D. Rechid, A. R. Remedio, F. Saeed, K. Sieck, C. Teichmann & C. Wilhelm (2012). „Assessing the Transferability of the Regional Climate Model REMO to Different COordinated Regional Climate Downscaling EXperiment (CORDEX) Regions“. In: *Atmosphere* 3.4, S. 181–199. DOI: 10.3390/atmos3010181.
- Jacob, D., C. Kottmeier, J. Petersen, D. Rechid & C. Teichmann (2017). „Regionale Klimamodellierung“. In: *Klimawandel in Deutschland*. Hrsg. von G. Brasseur, D. Jacob & S. Schuck-Zöller. Berlin: Springer Spektrum, S. 27–35. DOI: 10.1007/978-3-662-50397-3\_4.
- Jacob, D. & R. Podzun (1997). „Sensitivity studies with the regional climate model REMO“. In: *Meteorology and Atmospheric Physics* 63.1-2, S. 119–129. DOI: 10.1007/BF01025368.
- Jacob, D. et al. (2014). „EURO-CORDEX: new high-resolution climate change projections for European impact research“. In: *Regional Environmental Change* 14.2, S. 563–578. DOI: 10.1007/s10113-013-0499-2.
- Jacob, D. et al. (2020). „Regional climate downscaling over Europe: perspectives from the EURO-CORDEX community“. In: *Regional Environmental Change* 20.2. DOI: 10.1007/s10113-020-01606-9.

- Jia, G., E. Shevliakova, P. Artaxo, N. De Noblet-Ducoudré, R. Houghton, J. House, K. Kitajima, C. Lennard, A. Popp, A. Sirin, R. Sukumar & I. Lerchot (2019). „Land–climate interactions“. In: *Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems*. Hrsg. von P. R. Shukla, J. Skea, E. Calvo Buendia, V. Masson-Delmotte, H. O. Pörtner, D. C. Roberts, P. Zhai, R. Slade, S. Connors & R. van Diemen, S. 131–247.
- Jing, C., A. Shortridge, S. Lin & J. Wu (2013). „Comparison and validation of SRTM and ASTER GDEM for a subtropical landscape in Southeastern China“. In: *International Journal of Digital Earth* 7.12, S. 969–992. DOI: 10.1080/17538947.2013.807307.
- Johansen, O. (1975). „Thermal Conductivity of Soils“. Dissertation. University of Trondheim. DOI: 10.21236/ADA044002.
- Jolliffe, I. T. & D. B. Stephenson, Hrsg. (2012). *Forecast verification: A practitioner’s guide in atmospheric science*. 2. ed. Chichester: Wiley-Blackwell.
- Jones, P. W. (1998). „A User’s Guide for SCRIP: A Spherical Coordinate Remapping and Interpolation Package, Version 1.4“. In: *Theoretical Division, Los Alamos National Laboratory*.
- Kalnay, E. et al. (1996). „The NCEP/NCAR 40-Year Reanalysis Project“. In: *Bulletin of the American Meteorological Society* 77.3, S. 437–471. DOI: 10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.
- Kamat, N. & A. Nandi (2017). „A closer look at variance implementations in modern database systems“. In: *ACM SIGMOD Record* 45.4, S. 28–33.
- Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. J. Hnilo, M. Fiorino & G. L. Potter (2002). „NCEP-DOE AMIP-II REANALYSIS (R-2)“. In: *Bulletin of the American Meteorological Society* 83.11, S. 1631–1644.
- Karstens, U., R. Nolte-Holube & B. Rockel (1996). „Calculation of the water budget over the Baltic Sea catchment area using the regional forecast model REMO for June 1993“. In: *Tellus A: Dynamic Meteorology and Oceanography* 48.5, S. 684–692. DOI: 10.3402/tellusa.v48i5.12171.
- Kawamiya, M., T. Hajima, K. Tachiiri, S. Watanabe & T. Yokohata (2020). „Two decades of Earth system modeling with an emphasis on Model for Interdisciplinary Research on Climate (MIROC)“. In: *Progress in Earth and Planetary Science* 7.1. DOI: 10.1186/s40645-020-00369-5.
- Keller, J. & D. Karmeli (1975). *Trickle irrigation design Parameters: Rain Bird Sprinkler Manufacturing Corporation*. Glendora.
- Kersten, M. S. (1949). *Thermal Properties of Soil*.
- Kiehl, J. T. & K. E. Trenberth (1997). „Earth’s annual global mean energy budget“. In: *Bulletin of the American Meteorological Society* 78.2, S. 197–208.
- Knorr, W. (1998). „Constraining a global mechanistic vegetation model with satellite data“. In: *IEEE International Geoscience and Remote Sensing Symposium*. IEEE. DOI: 10.1109/IGARSS.1998.699643.
- Knorr, W., K.-G. Schnitzler & Y. Govaerts (2001). „The role of Bright Desert Regions in shaping North African climate“. In: *Geophysical Research Letters* 28.18, S. 3489–3492. DOI: 10.1029/2001GL013283.
- Korona, J., E. Berthier, M. Bernard, F. Rémy & E. Thouvenot (2009). „SPIRIT. SPOT 5 stereoscopic survey of Polar Ice: Reference Images and Topographies during the fourth International Polar Year (2007–2009)“. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 64.2, S. 204–212. DOI: 10.1016/j.isprsjprs.2008.10.005.
- Kothe, S., A. Dobler, A. Beck & B. Ahrens (2011). „The radiation budget in a regional climate model“. In: *Climate Dynamics* 36.5-6, S. 1023–1036. DOI: 10.1007/s00382-009-0733-2.
- Kotlarski, S. (2007). „A Subgrid Glacier Parameterisation for Use in Regional Climate Modelling“. In: *Berichte zur Erdsystemforschung* 42.
- Kowalczyk, E. A., Y. P. Wang, R. M. Law, H. L. Davies, J. L. McGregor & G. Abramowitz (2006). „The CSIRO Atmosphere Biosphere Land Exchange (CABLE) model for use in climate models and as an offline model“. In: *CSIRO Marine and Atmospheric Research Paper* 13, S. 42.

- Kracher, D., H.-T. Mengelkamp & T. Foken (2009). „The residual of the energy balance closure and its influence on the results of three SVAT models“. In: *Meteorologische Zeitschrift* 18.6, S. 647–661. DOI: 10.1127/0941-2948/2009/0412.
- Kraus, H. (2004). *Die Atmosphäre der Erde: Eine Einführung in die Meteorologie*. Berlin, Heidelberg: Springer Berlin Heidelberg. DOI: 10.1007/3-540-35017-9.
- Krinner, G., N. Viovy, N. de Noblet-Ducoudré, J. Ogée, J. Polcher, P. Friedlingstein, P. Ciais, S. Sitch & I. C. Prentice (2005). „A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system“. In: *Global Biogeochemical Cycles* 19.1. DOI: 10.1029/2003GB002199.
- Kumar, P., S. Kotlarski, C. Moseley, K. Sieck, H. Frey, M. Stoffel & D. Jacob (2015). „Response of Karakoram–Himalayan glaciers to climate variability and climatic change: A regional climate model assessment“. In: *Geophysical Research Letters* 42.6, S. 1818–1825. DOI: 10.1002/2015GL063392.
- Kumar, P., R. Podzun, S. Hagemann, D. Jacob & R. Podzun (2014). „Impact of modified soil thermal characteristic on the simulated monsoon climate over south Asia“. In: *J. Earth Syst. Sci* 123.1, S. 151–160. DOI: 10.1007/s12040-013-0381-0.
- Kunz, M., S. Mohr, M. Rauthe, R. Lux & C. Kottmeier (2010). „Assessment of extreme wind speeds from Regional Climate Models – Part 1: Estimation of return values and their evaluation“. In: *Natural Hazards and Earth System Sciences* 10.4, S. 907–922. DOI: 10.5194/nhess-10-907-2010.
- Kushwa, V., K. M. Hati, N. K. Sinha, R. K. Singh, M. Mohanty, J. Somasundaram, R. C. Jain, R. S. Chaudhary, A. K. Biswas & A. K. Patra (2016). „Long-term Conservation Tillage Effect on Soil Organic Carbon and Available Phosphorous Content in Vertisols of Central India“. In: *Agricultural Research* 5.4, S. 353–361. DOI: 10.1007/s40003-016-0223-9.
- Kyselý, J. & E. Plavcová (2010). „A critical remark on the applicability of E-OBS European gridded temperature data set for validating control climate simulations“. In: *Journal of Geophysical Research* 115.D23, S. 4211. DOI: 10.1029/2010JD014123.
- Lal, R. (1979). „Physical properties and moisture retention characteristics of some nigerian soils“. In: *Geoderma* 21.3, S. 209–223. DOI: 10.1016/0016-7061(78)90028-9.
- Larsén, X. G., J. Mann, J. Berg, H. Göttel & D. Jacob (2010). „Wind climate from the regional climate model REMO“. In: *Wind Energy* 13.4, S. 279–296. DOI: 10.1002/we.337.
- Lawrence, D. M. & A. G. Slater (2008). „Incorporating organic soil into a global climate model“. In: *Climate Dynamics* 30.2-3, S. 145–160. DOI: 10.1007/s00382-007-0278-1.
- Lawrence, D. M. et al. (2019). „The Community Land Model Version 5: Description of New Features, Benchmarking, and Impact of Forcing Uncertainty“. In: *Journal of Advances in Modeling Earth Systems* 11.12, S. 4245–4287. DOI: 10.1029/2018MS001583.
- le Moigne, P. (2009). „SURFEX scientific documentation“. In: *Note de centre (CNRM/GMME), Météo-France, Toulouse, France*.
- Le Treut, H., R. Somerville, U. Cubasch, Y. Ding, C. Mauritzen, A. Mokssit, T. Peterson & M. Prather (2007). „Chapter 1: Historical Overview of Climate Change“. In: *IPCC fourth assessment report (AR4)*. Hrsg. von S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. Tignor & H. Miller, S. 93–127.
- Leenaars, J. G. B., T. Hengl, M. R. Gonzalez, J. Mendes de Jesus, G. B. M. Heuvelink, J. Wolf, van Bussel, L. G. J., L. Claessens, H. Yang & K. G. Cassman (2015). „Root zone plant-available water holding capacity of the Sub-Saharan Africa Soil: Gridded functional soil information (dataset RZ-PAWHC SSA version 1.0): Collaboration project of Africa Soil Information Service and Global Yield Gap and Water Productivity Atlas. ISRIC - World Soil Information“. In: *ISRIC Report 2*. DOI: 10.13140/RG.2.1.3950.9209.
- Legates, D. R. & C. J. Willmott (1990). „Mean seasonal and spatial variability in global surface air temperature“. In: *Theoretical and Applied Climatology* 41.1-2, S. 11–21.

- Lemoine, F. G., S. C. Kenyon, J. K. Factor, R. G. Trimmer, N. K. Pavlis, D. S. Chinn, C. M. Cox, S. M. Klosko, S. B. Luthcke, M. H. Torrence, Y. M. Wang, R. G. Williamson, E. C. Pavlis, R. H. Rapp & T. R. Olson (1998). *The Development of the Joint NASA GSFC and the National Imagery and Mapping Agency (NIMA) Geopotential Model EGM96*. Hrsg. von National Aeronautics and Space Administration. NASA Goddard Space Flight Cent., Washington, D. C.
- Letts, M. G., N. T. Roulet, N. T. Comer, M. R. Skarupa & D. L. Versegny (2000). „Parametrization of peatland hydraulic properties for the Canadian land surface scheme“. In: *Atmosphere-Ocean* 38.1, S. 141–160. DOI: 10.1080/07055900.2000.9649643.
- Li, M., P. Wu & Z. Ma (2020). „A comprehensive evaluation of soil moisture and soil temperature from third-generation atmospheric and land reanalysis data sets“. In: *International Journal of Climatology* 40.13, S. 5744–5766. DOI: 10.1002/joc.6549.
- Li, P., C. Shi, Z. Li, J.-P. Muller, J. Drummond, X. Li, T. Li, Y. Li & J. Liu (2012). „Evaluation of ASTER GDEM using GPS benchmarks and SRTM in China“. In: *International Journal of Remote Sensing* 34.5, S. 1744–1771. DOI: 10.1080/01431161.2012.726752.
- Li, R., L. Zhao, T. Wu, Q. Wang, Y. Ding, J. Yao, X. Wu, G. Hu, Y. Xiao, Y. Du, X. Zhu, Y. Qin, S. Yang, R. Bai, E. Du, G. Liu, D. Zou, Y. Qiao & J. Shi (2019). „Soil thermal conductivity and its influencing factors at the Tanggula permafrost region on the Qinghai–Tibet Plateau“. In: *Agricultural and Forest Meteorology* 264, S. 235–246. DOI: 10.1016/j.agrformet.2018.10.011.
- Li, R., L. Zhao, T. Wu, X. Wu, Y. Xiao, Y. Du & Y. Qin (2016). „The impacts of net long-wave radiation on the surface soil thermal regimes over the Qinghai–Tibetan Plateau, China“. In: *Environmental Earth Sciences* 75.3. DOI: 10.1007/s12665-015-4861-1.
- Likos, W. J. (2015). „Pore-Scale Model for Thermal Conductivity of Unsaturated Sand“. In: *Geotechnical and Geological Engineering* 33.2, S. 179–192. DOI: 10.1007/s10706-014-9744-9.
- Lilliefors, H. W. (1967). „On the Kolmogorov-Smirnov test for normality with mean and variance unknown“. In: *Journal of the American statistical Association* 62.318, S. 399–402.
- Lloyd, E. A., M. Bukovsky & L. O. Mearns (2020). „An analysis of the disagreement about added value by regional climate models“. In: *Synthese*. DOI: 10.1007/s11229-020-02821-x.
- Lohmann, U. & E. Roeckner (1996). „Design and performance of a new cloud microphysics scheme developed for the ECHAM general circulation model“. In: *Climate Dynamics* 12.8, S. 557–572. DOI: 10.1007/BF00207939.
- Lorenz, P. (2011). „Einfluss der Zwei-Wege Nestung auf die Ergebnisse globaler und regionaler Klimasimulationen“. In: *Berichte zur Erdsystemforschung* 104.
- Louis, J.-F. (1979). „A parametric model of vertical eddy fluxes in the atmosphere“. In: *Boundary Layer Meteorology*, S. 187–202.
- Lu, S., T. Ren, Y. Gong & R. Horton (2007). „An Improved Model for Predicting Soil Thermal Conductivity from Water Content at Room Temperature“. In: *Soil Science Society of America Journal* 71.1, S. 8–14. DOI: 10.2136/sssaj2006.0041.
- Lu, Y., S. Lu, R. Horton & T. Ren (2014). „An Empirical Model for Estimating Soil Thermal Conductivity from Texture, Water Content, and Bulk Density“. In: *Soil Science Society of America Journal* 78.6, S. 1859. DOI: 10.2136/sssaj2014.05.0218.
- Lucht, W., C. B. Schaaf & A. H. Strahler (2000). „An algorithm for the retrieval of albedo from space using semiempirical BRDF models“. In: *IEEE Transactions on Geoscience and Remote Sensing* 38.2, S. 977–998. DOI: 10.1109/36.841980.
- Lundquist, J. D. & D. R. Cayan (2007). „Surface temperature patterns in complex terrain: Daily variations and long-term change in the central Sierra Nevada, California“. In: *Journal of Geophysical Research* 112.D11. DOI: 10.1029/2006JD007561.

- Lutz, F., T. Herzfeld, J. Heinke, S. Rolinski, S. Schaphoff, W. von Bloh, J. J. Stoorvogel & C. Müller (2019). „Simulating the effect of tillage practices with the global ecosystem model LPJmL (version 5.0-tillage)“. In: *Geoscientific Model Development* 12.6, S. 2419–2440. DOI: 10.5194/gmd-12-2419-2019.
- Majewski, D. (1991). „The Europa-Modell of the Deutscher Wetterdienst“. In: *ECMWF Seminar on Numerical Methods in Atmospheric Models 2*.
- Manabe, S. (1969). „CLIMATE AND THE OCEAN CIRCULATION: The Atmospheric Circulation and the Hydrology of the Earth’s Surface“. In: *Monthly Weather Review* 97.11, S. 739–774. DOI: 10.1175/1520-0493(1969)097<0806:CATOC>2.3.CO;2.
- Maraun, D. & M. Widmann (2018). *Statistical downscaling and bias correction for climate research*. Cambridge University Press.
- Markert, A. (2012). „Einfluss von Substrateigenschaften auf die Wärmeleitfähigkeit von Böden: Bodenphysikalische Nutzungsaspekte von Erdwärmekollektoren“. Diplomarbeit. TU-Berlin.
- Márquez, J. M. A., M. Á. M. Bohórquez & S. G. Melgar (2016). „Ground Thermal Diffusivity Calculation by Direct Soil Temperature Measurement. Application to very Low Enthalpy Geothermal Energy Systems“. In: *Sensors (Basel, Switzerland)* 16.3, S. 306. DOI: 10.3390/s16030306.
- Marron, J. S. (1994). „Visual Understanding of Higher-Order Kernels“. In: *Journal of Computational and Graphical Statistics* 3.4, S. 447. DOI: 10.2307/1390905.
- Martens, B., D. L. Schumacher, H. Wouters, J. Muñoz-Sabater, N. E. C. Verhoest & D. G. Miralles (2020). „Evaluating the land-surface energy partitioning in ERA5“. In: *Geoscientific Model Development* 13.9, S. 4159–4181. DOI: 10.5194/gmd-13-4159-2020.
- Masson, V., J.-L. Champeaux, F. Chauvin, C. Meriguet & R. Lacazze (2003). „A Global Database of Land Surface Parameters at 1-km Resolution in Meteorological and Climate Models“. In: *Journal of Climate* 16.9, S. 1261–1282. DOI: 10.1175/1520-0442(2003)16<1261:AGDOLS>2.0.CO;2.
- Matte, D., R. Laprise & J. M. Thériault (2016). „Comparison between high-resolution climate simulations using single- and double-nesting approaches within the Big-Brother experimental protocol“. In: *Climate Dynamics* 47.12, S. 3613–3626. DOI: 10.1007/s00382-016-3031-9.
- Maurer, E. P., A. W. Wood, J. C. Adam, D. P. Lettenmaier & B. Nijssen (2002). „A Long-Term Hydrologically Based Dataset of Land Surface Fluxes and States for the Conterminous United States\*“. In: *Journal of Climate* 15.22, S. 3237–3251. DOI: 10.1175/1520-0442(2002)015<3237:ALTHBD>2.0.CO;2.
- McCornack, R. L. (1965). „Extended Tables of the Wilcoxon Matched Pair Signed Rank Statistic“. In: *Journal of the American statistical Association* 60.311, S. 864–871. DOI: 10.1080/01621459.1965.10480835.
- Mccumber, M. C. & R. A. Pielke (1981). „Simulation of the effects of surface fluxes of heat and moisture in a mesoscale numerical model: 1. Soil layer“. In: *Journal of Geophysical Research* 86.C10, S. 9929–9938.
- McHugh, M. L. (2012). „Interrater reliability: the kappa statistic“. In: *Biochemia medica: Biochemia medica* 22.3, S. 276–282.
- Medeiros, J. C., M. Cooper, J. Dalla Rosa, M. Grimaldi & Y. Coquet (2014). „Assessment of pedotransfer functions for estimating soil water retention curves for the amazon region“. In: *Revista Brasileira de Ciência do Solo* 38.3, S. 730–743. DOI: 10.1590/S0100-06832014000300005.
- Mesinger, F. & A. Arakawa (1976). „Numerical Methods Used In Atmospheric Models“. In: *Global Atmospheric Research Programme (GARP) Publications Series* 17.1.
- Meyer, D. J., T. Tachikawa, M. Abrams, R. Crippen, T. Krieger, D. B. Gesch & C. C. Carabajal (2012). „SUMMARY OF THE VALIDATION OF THE SECOND VERSION OF THE ASTER GDEM“. In: *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XXXIX-B4, S. 291–293. DOI: 10.5194/isprsarchives-XXXIX-B4-291-2012.
- Mikhailsoy, F. D. & E. V. Shein (2019). „Comparison of four classical algorithms to determine the apparent thermal diffusivity of heavy clay soil in field and laboratory column experiments“. In: *IOP Conference Series: Earth and Environmental Science* 368. DOI: 10.1088/1755-1315/368/1/012033.

- Minder, J. R., P. W. Mote & J. D. Lundquist (2010). „Surface temperature lapse rates over complex terrain: Lessons from the Cascade Mountains“. In: *Journal of Geophysical Research* 115.D14. DOI: 10.1029/2009jd013493.
- Molin, P. & H. Abdi (2007). „Lillifors Test For Normality“. In: *Encyclopedia of measurement and statistics*. Hrsg. von N. J. Salkind & K. Rasmussen. Thousand Oaks, Calif.: SAGE Publications, S. 540–544.
- Montávez, J. P., J. M. Lopez-Romero, S. Jerez, J. J. Gómez-Navarro & P. Jimenez-Guerrero (2017). „How much spin-up period is really necessary in regional climate simulations?“ In: *EGU General Assembly Conference Abstracts*. EGU General Assembly Conference Abstracts, S. 15806.
- Monteith, J. L. & M. H. Unsworth (2013). *Principles of environmental physics: Plants, animals, and the atmosphere*. 4. ed. Amsterdam: Elsevier/Acad. Press.
- Montzka, C., M. Herbst, L. Weihermüller, A. Verhoef & H. Vereecken (2017). „A global data set of soil hydraulic properties and sub-grid variability of soil water retention and hydraulic conductivity curves“. In: *Earth System Science Data* 9.2, S. 529–543. DOI: 10.5194/essd-9-529-2017.
- Morais, J. D., T. S. Faria, M. A. T. Elmiro, M. A. Nero, A. d. A. Silva & R. A. d. A. Nobrega (2017). „ALTIMETRY ASSESSMENT OF ASTER GDEM v2 AND SRTM v3 DIGITAL ELEVATION MODELS: A CASE STUDY IN URBAN AREA OF BELO HORIZONTE, MG, BRAZIL“. In: *Boletim de Ciências Geodésicas* 23.4, S. 654–668. DOI: 10.1590/s1982-21702017000400043.
- Morcrette, J.-J., L. Smith & Y. Fouquart (1986). „Pressure and temperature dependence of the absorption in longwave radiation parameterizations“. In: *BePhA* 59, S. 455–469.
- Moriassi, D. N., J. G. Arnold, M. W. van Liew, R. L. Bingner, R. D. Harmel & T. L. Veith (2007). „Model evaluation guidelines for systematic quantification of accuracy in watershed simulations“. In: *Transactions of the ASABE* 50.3, S. 885–900.
- Müller, G., B. Mühr & C. Kottmeier (2003). *Der Rekordsommer 2003: Der Sommer 2003 - Zahlen und Fakten*. URL: <https://www.imk-tro.kit.edu/5293.php>.
- Müller, M. D. & D. Scherer (2005). „A Grid- and Subgrid-Scale Radiation Parameterization of Topographic Effects for Mesoscale Weather Forecast Models“. In: *Monthly Weather Review* 133.6, S. 1431–1442. DOI: 10.1175/MWR2927.1.
- Munz, C.-D. & T. Westermann (2019). *Numerische Behandlung gewöhnlicher und partieller Differenzialgleichungen*. Berlin, Heidelberg: Springer Berlin Heidelberg. DOI: 10.1007/978-3-662-55886-7.
- Nachtergaele, F., H. van Velthuisen, L. Verekst & D. Widberg (2012). „Harmonized World Soil Database v 1.2“. In: *Food and Agriculture Organization of the United Nations, Rome*.
- Nachtergaele, F., H. van Velthuisen, L. Verelst & D. Wiberg (2009). „Harmonized World Soil Database“. In: *Food and Agriculture Organization of the United Nations, Rome*.
- NASA JPL (2020). *NASADEM: Merged DEM Global 1 arc second V001*. DOI: 10.5067/MEaSURES/NASADEM/NASADEM\_HGT.001.
- Nash, J. E. & J. V. Sutcliffe (1970). „River flow forecasting through conceptual models part I — A discussion of principles“. In: *Journal of Hydrology* 10.3, S. 282–290. DOI: 10.1016/0022-1694(70)90255-6.
- National Research Council (2012). *A National Strategy for Advancing Climate Modeling*. Washington, D.C.: National Academies Press. DOI: 10.17226/13430.
- Navarro-Racines, C., J. Tarapues, P. Thornton, A. Jarvis & J. Ramirez-Villegas (2020). „High-resolution and bias-corrected CMIP5 projections for climate change impact assessments“. In: *Scientific data* 7.1, S. 7. DOI: 10.1038/s41597-019-0343-8.
- Nelli, N. R., M. Temimi, R. M. Fonseca, M. J. Weston, M. S. Thota, V. K. Valappil, O. Branch, V. Wulfmeyer, Y. Wehbe, T. Al Hosary, A. Shalaby, N. Al Shamsi & H. Al Naqbi (2020). „Impact of Roughness Length on WRF Simulated Land–Atmosphere Interactions Over a Hyper–Arid Region“. In: *Earth and Space Science* 7.6. DOI: 10.1029/2020EA001165.

- Nemes, A. (2011). „Databases of Soil Physical and Hydraulic Properties“. In: *Encyclopedia of Agrophysics*. Hrsg. von J. Gliński, J. Horabik & J. Lipiec. Encyclopedia of Earth Sciences Series. Dordrecht: Springer Netherlands, S. 194–199. DOI: 10.1007/978-90-481-3585-1\_39.
- El-Nesr, M. N. (2006). „Subsurface drip irrigation system development and modeling of wetting pattern distribution“. Dissertation. Alexandria University.
- Nikolakopoulos, K. G., E. K. Kamaratakis & N. Chrysoulakis (2007). „SRTM vs ASTER elevation products. Comparison for two regions in Crete, Greece“. In: *International Journal of Remote Sensing* 27.21, S. 4819–4838. DOI: 10.1080/01431160600835853.
- Niu, G.-Y. & X. Zeng (2012). „Earth System Model, Modeling the Land Component of“. In: *Climate Change Modeling Methodology*. Hrsg. von P. J. Rasch. New York, NY: Springer New York, S. 139–168. DOI: 10.1007/978-1-4614-5767-1\_7.
- NOAA National Geophysical Data Center (2006). *2-minute Gridded Global Relief Data (ETOPO2) v2*. DOI: 10.7289/V5J1012Q.
- Nordeng, T.-E. (1994). *Extended versions of the convective parametrization scheme at ECMWF and their impact on the mean and transient activity of the model in the tropics*. Hrsg. von European Centre for Medium Range Weather Forecasts. DOI: 10.21957/e34xwhysw.
- Nossent, J. & W. Bauwens (2012). „Application of a normalized Nash-Sutcliffe efficiency to improve the accuracy of the Sobol’ sensitivity analysis of a hydrological model“. In: *Geophysical Research Abstracts*. Hrsg. von EGU.
- Ochsner, T. E., R. Horton & T. Ren (2001). „A New Perspective on Soil Thermal Properties“. In: *Soil Science Society of America Journal* 65.6, S. 1641–1647. DOI: 10.2136/sssaj2001.1641.
- Oishi, C. M., J. Y. Yuan, J. A. Cuminato & D. E. Stewart (2015). „Stability analysis of Crank–Nicolson and Euler schemes for time-dependent diffusion equations“. In: *BIT Numerical Mathematics* 55.2, S. 487–513. DOI: 10.1007/s10543-014-0509-x.
- Oke, T. R. (1987). *Boundary layer climates*. 2. ed. Bd. 632. University paperbacks. London: Methuen.
- Oleson, K. W. & D. M. Lawrence (2013). „Technical Description of version 4.5 of the Community Land Model (CLM)“. In: DOI: 10.5065/D6RR1W7M.
- Oleson, K. W., D. M. Lawrence, G. B. Bonan, M. G. Flanner, E. Kluzek, J. Peter, S. Levis, S. C. Swenson, E. Thornton & J. J. Feddema (2010). „Technical description of version 4.0 of the Community Land Model (CLM)“. In: DOI: 10.5065/D6FB50WZ.
- Oliphant, A. J., R. A. Spronken-Smith, A. P. Sturman & I. F. Owens (2003). „Spatial Variability of Surface Radiation Fluxes in Mountainous Terrain“. In: *Journal of Applied Meteorology* 42.1, S. 113–128. DOI: 10.1175/1520-0450(2003)042<0113:SVOSRF>2.0.CO;2.
- Oliveira, L. B., Ribeiro, P. K. Jacomine, J. J. Rodrigues & F. A. Marques (2002). „Funções de pedotransferência para predição da umidade retida a potenciais específicos em solos do estado de Pernambuco“. In: *Revista Brasileira de Ciencia do Solo (Brazil)* 26.2, S. 315–323.
- Olofsson, P., G. M. Foody, M. Herold, S. V. Stehman, C. E. Woodcock & M. A. Wulder (2014). „Good practices for estimating area and assessing accuracy of land change“. In: *Remote Sensing of Environment* 148, S. 42–57. DOI: 10.1016/j.rse.2014.02.015.
- Olson, J. S. (1994). „Global ecosystem framework-definitions“. In: *USGS EROS Data Center Internal Report, Sioux Falls, SD* 37, S. 1994.
- Omuto, C., F. Nachtergaele & R. Vargas Rojas (2013). *State of the art report on global and regional soil information: Where are we? Where to go?* Global soil partnership technical report. Rome: Food and Agriculture Organization of the United Nations.
- Onogi, K., J. TSUTSUI, H. KOIDE, M. SAKAMOTO, S. Kobayashi, H. HATSUSHIKA, T. MATSUMOTO, N. YAMAZAKI, H. Kamahori, K. Takahashi, S. KADOKURA, K. WADA, K. KATO, R. OYAMA, T. OSE, N. MANNOJI & R. TAIRA (2007). „The JRA-25 Reanalysis“. In: *Journal of the Meteorological Society of Japan. Ser. II* 85.3, S. 369–432. DOI: 10.2151/jmsj.85.369.

- Onwuka, B. M. (2018). „Effects of Soil Temperature on Some Soil Properties and Plant Growth“. In: *Advances in Plants & Agriculture Research* 8.1. DOI: 10.15406/apar.2018.08.00288.
- Organick, E. I. (1966). *A Fortran IV Primer*. Addison-Wesley series in computer science and information processing. Addison-Wesley.
- Ostovari, Y., K. Asgari & W. Cornelis (2015). „Performance Evaluation of Pedotransfer Functions to Predict Field Capacity and Permanent Wilting Point Using UNSODA and HYPRES Datasets“. In: *Arid Land Research and Management* 29.4, S. 383–398. DOI: 10.1080/15324982.2015.1029649.
- Pachepsky, Y. & W. J. Rawls, Hrsg. (2004). *Development of Pedotransfer Functions in Soil Hydrology*. Developments in soil science. Elsevier.
- Padarian, J., J. Morris, B. Minasny & A. B. McBratney (2018). „Pedotransfer Functions and Soil Inference Systems“. In: *Pedometrics*. Hrsg. von A. B. McBratney, B. Minasny & U. Stockmann. Progress in Soil Science. Cham: Springer International Publishing, S. 195–220. DOI: 10.1007/978-3-319-63439-5\_7.
- Paeth, H. (2011). „Postprocessing of simulated precipitation for impact research in West Africa. Part I: model output statistics for monthly data“. In: *Climate Dynamics* 36.7-8, S. 1321–1336. DOI: 10.1007/s00382-010-0760-z.
- Paeth, H., K. Born, R. Girmes, R. Podzun & D. Jacob (2009). „Regional Climate Change in Tropical and Northern Africa due to Greenhouse Forcing and Land Use Changes“. In: *Journal of Climate* 22.1, S. 114–132. DOI: 10.1175/2008JCLI2390.1.
- Pahud, D. (2002). „Geothermal energy and heat storage“. In: *Cannobio: SUPSI DCT LEEE. Scuola Universitaria Professionale della Svizzera Italiana*.
- Papadopoulos, K. H., C. G. Helmis & G. T. Amanatidis (1992). „An Analysis of Wind Direction and Horizontal Wind Component Fluctuations over Complex Terrain“. In: *Journal of Applied Meteorology* 31.9, S. 1033–1040. DOI: 10.1175/1520-0450(1992)031<1033:AAOWDA>2.0.CO;2.
- Patterson, K. A. (1990). „Global distributions of total and total-available soil water-holding capacities“. Master Thesis. Newark: University of Delaware.
- Peters-Lidard, C. D., E. Blackburn, X. Liang & E. F. Wood (1998). „The Effect of Soil Thermal Conductivity Parameterization on Surface Energy Fluxes and Temperatures“. In: *Journal of the Atmospheric Sciences* 55.7, S. 1209–1224. DOI: 10.1175/1520-0469(1998)055<1209:TEOSTC>2.0.CO;2.
- Petersen, J. & D. Rechid (2014). „Rückwirkungen von Feldberegnung und Waldumbau auf das regionale Klima – eine Modellrechnung“. In: *Klimawandel in der Lüneburger Heide*. Hrsg. von B. Urban, J. Becker & I. Mersch. Berichte aus den KLIMZUG-NORD Modellgebieten. Hamburg: TuTech Verl.
- Pfeifer, S. (2006). „Modeling cold cloud processes with the regional climate model REMO“. Dissertation. Universität Hamburg.
- Pidgeon, J. D. (1972). „THE MEASUREMENT AND PREDICTION OF AVAILABLE WATER CAPACITY OF FERRALLITIC SOILS IN UGANDA“. In: *Journal of Soil Science* 23.4, S. 431–441. DOI: 10.1111/j.1365-2389.1972.tb01674.x.
- Pielke, R. A. (1984). *Mesoscale Meteorological Modeling*. International Geophysics. Saint Louis: Elsevier Science.
- Pielke, R. A. (2013). *Mesoscale Meteorological Modeling*. 3rd ed. International Geophysics. Burlington: Elsevier Science.
- Pietikäinen, J.-P., T. Markkanen, K. Sieck, D. Jacob, J. Korhonen, P. Räisänen, Y. Gao, J. Ahola, H. Korhonen, A. Laaksonen & J. Kaurola (2017). „The regional climate model REMO (v2015) coupled with the 1-D freshwater model FLake (v1): Fenno-Scandinavian climate and lakes“. In: *Geoscientific Model Development Discussions*, S. 1–28. DOI: 10.5194/gmd-2017-245.
- Pietikäinen, J.-P., D. O’Donnell, C. Teichmann, U. Karstens, S. Pfeifer, J. Kazil, R. Podzun, S. Fiedler, H. Kokkola, W. Birmili, C. O’Dowd, U. Baltensperger, E. Weingartner, R. Gehrig, G. Spindler, M. Kulmala, J. Feichter, D. Jacob & A. Laaksonen (2012). „The regional aerosol-climate model REMO-HAM“. In: *Geoscientific Model Development Discussions* 5.1, S. 737–779. DOI: 10.5194/gmdd-5-737-2012.



- Pineda, M. C., J. Vilorio, J. A. Martínez-Casasnovas, A. Valera, D. Lobo, L. C. Timm, L. F. Pires & D. Gabriels (2018). „Predicting soil water content at - 33 kPa by pedotransfer functions in stoniness 1 soils in northeast Venezuela“. In: *Environmental monitoring and assessment* 190.3, S. 161. DOI: 10.1007/s10661-018-6528-3.
- Pinty, B., F. Roveda, M. M. Verstraete, N. Gobron, Y. Govaerts, J. V. Martonchik, D. J. Diner & R. A. Kahn (2000). „Surface albedo retrieval from Meteosat: 2. Applications“. In: *Journal of Geophysical Research* 105.D14, S. 18113–18134. DOI: 10.1029/2000JD900114.
- Pitman, A. J. (2003). „The evolution of, and revolution in, land surface schemes designed for climate models“. In: *International Journal of Climatology* 23.5, S. 479–510. DOI: 10.1002/joc.893.
- Pollinger, F., K. Ziegler, D. Abel & H. Paeth (2020). „Effects of a new land surface parametrization scheme on thermal extremes in a Regional Climate Model“. In: *General Assembly 2020*. Hrsg. von EGU. DOI: 10.5194/egusphere-egu2020-16320.
- Pontius, R. G. & M. Millones (2011). „Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment“. In: *International Journal of Remote Sensing* 32.15, S. 4407–4429. DOI: 10.1080/01431161.2011.552923.
- Porter, C. et al. (2018). *ArcticDEM: Release 7*. DOI: 10.7910/DVN/OHHUKH.
- Prein, A. F., W. Langhans, G. Fosser, A. Ferrone, N. Ban, K. Goergen, M. Keller, M. Tölle, O. Gutjahr, F. Feser, E. Brisson, S. Kollet, J. Schmidli, N. P. M. van Lipzig & R. Leung (2015). „A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges“. In: *Reviews of geophysics (Washington, D.C. : 1985)* 53.2, S. 323–361. DOI: 10.1002/2014RG000475.
- Preuschmann, S. (2012). „Regional surface albedo characteristics - analysis of albedo data and application to land-cover changes for a regional climate model“. Dissertation. Max-Planck-Institut für Meteorologie.
- Pryor, S. C., G. Nikulin & C. Jones (2012). „Influence of spatial resolution on regional climate model derived wind climates“. In: *Journal of Geophysical Research* 117.D3. DOI: 10.1029/2011JD016822.
- Qiao, J., Y. Zhu, X. Jia, L. Huang & M. Shao (2019). „Pedotransfer functions for estimating the field capacity and permanent wilting point in the critical zone of the Loess Plateau, China“. In: *Journal of Soils and Sediments* 19.1, S. 140–147. DOI: 10.1007/s11368-018-2036-x.
- Qin, Z., P. Berliner & A. Karnieli (2002). „Numerical solution of a complete surface energy balance model for simulation of heat fluxes and surface temperature under bare soil environment“. In: *Applied Mathematics and Computation* 130, S. 171–200.
- Quante, M. & F. Colijn (2016). *North Sea Region Climate Change Assessment*. Cham: Springer International Publishing. DOI: 10.1007/978-3-319-39745-0.
- Rahimi, H., Y. Khoshkhou, A. Khalili & P. Irannejad (2013). „Application of numerical method in the estimation of soil thermal diffusivity and soil temperature prediction under different textures and moisture contents“. In: *African Journal of Agricultural Research* 8.46, S. 5764–5770. DOI: 10.5897/AJAR09.717.
- Randall, D. A., R. A. Wood, S. Bony, R. Colman, T. Fiechfet, J. Fyfe, V. Kattov, A. Pitman, J. Shukla, J. Srinivasan, R. J. Stouffer, A. Sumi & K. E. Taylor (2007). „Chapter 8: Climate Models and Their Evaluation“. In: *IPCC fourth assessment report (AR4)*. Hrsg. von S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. Tignor & H. Miller, S. 588–662.
- Rasch, B., M. Friese, W. Hofmann & E. Naumann (2010). *Quantitative Methoden Band 2: Einführung in die Statistik für Psychologen und Sozialwissenschaftler*. 3., erweiterte Auflage. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg. DOI: 10.1007/978-3-642-05270-5.
- Raschke, E., U. Karstens, R. Nolte-Holube, R. Brandt, H.-J. Isemer, D. Lohmann, M. Lobmeyr, B. Rockel & R. Stuhlmann (1998). „The Baltic Sea Experiment BALTEX: A brief overview and some selected results of the authors“. In: *Surveys in Geophysics* 19.1, S. 1–22. DOI: 10.1023/A:1006567924820.
- Rawls, W. J., Y. A. Pachepsky, J. C. Ritchie, T. M. Sobecki & H. Bloodworth (2003). „Effect of soil organic carbon on soil water retention“. In: *Geoderma* 116.1-2, S. 61–76. DOI: 10.1016/S0016-7061(03)00094-6.

- Rawls, W. J. & D. L. Brakensiek (1982). „Estimating soil water retention from soil properties“. In: *Journal of the Irrigation and Drainage Division* 108.2, S. 166–171.
- Rechid, D., D. Jacob, R. Podzun, S. Kotlarski, T. Semmler, K. Bülow, P. Lorenz & S. Hagemann (2003). „Water budgets for European catchment areas simulated by the regional climate model REMO“. In: *EGS-AGU-EUG Joint Assembly*. Hrsg. von European Geophysical Society. Nice.
- Rechid, D. (2001). „Untersuchung zur Parameterisierung von Landoberflächen im regionalen Klimamodell REMO“. Diplomarbeit. Universität Hannover.
- Rechid, D. (2009). „On biogeophysical interactions between vegetation phenology and climate simulated over Europe“. Dissertation. Max-Planck-Institut für Meteorologie.
- Rechid, D. & D. Jacob (2006). „Influence of monthly varying vegetation on the simulated climate in Europe“. In: *Meteorologische Zeitschrift* 15.1, S. 99–116. DOI: 10.1127/0941-2948/2006/0091.
- Rechid, D., T. J. Raddatz & D. Jacob (2009). „Parameterization of snow-free land surface albedo as a function of vegetation phenology based on MODIS data and applied in climate modelling“. In: *Theoretical and Applied Climatology* 95.3-4, S. 245–255. DOI: 10.1007/s00704-008-0003-y.
- Reichert, J. M., J. A. Albuquerque, D. R. Kaiser, D. J. Reinert, F. L. Urach & R. Carlesso (2009). „Estimation of water retention and availability in soils of Rio Grande do Sul“. In: *Revista Brasileira de Ciencia do Solo (Brazil)* 33.6, S. 1547–1560. DOI: 10.1590/S0100-06832009000600004.
- Reijmer, C. H., E. van Meijgaard & M. R. van den Broeke (2004). „Numerical Studies with a Regional Atmospheric Climate Model Based on Changes in the Roughness Length for Momentum and Heat Over Antarctica“. In: *Boundary-Layer Meteorology* 111.2, S. 313–337. DOI: 10.1023/B:BOUN.0000016470.23403.ca.
- Remedio, A. R. (2013). „Connections of low level jets and mesoscale convective systems in South America“. In: *Berichte zur Erdsystemforschung* 141.
- Remke, T. & D. Jacob (2017). „Towards the assessment of climate change impacts on critical energy infrastructure applied for wind energy“. In: *Challenges in Weather and Climate Services for Energy*. Hrsg. von World Energy & Meteorology Council. Villa Romanazzi Carduei Bari, Italy.
- Ren, J., L. Men, W. Zhang & J. Yang (2019). „A new empirical model for the estimation of soil thermal conductivity“. In: *Environmental Earth Sciences* 78.12, S. 1285. DOI: 10.1007/s12665-019-8360-7.
- Rexer, M. & C. Hirt (2014). „Comparison of free high resolution digital elevation data sets (ASTER GDEM2, SRTM v2.1/v4.1) and validation against accurate heights from the Australian National Gravity Database“. In: *Australian Journal of Earth Sciences* 61.2, S. 213–226. DOI: 10.1080/08120099.2014.884983.
- Richtmyer, R. D. & K. W. Morton (1994). *Difference methods for initial-value problems*. Second edition. Malabar, Florida: Krieger publishing company.
- Ridder, K. de (2009). „Soil temperature spin-up in land surface schemes“. In: *Theoretical and Applied Climatology* 95.3-4, S. 341–347. DOI: 10.1007/s00704-008-0011-y.
- Riegler, G., S. D. Hennig & M. Weber (2015). „WORLDDEM – A NOVEL GLOBAL FOUNDATION LAYER“. In: *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XL-3/W2, S. 183–187. DOI: 10.5194/isprsarchives-XL-3-W2-183-2015.
- Rienecker, M. M. et al. (2011). „MERRA: NASA’s Modern-Era Retrospective Analysis for Research and Applications“. In: *Journal of Climate* 24.14, S. 3624–3648. DOI: 10.1175/JCLI-D-11-00015.1.
- Ring, C., F. Pollinger, I. Kaspar-Ott, E. Hertig, J. Jacobeit & H. Paeth (2018). „A comparison of metrics for assessing state-of-the-art climate models and implications for probabilistic projections of climate change“. In: *Climate Dynamics* 50.5-6, S. 2087–2106. DOI: 10.1007/s00382-017-3737-3.
- Rizzoli, P., M. Martone, C. Gonzalez, C. Wecklich, D. Borla Tridon, B. Bräutigam, M. Bachmann, D. Schulze, T. Fritz, M. Huber, B. Wessel, G. Krieger, M. Zink & A. Moreira (2017). „Generation and performance assessment of the global TanDEM-X digital elevation model“. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 132, S. 119–139. DOI: 10.1016/j.isprsjprs.2017.08.008.

- Robine, J.-M., S. L. K. Cheung, S. Le Roy, H. van Oyen, C. Griffiths, J.-P. Michel & F. R. Herrmann (2008). „Death toll exceeded 70,000 in Europe during the summer of 2003“. In: *Comptes rendus biologiques* 331.2, S. 171–178. DOI: 10.1016/j.crv.2007.12.001.
- Rodriguez, E., C. S. Morris & J. E. Belz (2006). „A global assessment of the SRTM performance“. In: *PHOTOGRAMMETRIC ENGINEERING & REMOTE SENSING* 72.3, S. 249–260.
- Roeckner, E., K. Arpe, L. Bengtsson, M. Christoph, M. Claussen, L. Dümenil, M. Esch, M. Giorgetta, U. Schlese & U. Schulzweida (1996). *The Atmospheric general circulation Model ECHAM-4: Model description and simulation of present-day climate*.
- Rontu, L. (2006). „A study on parametrization of orography-related momentum fluxes in a synoptic-scale NWP model“. In: *Tellus A: Dynamic Meteorology and Oceanography* 58.1, S. 69–81. DOI: 10.1111/j.1600-0870.2006.00162.x.
- Rummukainen, M. (2010). „State of the art with regional climate models“. In: *Wiley Interdisciplinary Reviews: Climate Change* 1, S. 82–96. DOI: 10.1002/wcc.008.
- Rummukainen, M. (2016). „Added value in regional climate modeling“. In: *Wiley Interdisciplinary Reviews: Climate Change* 7.1, S. 145–159. DOI: 10.1002/wcc.378.
- Rustanto, A., M. J. Booij, H. Wösten & A. Y. Hoekstra (2017). „Application and recalibration of soil water retention pedotransfer functions in a tropical upstream catchment: case study in Bengawan Solo, Indonesia“. In: *Journal of Hydrology and Hydromechanics* 65.3, S. 307–320. DOI: 10.1515/johh-2017-0020.
- Saha, S. K. (2006). „The influence of an improved soil scheme on the arctic climate in a regional climate model (RCM)“. In: *Berichte zur Polar-und Meeresforschung (Reports on Polar and Marine Research)* 519.
- Saha, S. et al. (2010). „The NCEP Climate Forecast System Reanalysis“. In: *Bulletin of the American Meteorological Society* 91.8, S. 1015–1058. DOI: 10.1175/2010BAMS3001.1.
- Sahid, M., A. Arifati, A. W. Nurrohman, H. M. Ihsan & M. Z. Arifin (2018). „VERTICAL ACCURACY ASSESSMENT FOR SRTM V.4 AND ASTER GLOBAL DIGITAL ELEVATION MODELS V.2: A CASE STUDY OF PADANG REGENCY, WEST SUMATERA“. In: *Seminar Nasional Geomatika* 2, S. 399–408. DOI: 10.24895/SNG.2017.2-0.435.
- Santanello, J. A., P. A. Dirmeyer, C. R. Ferguson, K. L. Findell, A. B. Tawfik, A. Berg, M. B. Ek, P. Gentine, B. P. Guillod, C. van Heerwaarden, J. Roundy & V. Wulfmeyer (2018). „Land–Atmosphere Interactions: The LoCo Perspective“. In: *Bulletin of the American Meteorological Society* 99.6, S. 1253–1272. DOI: 10.1175/BAMS-D-17-0001.1.
- Santillan, J. R. & M. Makinano-Santillan (2016). „VERTICAL ACCURACY ASSESSMENT OF 30-M RESOLUTION ALOS, ASTER, AND SRTM GLOBAL DEMS OVER NORTHEASTERN MINDANAO, PHILIPPINES“. In: *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLI-B4, S. 149–156. DOI: 10.5194/isprs-archives-XLI-B4-149-2016.
- Santra, P., M. Kumar, R. N. Kumawat, D. K. Painuli, K. M. Hati, G. B. M. Heuvelink & N. H. Batjes (2018). „Pedotransfer functions to estimate soil water content at field capacity and permanent wilting point in hot Arid Western India“. In: *Journal of Earth System Science* 127.3, S. 1–16.
- Satgé, F., M. P. Bonnet, F. Timouk, S. Calmant, R. Pillco, J. Molina, W. Lavado-Casimiro, A. Arsen, J. F. Crétaux & J. Garnier (2015). „Accuracy assessment of SRTM v4 and ASTER GDEM v2 over the Altiplano watershed using ICESat/GLAS data“. In: *International Journal of Remote Sensing* 36.2, S. 465–488. DOI: 10.1080/01431161.2014.999166.
- Sato, H., A. Ito, A. Ito, T. Ise & E. Kato (2015). „Current status and future of land surface models“. In: *Soil Science and Plant Nutrition* 61.1, S. 34–47. DOI: 10.1080/00380768.2014.917593.
- Sauer, T. J. & R. Horton (2005). „Soil Heat Flux“. In: *Micrometeorology in Agricultural Systems*. Hrsg. von J. L. Hatfield & J. M. Baker. Agronomy Monographs. Madison, WI, USA: American Society of Agronomy, Crop Science Society of America, and Soil Science Society of America, S. 131–154. DOI: 10.2134/agronmonogr47.c7.

- Saxton, K. E. & W. J. Rawls (2006). „Soil Water Characteristic Estimates by Texture and Organic Matter for Hydrologic Solutions“. In: *Soil Science Society of America Journal* 70.5, S. 1569. DOI: 10.2136/sssaj2005.0117.
- Saxton, K. E., W. J. Rawls, J. S. Romberger & R. I. Papendick (1986). „Estimating Generalized Soil-water Characteristics from Texture“. In: *Soil Science Society of America Journal* 50.4, S. 1031–1036. DOI: 10.2136/sssaj1986.03615995005000040039x.
- Schaaf, C. B. et al. (2002). „First operational BRDF, albedo nadir reflectance products from MODIS“. In: *Remote Sensing of Environment* 83.1-2, S. 135–148. DOI: 10.1016/S0034-4257(02)00091-3.
- Schaap, M. G., F. J. Leij & M. T. van Genuchten (2001). „rosetta : a computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions“. In: *Journal of Hydrology* 251.3-4, S. 163–176. DOI: 10.1016/S0022-1694(01)00466-8.
- Schär, C. & G. Jendritzky (2004). „Climate change: hot news from summer 2003“. In: *Nature* 432.7017, S. 559–560. DOI: 10.1038/432559a.
- Schubert, E. & M. Gertz (2018). „Numerically stable parallel computation of (co-)variance“. In: *Proceedings of the 30th International Conference on Scientific and Statistical Database Management*. Hrsg. von D. Sacharidis, J. Gamper & M. Böhlen. New York, NY, USA: ACM, S. 1–12. DOI: 10.1145/3221269.3223036.
- Schulla, J. (2021). *Modell description WaSiM-ETH. (Water balance Simulation Model): Version 10.06.00*. URL: [http://wasim.ch/downloads/doku/wasim/wasim\\_2021\\_en.pdf](http://wasim.ch/downloads/doku/wasim/wasim_2021_en.pdf).
- Schulzweida, U. (2020). *CDO User Guide: CDO Climate Data Operators, Version 1.9.9*. URL: <https://code.mpimet.mpg.de/projects/cdo/embedded/cdo.pdf>.
- Scott, D. W. (1992). *Multivariate density estimation: Theory, practice, and visualization*. 2. [print.] Wiley series in probability and mathematical statistics applied probability and statistics. New York, NY: Wiley.
- Sellers, Dickinson, Randall, A. K. Betts, Hall, Berry, Collatz, Denning, Mooney, Nobre, Sato, Field & Henderson-Sellers (1997). „Modeling the Exchanges of Energy, Water, and Carbon Between Continents and the Atmosphere“. In: *Science (New York, N. Y.)* 275.5299, S. 502–509. DOI: 10.1126/science.275.5299.502.
- Semmler, T. (2002). „Der Wasser- und Energiehaushalt der arktischen Atmosphäre“. Dissertation.
- Semmler, T. & D. Jacob (2004). „Modeling extreme precipitation events—a climate change simulation for Europe“. In: *Global and Planetary Change* 44.1-4, S. 119–127. DOI: 10.1016/j.gloplacha.2004.06.008.
- Semmler, T., D. Jacob, K. H. Schlünzen & R. Podzun (2004). „Influence of Sea Ice Treatment in a Regional Climate Model on Boundary Layer Values in the Fram Strait Region“. In: *Monthly Weather Review* 132.4, S. 985–999. DOI: 10.1175/1520-0493(2004)132<0985:IOSITI>2.0.CO;2.
- Sertel, E. (2010). *ACCURACY ASSESSMENT OF ASTER GLOBAL DEM OVER TURKEY*.
- Shangguan, W., Y. Dai, Q. Duan, B. Liu & H. Yuan (2014). „A global soil data set for earth system modeling“. In: *Journal of Advances in Modeling Earth Systems* 6.1, S. 249–263. DOI: 10.1002/2013MS000293.
- Shao, Y. & P. Irannejad (1999). „On the Choice of Soil Hydraulic Models in Land-Surface Schemes“. In: *Boundary-Layer Meteorology* 90.1, S. 83–115. DOI: 10.1023/A:1001786023282.
- Sheather, S. J. (2004). „Density Estimation“. In: *Statistical Science* 19.4, S. 588–597. DOI: 10.1214/088342304000000297.
- Shepard, D. (1968). „A two-dimensional interpolation function for irregularly-spaced data“. In: *Proceedings of the 1968 23rd ACM national conference on -*. Hrsg. von R. B. Blue & A. M. Rosenberg. New York, New York, USA: ACM Press, S. 517–524. DOI: 10.1145/800186.810616.
- Shi, X. Z., D. S. Yu, E. D. Warner, X. Z. Pan, G. W. Petersen, Z. G. Gong & D. C. Weindorf (2004). „Soil database of 1: 1,000,000 digital soil survey and reference system of the Chinese genetic soil classification system“. In: *Soil Survey Horizons* 45.4, S. 129–136.
- Sieck, K. (2013). „Internal Variability in the Regional Climate Model REMO“. In: *Berichte zur Erdsystemforschung* 142.

- Sillmann, J., V. V. Kharin, X. Zhang, F. W. Zwiers & D. Bronaugh (2013). „Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. Model evaluation in the present climate“. In: *Journal of Geophysical Research: Atmospheres* 118.4, S. 1716–1733. DOI: 10.1002/jgrd.50203.
- Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. Boston, MA: Springer US. DOI: 10.1007/978-1-4899-3324-9.
- Simmons, A. J. & D. M. Burridge (1981). „An energy and angular-momentum conserving vertical finite-difference scheme and hybrid vertical coordinates“. In: *Monthly Weather Review* 109.4, S. 758–766. DOI: 10.1175/1520-0493(1981)109<0758:AEAAMC>2.0.CO;2.
- Simmons, A. J., S. Uppala, D. P. Dee & S. Kobayashi (2007). *ERA-Interim: New ECMWF reanalysis products from 1989 onwards*. DOI: 10.21957/pocnex23c6.
- Simunek, J., M. T. van Genuchten & M. Sejna (2005). „The HYDRUS-1D software package for simulating the one-dimensional movement of water, heat, and multiple solutes in variably-saturated media“. In: *University of California-Riverside Research Reports* 3, S. 1–240.
- Skamarock, W. C. & J. B. Klemp (2008). „A time-split nonhydrostatic atmospheric model for weather research and forecasting applications“. In: *Journal of Computational Physics* 227.7, S. 3465–3485. DOI: 10.1016/j.jcp.2007.01.037.
- Sloan, S. W. (1985). „A point-in-polygon program“. In: *Advances in Engineering Software (1978)* 7.1, S. 45–47. DOI: 10.1016/0141-1195(85)90094-4.
- Smits, P. C., S. G. Dellepiane & R. A. Schowengerdt (1999). „Quality assessment of image classification algorithms for land-cover mapping: A review and a proposal for a cost-based approach“. In: *International Journal of Remote Sensing* 20.8, S. 1461–1486. DOI: 10.1080/014311699212560.
- Stahl, K., R. D. Moore, J. A. Floyer, M. G. Asplin & I. G. McKendry (2006). „Comparison of approaches for spatial interpolation of daily air temperature in a large region with complex topography and highly variable station density“. In: *Agricultural and Forest Meteorology* 139.3-4, S. 224–236. DOI: 10.1016/j.agrformet.2006.07.004.
- Stahr, K., E. Kandeler, L. Herrmann & T. Streck (2020). *Bodenkunde und Standortlehre*. 4., vollständig überarbeitete Auflage. Agrarwissenschaften, Geowissenschaften, Umweltwissenschaften. Stuttgart: Verlag Eugen Ulmer.
- Strahler, A. H., W. Lucht, C. B. Schaaf, T. Tsang, F. Gao, X. Li, J.-P. Muller, P. Lewis & M. J. Barnsley (1999). *MODIS BRDF/Albedo Product: Algorithm Theoretical Basis Document Version 5.0*. NASA Goddard Space Flight Cent. URL: [https://modis.gsfc.nasa.gov/data/atbd/atbd\\_mod09.pdf](https://modis.gsfc.nasa.gov/data/atbd/atbd_mod09.pdf).
- Sturm, C., G. Hoffmann & B. Langmann (2007). „Simulation of the Stable Water Isotopes in Precipitation over South America: Comparing Regional to Global Circulation Models“. In: *Journal of Climate* 20.15, S. 3730–3750. DOI: 10.1175/JCLI4194.1.
- Sun, Q., C. Miao, Q. Duan, H. Ashouri, S. Sorooshian & K.-L. Hsu (2018). „A Review of Global Precipitation Data Sets: Data Sources, Estimation, and Intercomparisons“. In: *Reviews of Geophysics* 56.1, S. 79–107. DOI: 10.1002/2017RG000574.
- Sun, W.-Y. & J.-D. Chern (2005). „Validation of A One-Dimensional Snow-Land Surface Model at the Sleepers River Watershed“. In: *Boundary-Layer Meteorology* 116.1, S. 95–115. DOI: 10.1007/s10546-004-7741-x.
- Sundqvist, H. (1978). „A parameterization scheme for non-convective condensation including prediction of cloud water content“. In: *Quarterly Journal of the Royal Meteorological Society* 104, S. 677–690.
- Sung, C. & J. Iba (2010). „Accuracy of the Saxton-Rawls method for estimating the soil water characteristics for mineral soils of Malaysia“. In: *Pertanika J. Trop. Agric. Sci.* 33.2, S. 297–302.
- Suwandana, E., K. Kawamura, Y. Sakuno, E. Kustiyanto & B. Raharjo (2012). „Evaluation of ASTER GDEM2 in Comparison with GDEM1, SRTM DEM and Topographic-Map-Derived DEM Using Inundation Area Analysis and RTK-dGPS Data“. In: *Remote Sensing* 4.8, S. 2419–2431. DOI: 10.3390/rs4082419.

- Szypuła, B. (2019). „Quality assessment of DEM derived from topographic maps for geomorphometric purposes“. In: *Open Geosciences* 11.1, S. 843–865. DOI: 10.1515/geo-2019-0066.
- Tachikawa, T., M. Hato, M. Kaku & A. Iwasaki (2011). „Characteristics of ASTER GDEM version 2“. In: *IEEE International Geoscience and Remote Sensing Symposium*. Hrsg. von IEEE, S. 3657–3660.
- Tachikawa, T., M. Kaku, A. Iwasaki, D. B. Gesch, M. J. Oimoen, Z. Zhang, J. J. Danielson, T. Krieger, B. Curtis, J. Haase, M. Abrams & C. C. Carabajal (2011). *ASTER Global Digital Elevation Model Version 2 - summary of validation results*. URL: <http://pubs.er.usgs.gov/publication/70005960>.
- Tadono, T., H. Ishida, F. Oda, S. Naito, K. Minakawa & H. Iwamoto (2014). „Precise Global DEM Generation by ALOS PRISM“. In: *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences* II-4, S. 71–76. DOI: 10.5194/isprsannals-II-4-71-2014.
- Tafasca, S., A. Ducharme & C. Valentin (2020). „Weak sensitivity of the terrestrial water budget to global soil texture maps in the ORCHIDEE land surface model“. In: *Hydrology and Earth System Sciences* 24.7, S. 3753–3774. DOI: 10.5194/hess-24-3753-2020.
- Takaku, J., T. Tadono & K. Tsutsui (2014). „Generation of High Resolution Global DSM from ALOS PRISM“. In: *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XL-4, S. 243–248. DOI: 10.5194/isprsarchives-XL-4-243-2014.
- Takaku, J., T. Tadono, K. Tsutsui & M. Ichikawa (2016). „VALIDATION OF 'AW3D' GLOBAL DSM GENERATED FROM ALOS PRISM“. In: *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences* III-4, S. 25–31. DOI: 10.5194/isprsannals-III-4-25-2016.
- Tang, W., J. Qin, K. Yang, F. Zhu & X. Zhou (2021). „Does ERA5 outperform satellite products in estimating atmospheric downward longwave radiation at the surface?“ In: *Atmospheric Research* 252, S. 105453. DOI: 10.1016/j.atmosres.2021.105453.
- Tarek, M., F. P. Brissette & R. Arsenault (2019). *Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modeling over North-America*. DOI: 10.5194/hess-2019-316.
- Tarnawski, V. R. & F. Gori (2002). „Enhancement of the cubic cell soil thermal conductivity model“. In: *INTERNATIONAL JOURNAL OF ENERGY RESEARCH* 26.2, S. 143–157. DOI: 10.1002/er.772.
- Teichmann, C., B. Eggert, A. Elizalde, A. Haensler, D. Jacob, P. Kumar, C. Moseley, S. Pfeifer, D. Rechid, A. Remedio, H. Ries, J. Petersen, S. Preuschmann, T. Raub, F. Saeed, K. Sieck & T. Weber (2013). „How Does a Regional Climate Model Modify the Projected Climate Change Signal of the Driving GCM: A Study over Different CORDEX Regions Using REMO“. In: *Atmosphere* 4.2, S. 214–236. DOI: 10.3390/atmos4020214.
- Tetzner, D., E. Thomas & C. Allen (2019). „A Validation of ERA5 Reanalysis Data in the Southern Antarctic Peninsula—Ellsworth Land Region, and Its Implications for Ice Core Studies“. In: *Geosciences* 9.7, S. 289. DOI: 10.3390/geosciences9070289.
- Thomas, L. H. (1949). „Elliptic problems in linear difference equations over a network“. In: *Watson Sci. Comput. Lab. Rept., Columbia University, New York* 1.
- Tiedtke, M. (1989). „A Comprehensive Mass Flux Scheme for Cumulus Parameterization in Large-Scale Models“. In: *Monthly Weather Review* 117, S. 1779–1800.
- Tomasella, J. & M. Hodnett (2004). „Pedotransfer functions for tropical soils“. In: *Development of Pedotransfer Functions in Soil Hydrology*. Hrsg. von Y. Pachepsky & W. J. Rawls. Bd. 30. Developments in soil science. Elsevier, S. 415–429. DOI: 10.1016/S0166-2481(04)30021-8.
- Tomasella, J. & M. G. Hodnett (1998). „Estimating soil water retention characteristics from limited data in Brazilian Amazonia“. In: *Soil Science* 163.3, S. 190–202.
- Tran, A. P., B. Dafflon & S. S. Hubbard (2017). „Coupled land surface–subsurface hydrogeophysical inverse modeling to estimate soil organic carbon content and explore associated hydrological and thermal dynamics in the Arctic tundra“. In: *The Cryosphere* 11.5, S. 2089–2109. DOI: 10.5194/tc-11-2089-2017.
- Trinks, S. (2010). „Einfluss des Wasser- und Wärmehaushaltes von Böden auf den Betrieb erdverlegter Energiekabel“. Dissertation. Berlin: Technische Universität Berlin. DOI: 10.14279/depositonce-2739.

- Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA.
- Undén, P., L. Rontu, H. Järvinen, P. Lynch, J. Calvo-Sanchez, G. Cats, J. Cuxart, K. Eerola, C. Fortelius & J. García-Moya (2002). *HIRLAM-5 scientific documentation*.
- University Corporation for Atmospheric Research (2011). *Growth of Climate Modeling (DI02399)*. URL: <http://n2t.net/ark:/85065/d7r20zjd>.
- University Of East Anglia Climatic Research Unit, I. C. Harris & P. D. Jones (2019). *CRU TS4.03: Climatic Research Unit (CRU) Time-Series (TS) version 4.03 of high-resolution gridded data of month-by-month variation in climate (Jan. 1901- Dec. 2018)*. DOI: 10.5285/10d3e3640f004c578403419aac167d82.
- Uppala, S. M. et al. (2005). „The ERA-40 re-analysis“. In: *Quarterly Journal of the Royal Meteorological Society* 131.612, S. 2961–3012. DOI: 10.1256/qj.04.176.
- Urraca, R., T. Huld, A. Gracia-Amillo, F. J. Martinez-de-Pison, F. Kaspar & A. Sanz-Garcia (2018). „Evaluation of global horizontal irradiance estimates from ERA5 and COSMO-REA6 reanalyses using ground and satellite-based data“. In: *Solar Energy* 164, S. 339–354. DOI: 10.1016/j.solener.2018.02.059.
- van den Berg, M., E. Klamt, L. P. van Reeuwijk & W. G. Sombroek (1997). „Pedotransfer functions for the estimation of moisture retention characteristics of Ferralsols and related soils“. In: *Geoderma* 78.3-4, S. 161–180. DOI: 10.1016/S0016-7061(97)00045-1.
- van den Besselaar, E. J. M., M. R. Haylock, G. van der Schrier & A. M. G. Klein Tank (2011). „A European daily high-resolution observational gridded data set of sea level pressure“. In: *Journal of Geophysical Research* 116.D11, S. 2717. DOI: 10.1029/2010JD015468.
- van den Hurk, B., P. Viterbo, A. Beljaars & A. K. Betts (2000). *Offline Validation of the ERA-40 Surface Scheme: ECMWF Tech. Memo. No. 295*.
- van Engelen, V. W. & J. A. Dijkshoorn (2012). „Global and National Soils and Terrain Databases (SOTER). Procedures Manual, version 2.0, draft for comments. ISRIC Report 2012/04, ISRIC–World Soil Information, Wageningen“. In: *ISRIC Report* 4.
- van Genuchten, M. T. (1980). „A Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils“. In: *Soil Science Society of America Journal* 44.5, S. 892–898.
- van Looy, K., J. Bouma, M. Herbst, J. Koestel, B. Minasny, U. Mishra, C. Montzka, A. Nemes, Y. A. Pachepsky, J. Padarian, M. G. Schaap, B. Tóth, A. Verhoef, J. Vanderborght, M. J. Ploeg, L. Weiermüller, S. Zacharias, Y. Zhang & H. Vereecken (2017). „Pedotransfer Functions in Earth System Science: Challenges and Perspectives“. In: *Reviews of geophysics (Washington, D.C. : 1985)* 55.4, S. 1199–1256. DOI: 10.1002/2017RG000581.
- van Vlaardingen, P. L. A. & C. E. Smit (2009). „Verification of moisture content in test soils“. In: *National Institute for Public Health and the Environment RIVM Letter report* 601516016.
- van Wijk, W. R., Hrsg. (1963). *Physics of Plant Environment*. New York: Wiley.
- van Wijk, W. R. & D. A. de Vries (1963). „Periodic temperature variations in a homogeneous soil“. In: *Physics of Plant Environment*. Hrsg. von W. R. van Wijk. New York: Wiley, S. 103–143.
- Veihmeyer, F. J. & A. H. Hendrickson (1928). „SOIL MOISTURE AT PERMANENT WILTING OF PLANTS“. In: *Plant physiology* 3.3, S. 355–357. DOI: 10.1104/pp.3.3.355.
- Veihmeyer, F. J. & A. H. Hendrickson (1931). „The moisture equivalent as a measure of the field capacity of soils“. In: *Soil Science* 32.3, S. 181–194.
- Vereecken, H., M. Weynants, M. Javaux, Y. Pachepsky, M. G. Schaap & M. van Genuchten (2010). „Using Pedotransfer Functions to Estimate the van Genuchten-Mualem Soil Hydraulic Properties: A Review“. In: *Vadose Zone Journal* 9.4, S. 795–820. DOI: 10.2136/vzj2010.0045.
- Vereecken, H. et al. (2016). „Modeling Soil Processes: Review, Key Challenges, and New Perspectives“. In: *Vadose Zone Journal* 15.5. DOI: 10.2136/vzj2015.09.0131.
- Verseghy, D. L. (1991). „CLASS—A Canadian land surface scheme for GCMs. I. Soil model“. In: *International Journal of Climatology* 11.2, S. 111–133.

- Viterbo, P. (2002). „A review of parametrization schemes for land surface processes“. In: *Meteorological Training Course Lecture Series*.
- Vondou, D. A. & A. Haensler (2017). „Evaluation of simulations with the regional climate model REMO over Central Africa and the effect of increased spatial resolution“. In: *International Journal of Climatology* 37, S. 741–760. DOI: 10.1002/joc.5035.
- Wagena, M. B., A. S. Collick, A. C. Ross, R. G. Najjar, B. Rau, A. R. Sommerlot, D. R. Fuka, P. J. A. Kleinman & Z. M. Easton (2018). „Impact of climate change and climate anomalies on hydrologic and biogeochemical processes in an agricultural catchment of the Chesapeake Bay watershed, USA“. In: *Science of the Total Environment* 637-638, S. 1443–1454. DOI: 10.1016/j.scitotenv.2018.05.116.
- Walter, A., K. Keuler, D. Jacob, R. Knoche, A. Block, S. Kotlarski, G. Müller-Westermeier, D. Rechid & W. Ahrens (2006). „A high resolution reference data set of German wind velocity 19512001 and comparison with regional climate model results“. In: *Meteorologische Zeitschrift* 15.6, S. 585–596. DOI: 10.1127/0941-2948/2006/0162.
- Wang, F., F. Cheruy & J.-L. Dufresne (2016). „The improvement of soil thermodynamics and its effects on land surface meteorology in the IPSL climate model“. In: *Geoscientific Model Development* 9.1, S. 363–381. DOI: 10.5194/gmd-9-363-2016.
- Wang, J. & V. R. Kotamarthi (2014). „Downscaling with a nested regional climate model in near-surface fields over the contiguous United States“. In: *Journal of Geophysical Research: Atmospheres* 119.14, S. 8778–8797. DOI: 10.1002/2014JD021696.
- Wang, J., H. He, M. Dyck & J. Lv (2020). „A Review and Evaluation of Predictive Models for Thermal Conductivity of Sands at Full Water Content Range“. In: *Energies* 13.5, S. 1083. DOI: 10.3390/en13051083.
- Wang, Y., L. Leung, J. McGregor, D.-K. Lee, W.-C. Wang, Y. Ding & F. Kimura (2004). „Regional Climate Modeling: Progress, Challenges, and Prospects“. In: *Journal of The Meteorological Society of Japan - J METEOROL SOC JPN* 82, S. 1599–1628. DOI: 10.2151/jmsj.82.1599.
- Warrilow, D. A., A. B. Sangster & A. Slingo (1986). *Modelling of land surface processes and their influence on European climate*. Met. O. 20,(Dynam. Climatology Branch).
- Watson, P. F. & A. Petrie (2010). „Method agreement analysis: a review of correct methodology“. In: *Therriogenology* 73.9, S. 1167–1179. DOI: 10.1016/j.theriogenology.2010.01.003.
- Watson, R. T. & Core Writing Team, Hrsg. (2001). *Climate change 2001: Synthesis Report: A Contribution of Working Groups I, II, and III to the Third Assessment Report of the Intergovernmental Panel on Climate Change*. IPCC. Cambridge: Cambridge Univ. Press.
- Węglarczyk, S. (2018). „Kernel density estimation and its application“. In: *ITM Web of Conferences* 23, S. 00037. DOI: 10.1051/itmconf/20182300037.
- Wessel, B., M. Huber, C. Wohlfart, U. Marschalk, D. Kosmann & A. Roth (2018). „Accuracy assessment of the global TanDEM-X Digital Elevation Model with GPS data“. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 139, S. 171–182. DOI: 10.1016/j.isprsjprs.2018.02.017.
- Wetter Online (2006). *Strenger Winter 2005/2006: Viel Frost und viel Schnee*. URL: <https://www.wetteronline.de/wetterrueckblick/strenger-winter-2005-2006-viel-frost-und-viel-schnee-2006-02-25-wi>.
- Whan, K., J. Zscheischler, R. Orth, M. Shongwe, M. Rahimi, E. O. Asare & S. I. Seneviratne (2015). „Impact of soil moisture on extreme maximum temperatures in Europe“. In: *Weather and Climate Extremes* 9, S. 57–67. DOI: 10.1016/j.wace.2015.05.001.
- Wilcoxon, F. (1945). „Individual Comparisons by Ranking Methods“. In: *Biometrics Bulletin* 1.6, S. 80. DOI: 10.2307/3001968.
- Wilhelm, C., D. Rechid & D. Jacob (2011). „Modeling land cover changes with the regional model REMO-iMOVE for Europe“. In: *General Assembly 2011*. Hrsg. von EGU.



- Wilhelm, C., D. Rechied & D. Jacob (2014). „Interactive coupling of regional atmosphere with biosphere in the new generation regional climate system model REMO-iMOVE“. In: *Geoscientific Model Development* 7.3, S. 1093–1114. DOI: 10.5194/gmd-7-1093-2014.
- Wilks, D. S. (2006). *Statistical methods in the atmospheric sciences*. 2nd ed. Bd. 91. International geophysics series. Amsterdam & Boston: Academic Press.
- Wilson, M. F. & A. Henderson-Sellers (1985). „A global archive of land cover and soils data for use in general circulation climate models“. In: *Journal of Climatology* 5.2, S. 119–143. DOI: 10.1002/joc.3370050202.
- Witt, K. J., Hrsg. (2018). *Grundbau-Taschenbuch*. 8. Auflage. Berlin: Ernst & Sohn a Wiley brand.
- Wösten, J., P. A. Finke & M. Jansen (1995). „Comparison of class and continuous pedotransfer functions to generate soil hydraulic characteristics“. In: *Geoderma* 66.3-4, S. 227–237. DOI: 10.1016/0016-7061(94)00079-P.
- Wösten, J., Y. Pachepsky & W. J. Rawls (2001). „Pedotransfer functions: bridging the gap between available basic soil data and missing soil hydraulic characteristics“. In: *Journal of Hydrology* 251.3-4, S. 123–150. DOI: 10.1016/S0022-1694(01)00464-4.
- Wu, M., G. Schurgers, A. Ahlström, M. Rummukainen, P. A. Miller, B. Smith & W. May (2017). „Impacts of land use on climate and ecosystem productivity over the Amazon and the South American continent“. In: *Environmental Research Letters* 12.5. DOI: 10.1088/1748-9326/aa6fd6.
- Yamazaki, D., D. Ikeshima, R. Tawatari, T. Yamaguchi, F. O’Loughlin, J. C. Neal, C. C. Sampson, S. Kanai & P. D. Bates (2017). „A high-accuracy map of global terrain elevations“. In: *Geophysical Research Letters* 44.11, S. 5844–5853. DOI: 10.1002/2017GL072874.
- Yang, Y., R. J. Donohue & T. R. McVicar (2016). „Global estimation of effective plant rooting depth: Implications for hydrological modeling“. In: *Water Resources Research* 52.10, S. 8260–8276. DOI: 10.1002/2016WR019392.
- Yang, Z.-L. (2004). „MODELING LAND SURFACE PROCESSES IN SHORT-TERM WEATHER AND CLIMATE STUDIES“. In: *Observation, Theory and Modeling of Atmospheric Variability*. Hrsg. von X. Zhu, X. Li, M. Cai, S. Zhou, Y. Zhu, F.-F. Jin, X. Zou & M. Zhang. Bd. 03. World Scientific Series on Asia-Pacific Weather and Climate. WORLD SCIENTIFIC, S. 288–313. DOI: 10.1142/9789812791139\_0014.
- Zhang, X., S. Liang, G. Wang, Y. Yao, B. Jiang & J. Cheng (2016). „Evaluation of the Reanalysis Surface Incident Shortwave Radiation Products from NCEP, ECMWF, GSFC, and JMA Using Satellite and Surface Observations“. In: *Remote Sensing* 8.3, S. 225. DOI: 10.3390/rs8030225.
- Zhang, Y. & M. G. Schaap (2017). „Weighted recalibration of the Rosetta pedotransfer model with improved estimates of hydraulic parameter distributions and summary statistics (Rosetta3)“. In: *Journal of Hydrology* 547, S. 39–53. DOI: 10.1016/j.jhydrol.2017.01.004.
- Zhang, Y., M. G. Schaap & Z. Wei (2020). „Development of Hierarchical Ensemble Model and Estimates of Soil Water Retention With Global Coverage“. In: *Geophysical Research Letters* 47.15. DOI: 10.1029/2020GL088819.
- Zobler, L. (1986). *A World Soil File for Global Climate Modeling: Technical Memorandum 87802*. Hrsg. von NASA Goddard Institute for Space Studies. New York, NY.

# Anhang

Umrechnung von regulären zu rotierten Koordinaten und zurück .....	287
Variablenliste in REMO .....	289
LSP II Hagemann (2002) .....	295
Aufbau und Zweck der einzelnen Skripte der REMO-Bodenbibliothek .....	298
Calltree von REMO .....	302
Zusätzliche Variablen der neuen Bodenbibliothek .....	304
Übersicht der klassenbasierten PTFs aus der Literatur .....	305
Literaturübersicht der PTF für die Feldkapazität (FCAP) .....	311
Literaturübersicht der PTF für den permanenten Welkepunkt (PWP) .....	314
Liste der Variablen mit signifikanten Unterschieden zwischen Modellläufen des Experimentplan 1 im GER-11 .....	316
Kernfunktionsdichteschätzung der klimatologischen Validationsdatensätze pro Jahreszeit für GER-11 .....	319
Vergleich von Modellläufen mit EOBSV21 .....	320

# A Umrechnung von regulären zu rotierten Koordinaten und zurück

0) Definitionen:

$lon$  = Longitude in regulären Koordinatensystem

$lat$  = Latitude in regulären Koordinatensystem

$\theta$  = Inklination (Rotation in y-Richtung)

$\phi$  = Azimut (Rotation in z-Richtung)

$rotpollon$  = reguläre Longitude des Nordpols im rotiertem Koordinatensystem

$rotpollat$  = reguläre Latitude des Nordpols im rotiertem Koordinatensystem

$rotlon$  = Longitude in rotiertem Koordinatensystem

$rotlat$  = Latitude in rotiertem Koordinatensystem

1) Umwandlung von Grad (Degree) in Bogensekunden (Radian):

regulär  $\rightarrow$  rotiert:

$$\begin{aligned}lon &= lon \cdot \frac{\pi}{180^\circ} \\lat &= lat \cdot \frac{\pi}{180^\circ} \\ \theta &= (rotpollat + 90^\circ) \cdot \frac{\pi}{180^\circ} \\ \phi &= rotpollon \cdot \frac{\pi}{180^\circ}\end{aligned}$$

rotiert  $\rightarrow$  regulär:

$$\begin{aligned}rotlon &= lon \cdot \frac{\pi}{180^\circ} \\rotlat &= lat \cdot \frac{\pi}{180^\circ} \\ \theta &= -(rotpollat + 90^\circ) \cdot \frac{\pi}{180^\circ} \\ \phi &= -rotpollon \cdot \frac{\pi}{180^\circ}\end{aligned}$$

2) Konvertieren von kartesischen zu sphärischen Koordinatensystem

regulär  $\rightarrow$  rotiert:

$$\begin{aligned}x &= \cos(lon) \cdot \cos(lat) \\y &= \sin(lon) \cdot \cos(lat) \\z &= \sin(lat)\end{aligned}$$

rotiert  $\rightarrow$  regulär:

$$\begin{aligned}x &= \cos(rotlon) \cdot \cos(rotlat) \\y &= \sin(rotlon) \cdot \cos(rotlat) \\z &= \sin(rotlat)\end{aligned}$$

3) Rotation:

regulär → rotiert:

$$\begin{pmatrix} xx \\ yy \\ zz \end{pmatrix} = \begin{pmatrix} \cos(\theta) \cdot \cos(\phi) & \cos(\theta) \cdot \sin(\phi) & \sin(\theta) \\ -\sin(\phi) & \cos(\phi) & 0 \\ -\sin(\theta) \cdot \cos(\phi) & -\sin(\theta) \cdot \sin(\phi) & \cos(\theta) \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

rotiert → regulär:

$$\begin{pmatrix} xx \\ yy \\ zz \end{pmatrix} = \begin{pmatrix} \cos(\theta) \cdot \cos(\phi) & \sin(\phi) & \sin(\theta) \cdot \cos(\phi) \\ -\cos(\theta) \cdot \sin(\phi) & \cos(\phi) & -\sin(\theta) \cdot \sin(\phi) \\ -\sin(\theta) & 0 & \cos(\theta) \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

4) Konvertieren von sphärisch zu kartesischen Koordinatensystem

regulär → rotiert:

$$rotlon = atan2(yy, xx)$$

$$rotlat = asin(zz)$$

rotiert → regulär:

$$lon = atan2(y, x)$$

$$lat = asin(z)$$

mit der Funktion atan2 definiert nach Organick (1966):

$$atan2(y, x) = \begin{cases} \arctan\left(\frac{y}{x}\right), & \text{für } x > 0 \\ \arctan\left(\frac{y}{x}\right) + \pi, & \text{für } x < 0 \text{ und } y > 0 \\ \arctan\left(\frac{y}{x}\right) - \pi, & \text{für } x < 0 \text{ und } y < 0 \\ \pm\pi, & \text{für } x < 0 \text{ und } y = 0 \\ \frac{\pi}{2}, & \text{für } x = 0 \text{ und } y > 0 \\ -\frac{\pi}{2}, & \text{für } x = 0 \text{ und } y < 0 \end{cases}$$

5) Umwandlung von Bogensekunden (Radian) zu Grad (Degree):

regulär → rotiert:

$$rotlon = lon \cdot \frac{\pi}{180^\circ}$$

$$rotlat = lat \cdot \frac{\pi}{180^\circ}$$

rotiert → regulär:

$$lon = rotlon \cdot \frac{180^\circ}{\pi}$$

$$lat = rotlat \cdot \frac{180^\circ}{\pi}$$

## B Variablenliste in REMO

Aggregationstyp:

sum: Summe / mean: Mittelwert / time: Wert des letzten Zeitschritts / -: Konstante / max: Maximumwert / min: Minimumwert

nl: Anzahl der Modelllevel / pl: Anzahl der Drucklevel für druckinterpolierte Daten

ID	Anzahl der Level	Abkürzung	Beschreibung	Einheit	Aggregationstyp							
						a-files (bisher)	a-files (neu)	t-files	f-files	e-files	n-files	m-files
1	1	WS1	soil wetness of layer 1	[m]	time			1	1		1	1
2	1	WS2	soil wetness of layer 2	[m]	time			1	1		1	1
3	1	WS3	soil wetness of layer 3	[m]	time			1	1		1	1
4	1	WS4	soil wetness of layer 4	[m]	time			1	1		1	1
5	1	WS5	soil wetness of layer 5	[m]	time			1	1		1	1
6	1	DZR	rooting depth	[m]	-			1				
7	1	DZS	soil depth until bedrock derived from textures	[m]	-			1				
8	1	FKSAT	saturated hydraulic conductivity (Clapp and Hornberger)	[ $\frac{m}{s}$ ]	-			1				
9	1	FMPOT	matrix potential	[m]	-			1				
10	1	BCLAPP	exponent b in Clapp and Hornberger		-			1				
11	1	VPOR	volumetric soil porosity	[ $\frac{m}{m}$ ]	-			1				
12	1	ETRANS	transpiration at the surface	[mm]	sum			1	1		1	1
13	1	EBSOIL	bare soil evaporation at the surface	[mm]	sum			1	1		1	1
14	nl-1	FTKVM	turbulent transfer coefficient of momentum in the atmosphere					1				
15	nl-1	FTKVH	turbulent transfer coefficient of heat in the atmosphere					1				
49	1	TRSOL	effective transmissivity					1				
51	2	EMTEF	emissivity (clear sky)					1				
52	2	TRSO	transmissivity (clear sky)					1				
53	1	DRAIN	drainage (part of 160)	[mm]	sum			1	1	1	1	1
54	1	TSL	surface temperature (land)	[K]	time	1	1	1	1	1	1	1
55	1	TSW	surface temperature (water)	[K]	time	1	1	1	1	1	1	1
56	1	TSI	surface temperature (ice)	[K]	time	1	1	1	1	1	1	1
57	1	USTR	surface u-stress (land)	[Pa]	mean			1	1		1	1
58	1	USTRW	surface u-stress (water)	[Pa]	mean			1	1		1	1
59	1	USTRI	surface u-stress (ice)	[Pa]	mean			1	1		1	1
60	1	VSTR	surface v-stress (land)	[Pa]	mean			1	1		1	1
61	1	VSTRW	surface v-stress (water)	[Pa]	mean			1	1		1	1
62	1	VSTRI	surface v-stress (ice)	[Pa]	mean			1	1		1	1
63	1	EVAPL	surface evaporation (land)	[mm]	sum			1	1		1	1

Fortsetzung auf der nächsten Seite

## B Variablenliste in REMO

Fortsetzung der vorherigen Seite

ID	Anzahl der Level	Abkürzung	Beschreibung	Einheit	Aggregationstyp							
						a-files (bisher)	a-files (neu)	t-files	f-files	e-files	n-files	m-files
64	1	EVAPW	surface evaporation (water)	[mm]	sum			1	1		1	1
65	1	EVAPI	surface evaporation (ice)	[mm]	sum			1	1		1	1
69	1	AHFSL	surface sensible heat flux (land)	$[\frac{W}{m^2}]$	mean			1				
70	1	AHFSW	surface sensible heat flux (water)	$[\frac{W}{m^2}]$	mean			1				
71	1	AHFSI	surface sensible heat flux (ice)	$[\frac{W}{m^2}]$	mean			1				
72	1	AZ0L	surface roughness length (land)	[m]	time			1	1		1	1
73	1	AZ0W	surface roughness length (water)	[m]	time			1	1		1	1
74	1	AZ0I	surface roughness length (ice)	[m]	time			1	1		1	1
75	1	ALSOL	surface albedo (land)	[frac]	time			1	1		1	1
76	1	ALSO W	surface albedo (water)	[frac]	time			1	1		1	1
77	1	ALSOI	surface albedo (ice)	[frac]	time			1	1		1	1
81	1	TMCHL	turbulent transfer coefficient of heat at the surface (land)					1				
82	1	TMCHW	turbulent transfer coefficient of heat at the surface (water)					1				
83	1	TMCHI	turbulent transfer coefficient of heat at the surface (ice)					1				
84	1	QDBL	specific humidity surface (land)	$[\frac{kg}{kg}]$	time	1	1	1	1	1	1	1
85	1	QDBW	specific humidity surface (water)	$[\frac{kg}{kg}]$	time			1	1		1	1
86	1	QDBI	specific humidity surface (ice)	$[\frac{kg}{kg}]$	time			1	1		1	1
87	1	BFLHSL	surface sensible heat flux (land) (= 69)	$[\frac{W}{m^2}]$	mean			1	1		1	1
88	1	BFLHSW	surface sensible heat flux (water) (= 70)	$[\frac{W}{m^2}]$	mean			1	1		1	1
89	1	BFLHSI	surface sensible heat flux (ice) (= 71)	$[\frac{W}{m^2}]$	mean			1	1		1	1
90	1	BFLQDSL	surface latent heat flux (land) (= 66)	$[\frac{W}{m^2}]$	mean			1	1		1	1
91	1	BFLQDSW	surface latent heat flux (water) (= 67)	$[\frac{W}{m^2}]$	mean			1	1		1	1
92	1	BFLQDSI	surface latent heat flux (ice) (= 68)	$[\frac{W}{m^2}]$	mean			1	1		1	1
93	1	AHFICE	sea-ice: conductive heat	$[\frac{W}{m}]$	mean			1	1		1	1
94	1	QRES	residual heat flux for melting sea ice	$[\frac{W}{m^2}]$	mean			1	1		1	1
95	1	SRFL	net solar radiative flux at the surface	$[\frac{W}{m^2}]$				1				
96	1	QDBOX S	horizontal transport of water vapour	$[\frac{kg}{m^2}]$	mean			1	1		1	1
97	1	QWBOX S	horizontal transport of cloud water	$[\frac{kg}{m^2}]$	mean			1	1		1	1
98	1	EKBOX S	horizontal transport of kinetic energy	$[\frac{3600J}{m^2}]$	mean			1	1		1	1

Fortsetzung auf der nächsten Seite

Fortsetzung der vorherigen Seite

ID	Anzahl der Level	Abkürzung	Beschreibung	Einheit	Aggregationstyp	Aggregationstyp						
						a-files (bisher)	a-files (neu)	t-files	f-files	e-files	n-files	m-files
99	1	FHBOXS	horizontal transport of sensible heat	$[\frac{3600J}{m^2}]$	mean			1	1		1	1
100	1	FIBOXS	horizontal transport of potential energy	$[\frac{3600J}{m^2}]$	mean			1	1		1	1
101	1	TLAMBDA	heat conductivity of dry soil	$[\frac{W}{K \cdot m}]$	-			1	1		1	1
103	1	DLAMBDA	parameter for increasing the heat conductivity of the soil due to soil moisture	$[\frac{W}{K \cdot m}]$	-			1	1		1	1
104	1	PORVOL	pore volume	[frac]	-			1	1		1	1
105	1	FCAP	field capacity of soil	$[\frac{m^3}{m^3}]$	-		1	1				
106	1	WI3	fraction of frozen soil of layer 1	[frac]	time			1	1		1	1
107	1	WI4	fraction of frozen soil of layer 2	[frac]	time			1			1	1
108	1	WI5	fraction of frozen soil of layer 3	[frac]	time			1	1		1	1
109	1	WI	fraction of frozen soil of layer 4	[frac]	time			1			1	1
110	1	WICL	fraction of frozen soil of layer 5	[frac]	time			1	1		1	1
112	1	QDB	specific humidity surface	$[\frac{kg}{kg}]$	time		1	1	1		1	1
129	1	FIB	surface geopotential (orography)	[m]	-	1	1	1				
130	nl	T	temperature	[K]		1	1	1	1			
131	nl	U	u-velocity	$[\frac{m}{s}]$		1	1	1	1			
132	nl	V	v-velocity	$[\frac{m}{s}]$		1	1	1	1			
133	nl	QD	specific humidity	$[\frac{kg}{kg}]$		1	1	1	1			
134	1	PS	surface pressure	[Pa]	time	1	1	1	1	1		1
135	pl	OMEGA	vertical velocity	$[\frac{Pa}{s}]$			1	1				
139	1	TS	surface temperature (mean over gridbox)	[K]	time		1	1	1		1	1
140	1	WS	soil wetness	[m]	time	1	1	1	1	1	1	1
141	1	SN	snow depth	[m]	time	1	1	1	1	1		1
142	1	APRL	large scale precipitation	[mm]	sum			1	1	1	1	1
143	1	APRC	convective precipitation	[mm]	sum			1	1	1	1	1
144	1	APRS	snow fall	[mm]	sum			1	1	1	1	1
145	1	VDIS	boundary layer dissipation	$[\frac{W}{m^2}]$	mean			1	1		1	1
146	1	AHFS	surface sensible heat flux	$[\frac{W}{m^2}]$	mean			1	1		1	1
147	1	AHFL	surface latent heat flux	$[\frac{W}{m^2}]$	mean			1	1		1	1
151	1	PSRED	mean sealevel pressure	[Pa]	time				1		1	1
153	nl	QW	liquid water content	$[\frac{kg}{kg}]$		1	1	1	1			
156	nl	FI	geopotential height	[m]		1	1	1	1			
159	1	USTAR3	TKE for ocean mixed layer (accumulated)	$[\frac{m^3}{s^3}]$	mean			1	1		1	1
160	1	RUNOFF	total runoff (surface runoff + drainage)	[mm]	sum			1	1	1	1	1
162	nl	ACLC	cloud cover (see also 223)	[frac]	time		1	1				
163	1	ACLCV	total cloud cover (see also 164)	[frac]	time			1	1		1	1
164	1	ACLCOV	total cloud cover	[frac]	mean			1	1		1	1

Fortsetzung auf der nächsten Seite

*B Variablenliste in REMO*

Fortsetzung der vorherigen Seite

ID	Anzahl der Level	Abkürzung	Beschreibung	Einheit	Aggregationstyp	Aggregationstyp							
						a-files (bisher)	a-files (neu)	t-files	f-files	e-files	n-files	m-files	s-files
165	1	U10	10m u-velocity (rotated grid)	$[\frac{m}{s}]$	mean				1	1		1	1
166	1	V10	10m v-velocity (rotated grid)	$[\frac{m}{s}]$	mean				1	1		1	1
167	1	TEMP2	2m temperature	[K]	mean				1	1	1	1	1
168	1	DEW2	2m dew point temperature	[K]	mean				1	1	1	1	1
169	1	TSURF	surface temperature (land) (see also 54)	[K]	mean				1	1		1	1
170	1	TD	soil temperature of layer 4	[K]	time	1	1	1	1	1		1	1
171	1	WIND10	10m windspeed	$[\frac{m}{s}]$	mean				1	1		1	1
172	1	BLA	land sea mask	[frac]	-	1		1	1				
173	1	AZ0	surface roughness length	[m]	time	1			1	1		1	1
174	1	ALB	surface background albedo	[frac]	-	1		1	1	1		1	1
175	1	ALBEDO	surface albedo	[frac]	time				1	1		1	1
176	1	SRADS	net surface solar radiation	$[\frac{W}{m^2}]$	mean				1	1		1	1
177	1	TRADS	net surface thermal radiation	$[\frac{W}{m^2}]$	mean				1	1		1	1
178	1	SRAD0	net top solar radiation	$[\frac{W}{m^2}]$	mean				1	1		1	1
179	1	TRAD0	top thermal radiation (OLR)	$[\frac{W}{m^2}]$	mean				1	1		1	1
180	1	USTR	surface u-stress	[Pa]	mean				1	1		1	1
181	1	VSTR	surface v-stress	[Pa]	mean				1	1		1	1
182	1	EVAP	surface evaporation	[mm]	sum				1	1	1	1	1
183	1	TDCL	soil temperature of layer 5	[K]	time	1	1	1	1	1		1	1
185	1	SRAFS	net surf. solar radiation (clear sky)	$[\frac{W}{m^2}]$	mean				1				
186	1	TRAFS	net surf. thermal radiation (clear sky)	$[\frac{W}{m^2}]$	mean				1	1		1	1
187	1	SRAF0	net top solar radiation (clear sky)	$[\frac{W}{m^2}]$	mean				1				
188	1	TRAF0	net top thermal radiation (clear sky)	$[\frac{W}{m^2}]$	mean				1	1		1	1
189	1	SCLFS	surface solar cloud forcing	$[\frac{W}{m^2}]$	mean				1				
190	1	TCLFS	surface thermal cloud forcing	$[\frac{W}{m^2}]$	mean				1	1		1	1
191	1	SCLF0	top solar cloud forcing	$[\frac{W}{m^2}]$	mean				1				
192	1	TCLF0	top thermal cloud forcing	$[\frac{W}{m^2}]$	mean				1	1		1	1
194	1	WL	skin reservoir content	[m]	time	1	1	1	1	1		1	1
195	1	USTRGW	u-gravity wave stress	[Pa]	mean				1				
196	1	VSTRGW	v-gravity wave stress	[Pa]	mean				1				
197	1	VDISGW	gravity wave dissipation	$[\frac{W}{m^2}]$	mean				1				
198	1	VGRAT	vegetation ratio	[frac]	-	1			1	1		1	1
199	1	VAROR	orographic variance (for surface runoff)	$[m^2]$	-	1		1	1				
200	1	VLT	leaf area index	$[\frac{m^2}{m^2}]$	-	1			1	1		1	1
201	1	T2MAX	maximum 2m-temperature	[K]	max				1	1	1	1	1
202	1	T2MIN	minimum 2m-temperature	[K]	min				1	1	1	1	1
203	1	SRAD0U	top solar radiation upward	$[\frac{W}{m^2}]$	mean				1	1		1	1
204	1	SRADSU	surface solar radiation upward	$[\frac{W}{m^2}]$	mean				1	1		1	1

Fortsetzung auf der nächsten Seite



Fortsetzung der vorherigen Seite

ID	Anzahl der Level	Abkürzung	Beschreibung	Einheit	Aggregationstyp										
						a-files (bisher)	a-files (neu)	t-files	f-files	e-files	n-files	m-files	s-files		
205	1	TRADSU	surface thermal radiation upward	$[\frac{W}{m^2}]$	mean				1	1					
206	1	TSN	snow temperature	$[K]$	time	1	1	1	1	1			1	1	
207	1	TD3	soil temperature of layer 1	$[K]$	time	1	1	1	1	1			1	1	
208	1	TD4	soil temperature of layer 2	$[K]$	time	1	1	1	1	1			1	1	
209	1	TD5	soil temperature of layer 3	$[K]$	time	1	1	1	1	1			1	1	
210	1	SEAICE	sea ice cover	[frac]	time	1	1	1	1	1			1	1	
211	1	SICED	sea ice depth	$[m]$	time			1	1	1			1	1	
212	1	FOREST	vegetation type	[frac]	-	1		1	1						
213	1	TEFF	(effective) sea-ice skin temperature	$[K]$	mean				1	1			1	1	
214	1	TSMAX	maximum surface temperature	$[K]$	max				1	1			1	1	
215	1	TSMIN	minimum surface temperature	$[K]$	min				1	1			1	1	
216	1	WIMAX	maximum 10m-wind speed (w/o gusts)	$[\frac{m}{s}]$	max				1	1			1	1	
217	1	TOPMAX	maximum height of convective cloud tops	$[Pa]$	max				1	1			1	1	
218	1	SNMEL	snow melt	$[mm]$	sum				1	1			1	1	
220	1	TSLIN	land: residual surface heat budget / sea-ice: conductive heat flux	$[\frac{W}{m^2}]$	mean				1	1			1	1	
221	1	DSNAC	snow depth change	$[mm]$	sum				1	1			1	1	
222	nl	EMTER	effective emissivity						1						
223	nl	ACLCAC	cloud cover	[frac]	mean				1						
224	nl	TKE	turbulent kinetic energy	$[\frac{m^2}{s^2}]$					1	1					
226	1	FAO	FAO texture class	[0-6]	-	1		1	1						
227	1	RGCGN	heat capacity of soil	$[\frac{J}{m^3 K}]$	-			1	1						
229	1	WSMX	field capacity of soil	$[m]$	-	1		1	1						
230	1	QVI	vertically integrated specific humidity	$[\frac{kg}{m^2}]$	mean				1	1			1	1	
231	1	ALWCVI	vertically integrated liquid water cont.	$[\frac{kg}{m^2}]$	mean				1	1			1	1	
232	1	GLAC	glacier mask (0: no/ 1: yes)		time	1	1	1	1						
265	1	U10er	10m u-velocity (regular grid)	$[\frac{m}{s}]$	mean					1			1	1	
266	1	V10er	10m v-velocity (regular grid)	$[\frac{m}{s}]$	mean					1			1	1	
271	1	GHPBL	height of planetary boundary layer	$[m]$	time					1			1	1	
272	1	BETA	shape parameter		-	1		1							
273	1	WMINLOK	minimal subgrid field capacity	$[m]$	-	1		1							
274	1	WMAXLOK	maximal subgrid field capacity	$[m]$	-	1		1							
275	1	VBM10M	maximum of the expected gust velocity near the surface	$[\frac{m}{s}]$	max				1	1			1	1	
276	1	CAPE	convective available potential energy	$[\frac{J}{kg}]$	mean				1	1			1	1	
281	nl	PINT	non-hydrostatic pressure	$[Pa]$					1	1					

Fortsetzung auf der nächsten Seite

## B Variablenliste in REMO

Fortsetzung der vorherigen Seite

ID	Anzahl der Level	Abkürzung	Beschreibung	Einheit	Aggregationstyp	Aggregationstyp							
						a-files (bisher)	a-files (neu)	t-files	f-files	e-files	n-files	m-files	s-files
282	nl	DWDT	non-hydrostatic acceleration	$[\frac{m^2}{s^2}]$				1	1				
283	nl	W	non-hydrostatic vertical velocity	$[\frac{m}{s}]$				1	1				
301	1	ESNOW	evaporation over snow at the surface	$[mm]$	sum				1	1		1	1
302	1	ESKIN	evaporation from skin reservoir at the surface	$[mm]$	sum				1	1		1	1
303	1	ERES			sum				1	1		1	1
334	nl	QI	cloud ice	$[\frac{kg}{m^3}]$					1				
336	1	QIVI	vertically integrated cloud ice	$[\frac{kg}{m^2}]$	mean				1	1		1	1
337	1	QIBOXS	horizontal transport of cloud ice	$[\frac{kg}{m^2}]$					1				
433	nl	RPR	warm cloud rain production		sum				1				
456	1	BFLHS	surface sensible heat flux (=146)	$[\frac{W}{m^2}]$	mean				1	1		1	1
457	1	BFLQDS	surface latent heat flux (=147)	$[\frac{W}{m^2}]$	mean				1	1		1	1
461	1	TMCM	turbulent transfer coefficient of momentum at the surface						1				
462	1	TMCH	turbulent transfer coefficient of heat at the surface						1				
614	1	PHI	latitude in real coordinates	$[^\circ N]$	-	1	1	1	1				
615	1	RLA	longitude in real coordinates	$[^\circ E]$	-	1	1	1	1				

## C LSP II Hagemann (2002)

**Anmerkung:** Fehlwerte oder nicht vorkommende Klassen sind mit -1. markiert, Klasse 5 und 26 sind zwar vom Namen doppelt, aber die Parameterwerte unterscheiden sich. In dem Skript, welches die Parameter auf die Pixel des jeweiligen Untersuchungsgebietes anwendet, werden die beiden Klassen abhängig von den Latitude- und Longitudewerten verwendet.

LS-Klasse	$\alpha_{veg}$	$z_{veg}$	$c_{veg_g}$	$c_{veg_d}$	$LAI_g$	$LAI_d$	$c_{for}$	$W_{ava}$	$f_{PWP}$
1 Urban	0.2	2.5	0.	0.	0.	0.	0.	0.	0.48
2 Low Sparse Grassland	0.19	0.03	0.55	0.04	1.75	0.2	0.	580.	0.45
3 Coniferous Forest	0.13	1.	0.96	0.95	9.2	9.0	0.9	130.	0.41
4 Deciduous Conifer Forest	0.13	1.	0.56	0.	3.7	0.1	0.56	155.	0.45
5 Deciduous Broadleaf Forest	0.16	1.	0.8	0.	5.1	0.1	0.8	300.	0.53
6 Evergreen Broadleaf Forests	0.16	0.68	0.96	0.95	9.9	9.5	0.95	200.	0.46
7 Tall Grasses and Shrubs	0.2	0.1	0.44	0.	1.5	0.1	0.	280.	0.50
8 Bare Desert	0.28	0.005	0.	0.	0.	0.	0.	100.	0.
9 Upland Tundra	0.17	0.03	0.51	0.	2.2	0.4	0.	120.	0.34
10 Irrigated Grassland	0.16	0.03	0.9	0.1	4.5	0.	0.	450.	0.49
11 Semi Desert	0.28	0.005	0.24	0.	0.7	0.	0.	300.	0.48
12 Glacier Ice	0.7	0.005	0.	0.	0.	0.	0.	0.	0.
13 Wooded Wet Swamp	0.12	0.03	0.73	0.67	3.4	3.0	0.	235.	0.50
14 Inland Water	0.07	0.0002	0.	0.	0.	0.	0.	0.	0.
15 Sea Water	0.07	0.0002	0.	0.	0.	0.	0.	0.	0.
16 Shrub Evergreen	0.165	0.55	0.39	0.17	5.1	1.7	0.24	410.	0.45
17 Shrub Deciduous	0.16	0.26	0.53	0.1	4.6	0.5	0.26	350.	0.32
18 Mixed Forest and Field	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
19 Evergreen Forest and Fields	0.16	0.45	0.85	0.4	6.	3.	0.5	200.	0.47
20 Cool Rain Forest	0.12	2.	0.96	0.96	9.3	9.3	0.95	80.	0.35
21 Conifer Boreal Forest	0.13	1.	0.52	0.52	6.	6.	0.46	140.	0.33
22 Cool Conifer Forest	0.13	1.	0.7	0.35	9.2	4.	0.66	380.	0.48
23 Cool Mixed Forest	0.15	1.	0.93	0.02	4.3	0.1	0.93	140.	0.40
24 Mixed Forest	0.16	0.68	0.97	0.3	7.	1.	0.83	220.	0.51
25 Cool Broadleaf Forest	0.16	1.	0.88	0.16	5.2	0.85	0.85	210.	0.43
26 Deciduous Broadleaf Forest	0.16	1.	0.95	0.19	5.3	0.95	0.92	320.	0.47
27 Conifer Forest	0.13	1.	0.87	0.31	9.7	4.0	0.84	250.	0.49
28 Montane Tropical Forests	0.15	0.55	0.9	0.77	4.8	4.0	0.52	700.	0.50
29 Seasonal Tropical Forest	0.12	2.	0.99	0.69	6.1	2.7	0.98	210.	0.53
30 Cool Crops and Towns	0.18	0.25	0.9	0.14	2.5	0.74	0.14	280.	0.475
31 Crops and Town	0.18	0.25	0.85	0.16	4.4	1.1	0.16	430.	0.50
32 Dry Tropical Woods	0.14	0.55	0.98	0.19	6.1	2.2	0.65	550.	0.44
33 Tropical Rainforest	0.12	2.	0.96	0.96	9.3	9.3	0.95	240.	0.52
34 Tropical Degraded Forest	0.14	0.55	0.87	0.24	6.	2.2	0.57	260.	0.52
35 Corn and Beans Cropland	0.18	0.1	0.9	0.08	2.5	0.4	0.	250.	0.495
36 Rice Paddy and Field	0.15	0.06	0.95	0.19	4.6	0.26	0.	350.	0.49
37 Hot Irrigated Cropland	0.18	0.05	0.83	0.28	4.4	1.4	0.	390.	0.51

Fortsetzung auf der nächsten Seite

Fortsetzung der vorherigen Seite

LS-Klasse	$\alpha_{veg}$	$z_{veg}$	$C_{veg}$	$C_{veg}$	$LAI_g$	$LAI_d$	$C_{for}$	$W_{ava}$	$f_{PWP}$
38 Cool Irrigated Cropland	0.18	0.05	0.6	0.	3.	0.	0.	370.	0.48
39 Cold Irrigated Cropland	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
40 Cool Grasses and Shrubs	0.19	0.06	0.59	0.01	1.8	0.05	0.	450.	0.44
41 Hot and Mild Grasses, Shrubs	0.2	0.1	0.58	0.16	1.85	0.85	0.	650.	0.40
42 Cold Grassland	0.19	0.03	0.98	0.02	3.	0.14	0.	260.	0.52
43 Savanna (Woods)	0.16	0.25	0.8	0.27	2.6	0.94	0.4	660.	0.485
44 Mire, Bog, Fen	0.12	0.03	0.67	0.	2.6	0.1	0.	160.	0.39
45 Marsh Wetland	0.12	0.03	0.85	0.	3.1	0.1	0.	800.	0.55
46 Mediterranean Scrub	0.15	0.46	0.8	0.66	4.3	0.28	0.4	620.	0.51
47 Dry Woody Scrub	0.16	0.26	0.34	0.2	4.	1.	0.2	610.	0.38
48 Dry Evergreen Woods	0.18	0.04	0.6	0.3	1.8	1.7	0.53	400.	0.46
49 Volcanic Rock	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
50 Sand Desert	0.28	0.005	0.	0.	0.	0.	0.	100.	0.
51 Semi Desert Shrubs	0.28	0.005	0.24	0.08	0.73	0.41	0.	470.	0.40
52 Semi Desert Sage	0.28	0.005	0.39	0.	1.18	0.	0.	840.	0.425
53 Barren Tundra	0.17	0.03	0.14	0.	1.6	0.	0.	60.	0.31
54 Cool South. Hemi. Mix-Forests	0.16	0.65	0.81	0.	4.7	0.1	0.75	260.	0.38
55 Cool Fields and Woods	0.18	0.17	0.9	0.3	3.	0.76	0.3	180.	0.45
56 Forest and Field	0.16	0.25	0.8	0.21	5.9	2.5	0.45	310.	0.50
57 Cool Forest and Field	0.16	0.25	0.9	0.2	4.1	1.	0.5	180.	0.50
58 Fields and Woody Savanna	0.18	0.17	0.95	0.35	5.	2.3	0.35	700.	0.48
59 Succulent and Thorn Scrub	0.2	0.03	0.85	0.24	4.7	0.9	0.4	740.	0.49
60 Small Leaf Mixed Woods	0.15	1.	0.47	0.	3.7	0.1	0.47	130.	0.43
61 Deciduous, Mix. Boreal Forest	0.16	0.65	0.64	0.	4.7	0.1	0.6	170.	0.46
62 Narrow Conifers	0.15	0.31	0.48	0.	3.4	0.1	0.27	180.	0.30
63 Wooded Tundra	0.18	0.05	0.47	0.15	3.1	0.5	0.34	100.	0.38
64 Heath Scrub	0.2	0.1	0.5	0.	4.6	0.1	0.	90.	0.45
65 Coastal Wetland, NW	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
66 Coastal Wetland, NE	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
67 Coastal Wetland, SE	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
68 Coastal Wetland, SW	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
69 Polar and Alpine Desert	0.28	0.005	0.	0.	0.	0.	0.	35.	0.43
70 Glacier Rock	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
71 Salt Playas	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
72 Mangrove	0.12	1.29	0.95	0.95	9.	9.	0.9	290.	0.50
73 Water and Island Fringe	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
74 Land, Water, and Shore (=14)	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
75 Land and Water, Rivers (=14)	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
76 Crop and Water Mixtures	0.15	0.06	0.65	0.13	4.4	0.16	0.	2000.	0.575
77 Southern Hemisphere Conifers	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
78 Southern Hemi. Mixed Forest	0.16	0.65	0.93	0.34	5.	2.6	0.87	235.	0.46
79 Wet Sclerophylic Forest	0.16	0.65	0.87	0.5	4.8	2.6	0.8	300.	0.55
80 Coastline Fringe	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
81 Beaches and Dunes	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
82 Sparse Dunes and Ridges	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
83 Bare Coastal Dunes	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
84 Residual Dunes and Beaches	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.

Fortsetzung auf der nächsten Seite

Fortsetzung der vorherigen Seite

LS-Klasse	$\alpha_{veg}$	$z_{veg}$	$c_{veg}$	$c_{vega}$	$LAI_g$	$LAI_d$	$c_{for}$	$W_{ava}$	$f_{PWP}$
85 Compound Coastlines	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
86 Rocky Cliffs and Slopes	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
87 Sandy Grassland and Shrubs	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
88 Bamboo	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
89 Moist Eucalyptus	0.16	0.65	0.75	0.59	4.7	3.1	0.7	210.	0.54
90 Rain Green Tropical Forest	0.12	2.	0.96	0.96	9.3	9.3	0.95	360.	0.42
91 Woody Savanna	0.16	0.25	0.95	0.32	1.9	1.	0.4	560.	0.485
92 Broadleaf Crops	0.17	0.175	0.95	0.12	5.	2.	0.3	240.	0.42
93 Grass Crops	0.185	0.1	0.91	0.2	2.5	1.1	0.2	240.	0.47
94 Crops, Grass, Shrubs	0.19	0.1	0.65	0.33	2.7	0.4	0.	530.	0.46
95 Evergreen Tree Crop	0.16	0.17	0.97	0.3	6.1	3.	0.5	250.	0.46
96 Deciduous Tree Crop	0.17	0.17	0.9	0.3	5.	2.	0.3	240.	0.51
97 not used	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
98 not used	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
99 not known	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.	-1.
100 Missing data (Small Islands)	0.2	0.03	0.53	0.2	4.1	1.	0.	180.	0.49

# D Aufbau und Zweck der einzelnen Skripte der REMO-Bodenbibliothek

Nach dem Erstellen der Ordner und der Verknüpfungen mit den jeweiligen Ausgangsdaten, dem Kopieren der 33 Topographiedaten in den dafür vorgesehenen Ordner sowie der Erzeugung des rotierten Gitters für das ausgewählte Untersuchungsgebiet, werden im ersten Skript der BodLibKit alle konstanten Landoberflächenparameter, die aus dem GLCCD generiert werden, erstellt. Dabei werden der Datensatz und die jeweilige Klassifikationstabelle (Standard: LSPII) eingelesen und anschließend für jeden Parameter mit einer flächenanteiligen Mittelwertbildung auf jeden Modellpixel interpoliert:

$$x_i = \sum_{k=1}^n p_k \cdot w_{k,i} \quad (\text{D.1})$$

mit:

$i$  = Pixelindex

$k$  = Landbedeckungsklasse

$n$  = maximale Anzahl an Klassen

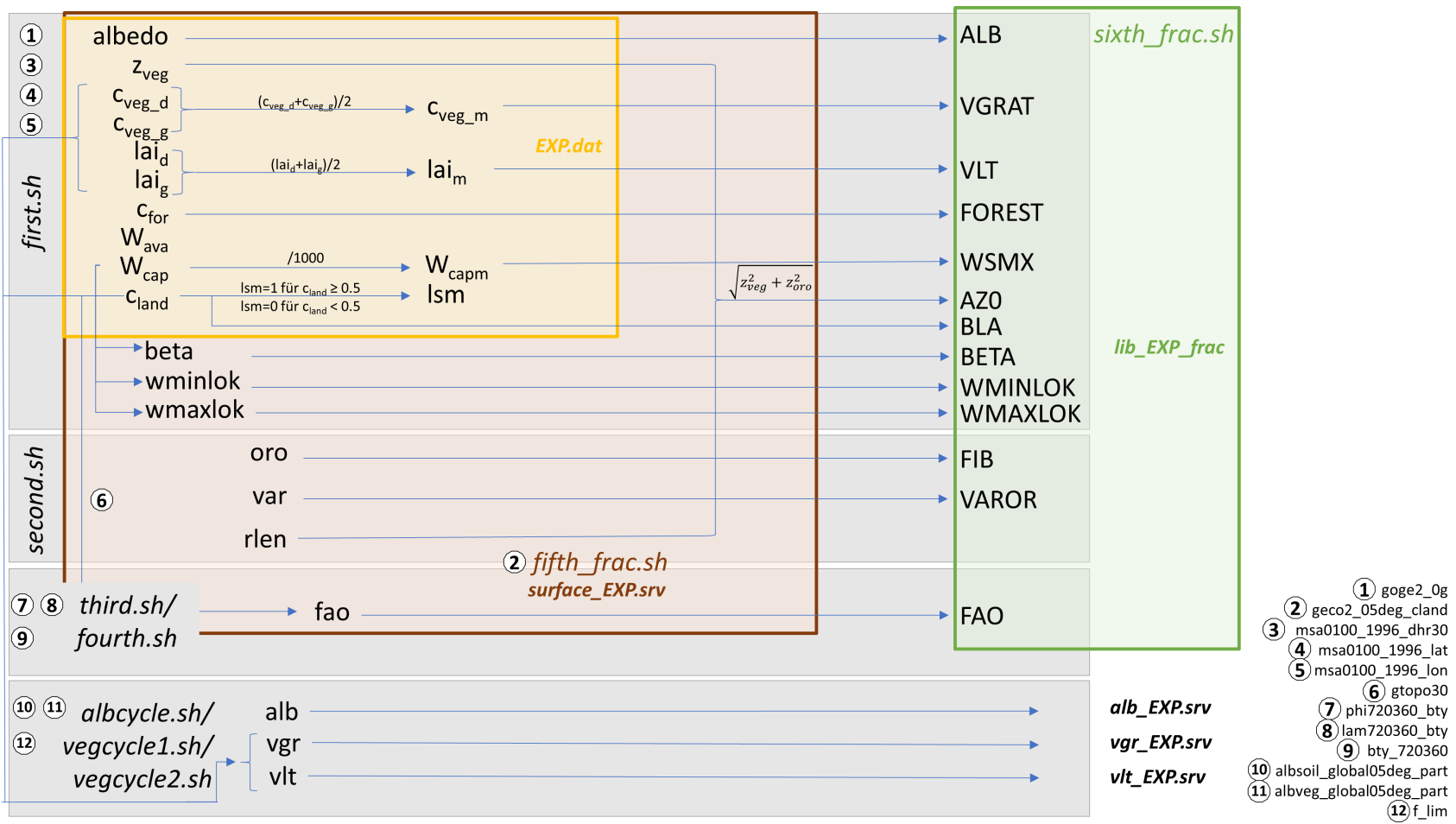
$w_{k,i} = \frac{a_{k,i}}{\sum_{j=1}^n a_{j,i}}$  = Flächenanteil der Klasse  $k$  pro Modellpixel  $i$

$a_{k,i}$  = Fläche der jeweiligen Klasse in einem Pixel  $i$

$p_k$  = Wert für den jeweiligen Parameter für Klasse  $k$

Es ergeben sich zehn Karten für das Untersuchungsgebiet (vgl. *first.sh* in unten stehender Abbildung). In einem weiteren Schritt wird der  $W_{cap}$ -Wert durch 1000 geteilt, um die SI-Einheit Meter zu erreichen. Die Variablen  $c_{veg_d/g}$  bzw.  $lai_{d/g}$  werden am Ende der Prozessierungskette für die Erstellung der Vegetationszyklusdaten als Eingangsdaten verwendet. Die Mittelwerte dieser aus der Growing- und Dormancy-Phase stammenden Maximal- und Minimalwerte für Vegetationsanteil und LAI werden ebenfalls in diesem Skript berechnet und als die Konstanten VLT und VGRAT in Skript 6 in die finale BodLibKit-Datei geschrieben.

Für die Albedo (ALB) wird der anhand des GLCCD und der LSPII Tabelle jedem Pixel zugewiesene Bodenalbedowert für einen Bereich über Afrika durch die ebenfalls eingelesenen Meteosat Albedo-Werte automatisch ersetzt, wenn dieser Bereich im Untersuchungsgebiet liegt. Allerdings werden, wie oben beschrieben, dieser im ersten Skript erzeugte Albedo-Wert und auch VGRAT und VLT standardmäßig seit 2009 durch die 12 Werte des Vegetationsjahreszyklus im Klimamodelllauf ersetzt. Die konstanten Mittelwerte werden trotzdem immer noch in die finale BodLibKit-Datei geschrieben.



Übersicht des BodLibKit Workflows mit den einzelnen Skripten und den jeweils für die Zwischenschritte verwendeten Datensätzen

Außerdem liefert das erste Skript eine binäre Land-Meer-Maske, die aus dem jeweiligen Anteil der Landfläche des Landbedeckungsdatensatzes innerhalb eines Modellpixels ( $c_{land}$ ) abgeleitet wird. Pixel mit einem Anteil an Landpunkten größer oder gleich 0.5 werden in der Datei „lsm“ als Land, der Rest als Meer klassifiziert. Diese Karte und auch die zusätzlich erzeugten Anteile verschiedener Landbedeckungsklassen, wie z. B. der Stadtanteil pro Pixel (nicht im Schaubild abgebildet), werden zwar herausgeschrieben, aber nicht weiter im Prozess verwendet und damit auch nicht ans Klimamodell übergeben.

Des Weiteren wird in der aktuellen Version der minimale und maximale Feldkapazitätswert ( $WMINLOK$ ,  $WMAXLOK$ ) mittels der auf der gewählten Modellauflösung berechneten Feldkapazität  $W_{cap}$  und der auf  $\frac{1}{10}$  dieser Modellauflösung berechneten Feldkapazität innerhalb eines jeden Pixels ermittelt. Diese werden anschließend dazu benutzt den die Verteilungskurve dieser Kapazitätswerte formenden Koeffizient  $BETA$  zu berechnen. Dieser wird im Modell für das Arno-Schema benötigt. Alle drei Variablen werden separat herausgeschrieben.

Aus den topographischen Daten werden im zweiten Skript der Mittelwert der Höhe (Formel D.2), die Varianz (Formel D.3) und die Rauigkeitslänge (Formel D.4) der Orographie in einem Pixel wie folgt berechnet:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad (D.2)$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (D.3)$$

$$z_{0,oro} = c \cdot s^2 \cdot atan\left(\sqrt{\frac{fläche}{L}}\right) \quad (D.4)$$

mit:

$n$  = Anzahl der Pixel von GTOPO, die in einem Modellpixel liegen

$x_i$  = Höhe in m

$\bar{x}$  = Mittelwert der Höhe von allen GTOPO-Pixeln, die in einem Modellpixel liegen

$s^2$  = Varianz der Orographie in einem Modellpixel in  $m^2$

$c$  = 0.00001

$L$  = 2.25

$fläche = (40000 \cdot \frac{DPHI}{360}) \cdot (40000 \cdot \frac{DLAM}{360})$

$DPHI$  = Modellauflösung in Longitude-Richtung

$DLAM$  = Modellauflösung in Latitude-Richtung

Die in den Klimamodelllauf eingehende Rauigkeitslänge setzt sich aus der Rauigkeitslänge der Orographie und der Vegetation, wie folgt zusammen:



---

$$z_0 = \sqrt{z_{0,Oro}^2 + z_{0,Veg}^2} \quad (D.5)$$

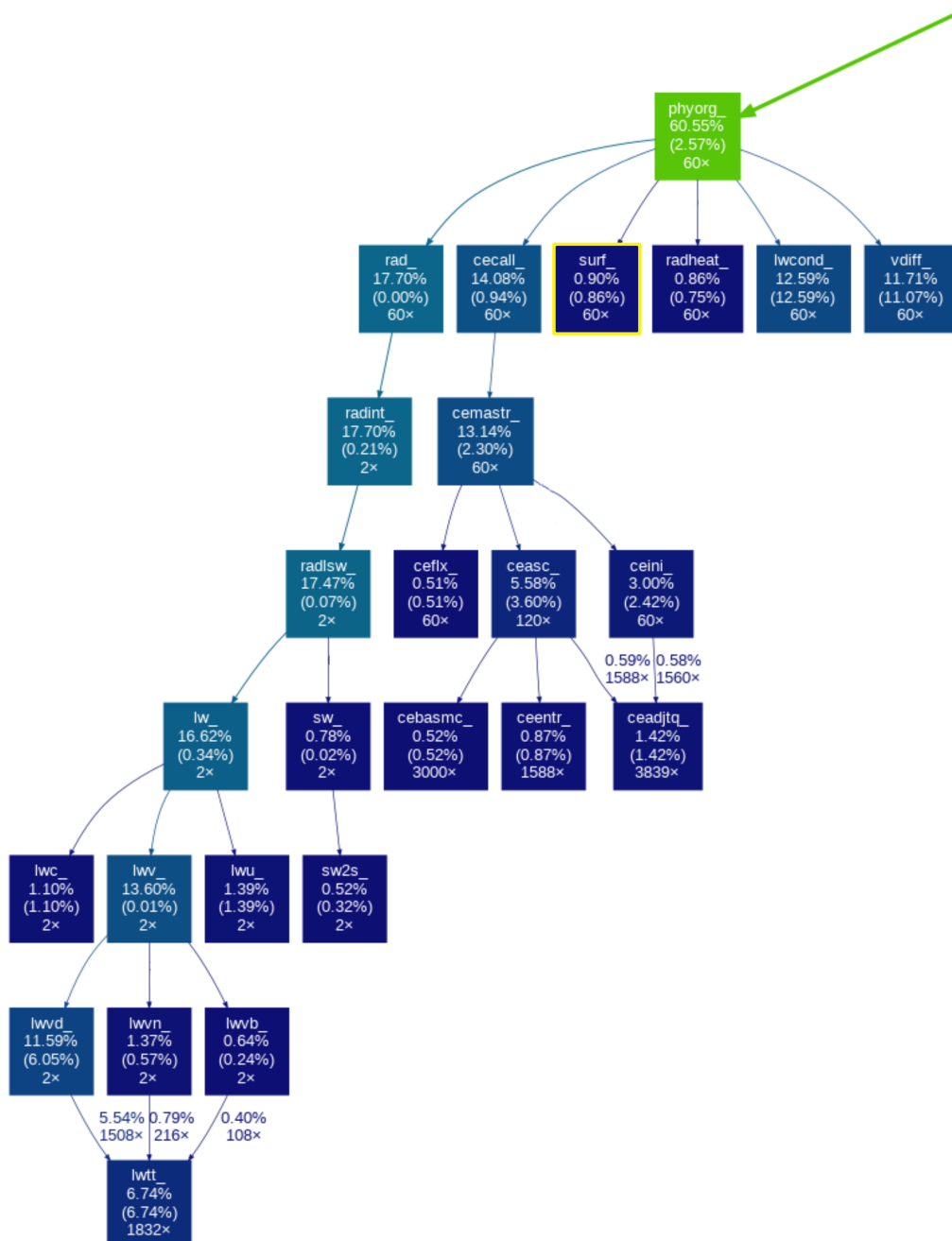
Diese Verknüpfung beider Rauigkeitslängen findet aber erst im fünften Skript statt. Zuvor werden im dritten Skript die Koordinatenindizes des Bodendatensatzes bestimmt, die das Modellgebiet abdecken. Im vierten Skript wird mit einem stark vereinfachten Cressmann-Schema (vgl. Cressman, 1959, Majewski, 1991) nach der Bodentexturklasse für jeden Pixel gesucht. Dabei werden die im dritten Skript erzeugten Indizes als Startwert für ein sich iterativ von dort ausweitendes Fenster benutzt, um im Fenster nach dem am nächsten liegenden Landpunkt, der durch die im ersten Skript generierte Variable  $c_{land}$  auf Modellauflösung definiert ist, gesucht (vgl. obere Reihe in Abb. 11). Da es sich bei den zu interpolierenden Werten um die nominal skalierte Bodentextur handelt und nicht um metrisch skalierte Variablen, fällt die im Cressmann-Schema übliche Distanzgewichtung weg und der Wert des gefundenen Pixels der Ausgangsdaten wird für den Modellpixel unverändert übernommen.

Im fünften Skript werden alle aus den unterschiedlichen Datenquellen abgeleitete Variablen zusammengebracht und in einer eigenen Datei gespeichert. In Skript 6 werden diese 12 in Tabelle 6 aufgelisteten Variablen aus der Sammeldatei aus Skript 5 (*surface\_EXP.srv*) eingelesen, die Bodentextur in die finalen sechs Klassen reklassifiziert (vgl. Tab. 5) und dann im Zielformat (ieg) so mit den jeweiligen Header abgespeichert, dass der Präprozessor zur Erzeugung der a-files diese finale BodLibKit-Datei einlesen kann.

Im siebten und achten Skript (*vegcycle1/2.sh*) werden die Daten für den Jahresgang des Vegetationsanteils und des LAIs mittels des eingelesenen Wachstumsfaktors und der minimalen und maximalen Werte pro Pixel berechnet und als srv-Datei abgespeichert. Zusammen mit dem bereits erwähnten im neunten Skript erzeugten Albedozyklus, werden diese Dateien bei Bedarf direkt von REMO eingelesen.

# E Calltree von REMO

Die linke und rechte Seite des Calltrees wird über die grüne Verbindungslinie zwischen den Subroutinen *phyec* und *phyorg* verbunden. Dargestellt sind nur Programme, die einen Rechenzeitanteil über 0.5% haben. Die Farbgebung ist abhängig von der jeweiligen Rechenzeit der Routine und deren Subroutinen. In den Boxen steht jeweils der Name, die kumulierte Rechenzeit des Programms und dessen Kinder, die Zeit des Programms und die Anzahl, wie oft dieses Programm aufgerufen wird. Wird eine Subroutine von mehr als einem Programm aufgerufen, sind die jeweiligen Zahlen am Pfeil aufgetragen. Die gelb umrandete Subroutine wurde in der Arbeit stark überarbeitet.





## F Zusätzliche Variablen der neuen Bodenbibliothek

Für alle aufgeführten Variablen außer DZR und DZS liegen jeweils sechs Werte vor. Der mit dem Index 0 versehene Wert entspricht dem gewichteten Mittelwert über die gesamte Bodentiefe, während die Werte mit Index 1-5 sich auf die jeweiligen REMO-Bodenschichten beziehen.

<sup>1</sup>: Y. Yang et al. (2016)

<sup>2</sup>: SoilGrids1km, Hengl et al. (2014)

<sup>3</sup>: Montzka et al. (2017)

<sup>4</sup>: abgeleitet aus den jeweiligen SAND-, CLAY- und FORG-Werten (vgl. Abel et al. (2020))

Name	Beschreibung	Einheit
SAND0-5	Sandanteil	%
CLAY0-5	Tonanteil	%
FORG0-5	organischer Anteil	-
DZR <sup>1</sup>	Wurzeltiefe	m
DZS <sup>2</sup>	Bodentiefe	m
FKSAT0-5 <sup>3</sup>	gesättigte hydraulische Konduktivität	$\frac{\text{m}}{\text{s}}$
FMPOT0-5 <sup>4</sup>	gesättigtes Matrixpotential	m
BCLAPP0-5 <sup>4</sup>	Clapp Hornberger Parameter	-
VPOR0-5 <sup>4</sup>	Porenvolumen	$\frac{\text{m}^3}{\text{m}^3}$

# G Übersicht der klassenbasierten PTFs aus der Literatur

Übersicht der Ergebnisse der Literaturrecherche für verschiedene Bodenparameter pro Texturklasse (Die fett gedruckten Werte entsprechen den bisherigen in REMO verwendeten Werten.)

Parameter	vol. Wärmekapazität des Bodens bei mittlerem Wassergehalt	vol. Wärmekapazität des trockenen Bodens	thermale Diffusivität bei mittlerem Wassergehalt	Wärmeleitfähigkeit von trockenem Boden	Amplitude der Wärmeleitfähigkeit	Porenvolumen	Feldkapazität	Quelle
	<i>RGCGN</i> ( <i>LWDIF=FALSE</i> )	<i>RGCGN</i> ( <i>LWDIF=TRUE</i> )	<i>TLAMBDA</i> ( <i>LWDIF=FALSE</i> )	<i>TLAMBDA</i> ( <i>LWDIF=TRUE</i> )	<i>DLAMBDA</i>	<i>PORVOL</i>	<i>FCAP</i>	
	1 $\frac{10^6 J}{m^3 K}$	2 $\frac{10^6 J}{m^3 K}$	3 $\frac{10^{-6} m^2}{s}$	4 $\frac{W}{m K}$	5 $\frac{W}{m K}$	6 $\frac{m^3}{m^3}$	7 $\frac{m^3}{m^3}$	
Cl	<b>2.48</b>	<b>1.63</b>	<b>0.67</b>	<b>0.18</b>	<b>1.50</b>	<b>0.507</b>	<b>0.463</b>	REMO = Majewski, 1991, Doms et al., 2011
		1.10						F. J. Braun (2002)
			1.60	0.50	1.20			Márquez et al. (2016)
	2.45	1.55		0.50	1.20			Markert (2012) nach VDI 4640
		1.25				0.461		Beringer et al. (2001)
						0.461		Hagemann & Stacke (2015)
							0.384	Hagemann & Stacke (2015) nach Patterson (1990)
	2.25	1.42	0.53	0.25	1.33			Monteith & Unsworth (2013) nach van Wijk & de Vries (1963)
						0.380		Carsel & Parrish (1988)
						0.468		F. Chen & Dudhia (2001) nach Cosby et al. (1984)
							0.412	F. Chen & Dudhia (2001) nach Hillel (1980)
		1.42		0.25	1.33			Arya (2008) nach Oke (1987) und Garratt (1992)
					0.482		Clapp & Hornberger (1978)	

Fortsetzung auf der nächsten Seite

Fortsetzung der vorherigen Seite

Klasse	1	2	3	4	5	6	7	Quelle
Cl						0.503	0.445	El-Nesr (2006) nach Keller & Karmeli (1975), van Genuchten (1980) und Saxton et al. (1986)
		1.09		0.26		0.380	0.247	Guimberteau et al. (2018)
		1.09						Pielke (1984), Pielke (2013)
SiCl		1.09						F. J. Braun (2002)
		1.22				0.481		Beringer et al. (2001)
						0.481		Hagemann & Stacke (2015)
							0.338	Hagemann & Stacke (2015) nach Patterson (1990)
						0.360		Carsel & Parrish (1988)
						0.468		F. Chen & Dudhia (2001) nach Cosby et al. (1984)
							0.404	F. Chen & Dudhia (2001) nach Hillel (1980)
						0.492		Clapp & Hornberger (1978)
						0.500	0.423	El-Nesr (2006) nach Keller & Karmeli (1975), van Genuchten (1980) und Saxton et al. (1986)
SaCl		1.25						F. J. Braun (2002)
		1.29				0.424		Beringer et al. (2001)
						0.424		Hagemann & Stacke (2015)
							0.311	Hagemann & Stacke (2015) nach Patterson (1990)
						0.380		Carsel & Parrish (1988)
						0.406		F. Chen & Dudhia (2001) nach Cosby et al. (1984)
							0.338	F. Chen & Dudhia (2001) nach Hillel (1980)
						0.426		Clapp & Hornberger (1978)
						0.465	0.328	El-Nesr (2006) nach Keller & Karmeli (1975), van Genuchten (1980) und Saxton et al. (1986)
ClLo	<b>2.36</b>	<b>1.50</b>	<b>0.71</b>	<b>0.21</b>	<b>1.55</b>	<b>0.475</b>	<b>0.370</b>	REMO = Majewski (1991), Doms et al. (2011)
		1.29						F. J. Braun (2002)
		1.25				0.449		Beringer et al. (2001)
ClLo						0.449		Hagemann & Stacke (2015)
							0.317	Hagemann & Stacke (2015) nach Patterson (1990)
						0.410		Carsel & Parrish (1988)
						0.465		F. Chen & Dudhia (2001) nach Cosby et al. (1984)

Fortsetzung auf der nächsten Seite

Fortsetzung der vorherigen Seite

Klasse	1	2	3	4	5	6	7	Quelle
<b>CI<sub>Lo</sub></b>							0.382	F. Chen & Dudhia (2001) nach Hillel (1980)
						0.476		Clapp & Hornberger (1978)
						0.479	0.330	El-Nesr (2006) nach Keller & Karmeli (1975), van Genuchten (1980) und Saxton et al. (1986)
		1.23		0.23		0.410	0.270	Guimberteau et al. (2018)
		1.23				0.476		Pielke (1984), Pielke (2013)
		1.18						Ochsner et al. (2001)
<b>SiCI<sub>Lo</sub></b>		1.16						F. J. Braun (2002)
		1.22				0.476		Beringer et al. (2001)
						0.476		Hagemann & Stacke (2015)
							0.345	Hagemann & Stacke (2015) nach Patterson (1990)
						0.430		Carsel & Parrish (1988)
						0.464		F. Chen & Dudhia (2001) nach Cosby et al. (1984)
							0.387	F. Chen & Dudhia (2001) nach Hillel (1980)
						0.477		Clapp & Hornberger (1978)
						0.503	0.384	El-Nesr (2006) nach Keller & Karmeli (1975), van Genuchten (1980) und Saxton et al. (1986)
		1.32		0.22		0.390	0.169	Guimberteau et al. (2018)
	1.32				0.477		Pielke (2013)	
	1.29						Ochsner et al. (2001)	
<b>SaCI<sub>Lo</sub></b>		1.38						F. J. Braun (2002)
		1.29				0.416		Beringer et al. (2001)
						0.416		Hagemann & Stacke (2015)
							0.280	Hagemann & Stacke (2015) nach Patterson (1990)
						0.390		Carsel & Parrish (1988)
						0.404		F. Chen & Dudhia (2001) nach Cosby et al. (1984)
							0.314	F. Chen & Dudhia (2001) nach Hillel (1980)
						0.420		Clapp & Hornberger (1978)
						0.450	0.268	El-Nesr (2006) nach Keller & Karmeli (1975), van Genuchten (1980) und Saxton et al. (1986)
		1.18		0.25		0.390	0.169	Guimberteau et al. (2018)
	1.18				0.420		Pielke (2013)	
<b>Lo</b>	<b>2.25</b>	<b>1.42</b>	<b>0.74</b>	<b>0.25</b>	<b>1.58</b>	<b>0.455</b>	<b>0.340</b>	REMO = Majewski (1991), Doms et al. (2011)
		1.24						F. J. Braun (2002)
		1.25				0.435		Beringer et al. (2001)
						0.435		Hagemann & Stacke (2015)

Fortsetzung auf der nächsten Seite

Fortsetzung der vorherigen Seite

Klasse	1	2	3	4	5	6	7	Quelle
<b>Lo</b>							0.288	Hagemann & Stacke (2015) nach Patterson (1990)
						0.430		Carsel & Parrish (1988)
						0.439		F. Chen & Dudhia (2001) nach Cosby et al. (1984)
							0.329	F. Chen & Dudhia (2001) nach Hillel (1980)
						0.451		Clapp & Hornberger (1978)
						0.443	0.266	El-Nesr (2006) nach Keller & Karmeli (1975), van Genuchten (1980) und Saxton et al. (1986)
		1.21		0.22		0.430	0.165	Guimberteau et al. (2018)
		1.21				0.451		Pielke (1984), Pielke (2013)
<b>SiLo</b>		1.15						F. J. Braun (2002)
		1.19				0.468		Beringer et al. (2001)
						0.468		Hagemann & Stacke (2015)
							0.341	Hagemann & Stacke (2015) nach Patterson (1990)
						0.450		Carsel & Parrish (1988)
						0.476		F. Chen & Dudhia (2001) nach Cosby et al. (1984)
							0.360	F. Chen & Dudhia (2001) nach Hillel (1980)
						0.485		Clapp & Hornberger (1978)
						0.453	0.308	El-Nesr (2006) nach Keller & Karmeli (1975), van Genuchten (1980) und Saxton et al. (1986)
		1.27		0.20		0.450	0.240	Guimberteau et al. (2018)
		1.27				0.485		Pielke (2013)
	1.17						Ochsner et al. (2001)	
<b>SaLo</b>		1.26						F. J. Braun (2002)
		1.27				0.416		Beringer et al. (2001)
						0.416		Hagemann & Stacke (2015)
							0.229	Hagemann & Stacke (2015) nach Patterson (1990)
						0.410		Carsel & Parrish (1988)
						0.434		F. Chen & Dudhia (2001) nach Cosby et al. (1984)
							0.312	F. Chen & Dudhia (2001) nach Hillel (1980)
						0.435		Clapp & Hornberger (1978)
						0.413	0.208	El-Nesr (2006) nach Keller & Karmeli (1975), van Genuchten (1980) und Saxton et al. (1986)
		1.34		0.23		0.410	0.122	Guimberteau et al. (2018)
		1.34				0.435		Pielke (1984), Pielke (2013)
	1.13						Ochsner et al. (2001)	

Fortsetzung auf der nächsten Seite



Fortsetzung der vorherigen Seite

Klasse	1	2	3	4	5	6	7	Quelle	
Si		1.15						F. J. Braun (2002)	
		1.60		0.50	1.20			Márquez et al. (2016)	
		2.45	1.55		0.50	1.20		Markert (2012) nach VDI 4640	
							0.460	Carsel & Parrish (1988)	
							0.476	F. Chen & Dudhia (2001) nach Cosby et al. (1984)	
								0.360	F. Chen & Dudhia (2001) nach Hillel (1980)
								0.485	Clapp & Hornberger (1978)
							0.429	0.306	El-Nesr (2006) nach Keller & Karmeli (1975), van Genuchten (1980) und Saxton et al. (1986)
			1.21		0.20		0.460	0.258	Guimberteau et al. (2018)
LoSa	<b>2.10</b>	<b>1.35</b>	<b>0.8</b>	<b>0.28</b>	<b>2.40</b>	<b>0.445</b>	<b>0.260</b>	REMO = Majewski (1991), Doms et al. (2011)	
		1.35						F. J. Braun (2002)	
		1.32					0.386	Beringer et al. (2001)	
							0.386	Hagemann & Stacke (2015)	
								0.172	Hagemann & Stacke (2015) nach Patterson (1990)
							0.430	Carsel & Parrish (1988)	
							0.421	F. Chen & Dudhia (2001) nach Cosby et al. (1984)	
								0.283	F. Chen & Dudhia (2001) nach Hillel (1980)
							0.410	Clapp & Hornberger (1978)	
							0.387	0.156	El-Nesr (2006) nach Keller & Karmeli (1975), van Genuchten (1980) und Saxton et al. (1986)
		1.41		0.23		0.410	0.0071	Guimberteau et al. (2018)	
		1.41				0.410		Pielke (2013)	
Sa	<b>1.93</b>	<b>1.28</b>	<b>0.87</b>	<b>0.30</b>	<b>2.40</b>	<b>0.364</b>	<b>0.196</b>	REMO = Majewski (1991), Doms et al. (2011)	
		1.35						F. J. Braun (2002)	
	2.75		0.73					Witt (2018) nach B. Braun et al. (1985)	
		1.60		0.40	2.00			Márquez et al. (2016)	
	2.30	1.45		0.40	2.00			Markert (2012) nach VDI 4640	
		1.34					0.373	Beringer et al. (2001)	
							0.373	Hagemann & Stacke (2015)	
								0.093	Hagemann & Stacke (2015) nach Patterson (1990)
	2.12	1.28	0.85	0.30	1.90			Monteith & Unsworth (2013) nach van Wijk & de Vries (1963)	

Fortsetzung auf der nächsten Seite

Fortsetzung der vorherigen Seite

Klasse	1	2	3	4	5	6	7	Quelle
<b>Sa</b>						0.430		Carsel & Parrish (1988)
						0.339		F. Chen & Dudhia (2001) nach Cosby et al. (1984)
							0.236	F. Chen & Dudhia (2001) nach Hillel (1980)
		1.28		0.30	1.90			Arya (2008) nach Oke (1987) und Garratt (1992)
						0.395		Clapp & Hornberger (1978)
						0.376	0.137	El-Nesr (2006) nach Keller & Karmeli (1975), van Genuchten (1980) und Saxton et al. (1986)
		1.47		0.22		0.430	0.0049	Guimberteau et al. (2018)
		1.47				0.395		Pielke (1984), Pielke (2013)
<b>Torf</b>	<b>2.59</b>	<b>0.58</b>	<b>0.65</b>	<b>0.06</b>	<b>0.50</b>	<b>0.863</b>	<b>0.763</b>	REMO = Majewski (1991), Doms et al. (2011)
		0.83						F. J. Braun (2002)
						0.700		Beringer et al. (2001)
						0.880		Hagemann & Stacke (2015) nach Letts et al. (2000)
	2.18	0.58	0.13	0.06	0.44			Monteith & Unsworth (2013) nach van Wijk & de Vries (1963)
						0.863		Carsel & Parrish (1988)
						0.439		F. Chen & Dudhia (2001) nach Cosby et al. (1984)
							0.329	F. Chen & Dudhia (2001) nach Hillel (1980)
		0.58		0.06	0.44			Arya (2008) nach Oke (1987) und Garratt (1992)
							0.880	El-Nesr (2006) nach Keller & Karmeli (1975), van Genuchten (1980) und Saxton et al. (1986)
		2.50		0.05		0.920		Guimberteau et al. (2018)
	0.84				0.863		Pielke (1984), Pielke (2013)	

# H Literaturübersicht der PTF für die Feldkapazität (FCAP)

Auswahl an Pedotransferfunktionen für die Feldkapazität (FCAP) aus der Literatur

Sa: Sandanteil in % / Cl: Tonanteil in % / Si: Schluffanteil in % /  $p_d$ : Trockenrohdichte in  $\frac{g}{cm^3}$  / SOC: Soil Organic Carbon / SOM: Soil Organic Matter

x: Variable wird für die Gleichung benötigt.

(x): Trockenrohdichte wird zusätzlich zu den in der originalen Formel verwendeten Variablen zur Umwandlung von gravimetrischer zur volumetrischer FCAP benötigt.

Einheit von FCAP:

vol: volumetrisch [ $\frac{cm^3}{cm^3}$  oder  $\frac{m^3}{m^3}$ ] / grav: gravimetrisch [ $\frac{g}{g}$  oder  $\frac{kg}{kg}$ ]

### Anmerkungen:

<sup>1</sup>: Formel wird in der Veröffentlichung in einer anderen Einheit angegeben bzw. umgerechnet.

<sup>2</sup>: Formel weicht leicht von der Originalquelle ab (Übertragungsfehler wie z. B. falsches Vorzeichen, Dezimalstellenverschiebung, andere Einheit von Sand- und Tonanteil oder andere organische Variable SOM/SOC).

<sup>3</sup>: Formel ist umformuliert.

ID	Einheit	Sa	Cl	Si	$p_d$	SOC/SOM	Formel	Referenzen
1	vol	x		x			$0.5637 - 0.0027 \cdot Si - 0.0051 \cdot Sa$	Adhikary et al. (2008) in Rustanto et al. (2017); Santra et al. (2018); Qiao et al. (2019)
2	grav	x			x		$0.6788 - 0.0055 \cdot Sa - 0.0013 \cdot Sa \cdot p_d$	Aina & Periaswamy (1985) in van den Berg et al. (1997) <sup>1,3</sup> ; Tomasella & M. Hodnett (2004) <sup>1,2</sup> ; Botula et al. (2012) <sup>1,2</sup> ; Medeiros et al. (2014) <sup>1,2</sup> ; Ostovari et al. (2015) <sup>1,2</sup> ; Rustanto et al. (2017) <sup>2</sup>
3	grav		x	x	(x)		$0.0307439 + 0.00629 \cdot (Cl + Si) - 0.0000343813 \cdot (Cl + Si)^2$	Arruda et al. (1987) in Oliveira et al. (2002); Reichert et al. (2009); Medeiros et al. (2014); Pineda et al. (2018)
4	grav		x	x	(x)		$0.0671198 \cdot \exp(4.28 \cdot 10^{-2}(Si + Cl) - 3.949 \cdot 10^{-4}(Si + Cl)^2 + 6.68 \cdot 10^{-7}(Si + Cl)^3)$	Arruda et al. (1987) in Tomasella & M. Hodnett (2004)

Fortsetzung auf der nächsten Seite

Fortsetzung der vorherigen Seite

ID	Einheit	Sa	Cl	Si	$p_d$	SOC/SOM	Formel	Referenzen
5	grav		x	x	(x)		$0.0993 + 0.0029 \cdot (Cl + Si)$	Arruda et al. (1987) in van den Berg et al. (1997); Botula et al. (2012); Rustanto et al. (2017) <sup>2</sup>
6	vol		x	x		x	$0.0046 \cdot Cl + 0.003045 \cdot Si + 0.020703 \cdot SOC[\%]$	Batjes (1996) in Medeiros et al. (2014)
7	vol	x					$0.4193 - 0.0035 \cdot Sa$	Botula (2013) in Rustanto et al. (2017) <sup>1,2</sup> ; Qiao et al. (2019)
8	vol		x				$\frac{0.043 + 0.004 \cdot Cl}{0.471 + 0.00411 \cdot Cl}$	Bruand et al. (1994) in Dai et al. (2013)
9	grav	x	x	x	(x)		$0.27447 + 0.00078 \cdot Cl + 0.00248 \cdot Si - 0.00241 \cdot Sa$	Chakraborty et al. (2011) in Santra et al. (2018)
10	vol	x			x		$0.156 - 0.00323 \cdot Sa + 0.169 \cdot p_d$	Dashtaki & Homae (2004) in Ostovari et al. (2015)
11	grav	x			(x)		$0.3697 - 0.0035 \cdot Sa$	Dijkerman (1988) in van den Berg et al. (1997) <sup>3</sup> ; Tomasella & M. Hodnett (2004); Medeiros et al. (2014); Ostovari et al. (2015) <sup>1,2</sup> ; Rustanto et al. (2017); Pineda et al. (2018); Qiao et al. (2019) <sup>1</sup>
12	grav	x			(x)		$0.334 - 0.003 \cdot Sa$	Lal (1979) in Pineda et al. (2018)
13	grav		x		(x)		$0.065 + 0.004 \cdot Cl$	Lal (1979) in Tomasella & M. Hodnett (2004) <sup>1</sup> ; Botula et al. (2012); Medeiros et al. (2014) <sup>1</sup> ; Ostovari et al. (2015) <sup>1</sup> ; Rustanto et al. (2017)
14	grav				(x)		$0.00333 \cdot Si + 0.00387 \cdot Cl$	Oliveira et al. (2002) in Tomasella & M. Hodnett (2004); Reichert et al. (2009); Botula et al. (2012); Medeiros et al. (2014); Ostovari et al. (2015) <sup>1</sup> ; Rustanto et al. (2017); Pineda et al. (2018) <sup>2</sup>
15	vol	x				x	$0.3815 - 0.0017 \cdot Sa + 0.0077 \cdot SOM[\%]$	Pidgeon (1972)

Fortsetzung auf der nächsten Seite

Fortsetzung der vorherigen Seite

ID	Einheit	Sa	Cl	Si	$p_d$	SOC/SOM	Formel	Referenzen
16	grav		x	x	(x)	x	$0.0738 + 0.0016 \cdot Si + 0.003 \cdot Cl + 0.015 \cdot SOM[\%]$	Pidgeon (1972) in van den Berg et al. (1997) <sup>2</sup> ; Tomasella & M. Hodnett (2004) <sup>2</sup> ; Botula et al. (2012) <sup>2</sup> ; Medeiros et al. (2014) <sup>3</sup> ; Rustanto et al. (2017) <sup>3</sup> ; Pineda et al. (2018) <sup>2,3</sup>
17	vol	x	x			x	$0.2576 + 0.0036 \cdot Cl - 0.0026 \cdot Sa + 0.0299 \cdot SOM[\%]$	Rawls & Brakensiek (1982) in Reichert et al. (2009); Botula et al. (2012); Dai et al. (2013) <sup>2</sup> ; Pineda et al. (2018)
18	vol		x	x			$0.04046 + 0.00404 \cdot Cl + 0.00426 \cdot Si$	Tomasella & M. G. Hodnett (1998) in Dai et al. (2013); Pineda et al. (2018); Santra et al. (2018)

# I Literaturübersicht der PTF für den permanenten Welkepunkt (PWP)

Auswahl an Pedotransferfunktionen für den Permanenten Welkepunkt (PWP) aus der Literatur  
 Sa: Sandanteil in % / Cl: Tonanteil in % / Si: Schluffanteil in % /  $p_d$ : Trockenrohdichte in  $\frac{g}{cm^3}$  /  
 SOC: Soil Organic Carbon / SOM: Soil Organic Matter

x: Variable wird für die Gleichung benötigt.

(x): Trockenrohdichte wird zusätzlich zu den in der originalen Formel verwendeten Variablen zur Umwandlung von gravimetrischer zur volumetrischer PWP benötigt.

Einheit von PWP:

vol: volumetrisch [ $\frac{cm^3}{cm^3}$  oder  $\frac{m^3}{m^3}$ ] / grav: gravimetrisch [ $\frac{g}{g}$  oder  $\frac{kg}{kg}$ ]

## Anmerkungen:

<sup>1</sup>: Formel wird in der Veröffentlichung in einer anderen Einheit angegeben bzw. umgerechnet.

<sup>2</sup>: Formel weicht leicht von der Originalquelle ab (Übertragungsfehler wie z. B. falsches Vorzeichen, Dezimalstellenverschiebung, andere Einheit von Sand- und Tonanteil oder andere organische Variable SOM/SOC).

<sup>3</sup>: Formel ist umformuliert.

ID	Einheit	Sa	Cl	Si	$p_d$	SOC/SOM	Formel	Referenzen
1	vol		x				$0.0071 + 0.0044 \cdot Cl$	Adhikary et al. (2008) in Rustanto et al. (2017); Santra et al. (2018); Qiao et al. (2019)
2	grav		x		(x)		$0.0213 + 0.0031 \cdot Cl$	Aina & Periaswamy (1985) in van den Berg et al. (1997) <sup>1,3</sup> ; Tomasella & M. Hodnett (2004) <sup>1,2</sup> ; Botula et al. (2012) <sup>1,2</sup> ; Medeiros et al. (2014) <sup>1,2</sup> ; Ostovari et al. (2015) <sup>1,2</sup> ; Rustanto et al. (2017)
3	grav				(x)		$\frac{398.889 \cdot (Si + Cl)}{1308.09 + (Si + Cl)} / 100$	Arruda et al. (1987) in Oliveira et al. (2002); Medeiros et al. (2014)
4	grav				(x)		$2.3662 \cdot 10^{-2} \cdot (Si + Cl)^{1.20408 - 0.0872025 \cdot \log(Si + Cl)}$	Arruda et al. (1987) in Tomasella & M. Hodnett (2004)

Fortsetzung auf der nächsten Seite

Fortsetzung der vorherigen Seite

ID	Einheit	Sa	Cl	Si	$p_d$	SOC/SOM	Formel	Referenzen
5	grav		x	x	(x)		$0.0107 + 0.0027 \cdot Cl + 0.0027 \cdot Si$	Arruda et al. (1987) in van den Berg et al. (1997); Reichert et al. (2009); Botula et al. (2012); Rustanto et al. (2017) <sup>2</sup> ;
6	vol						$0.003624 \cdot Cl + 0.00117 \cdot Si + 0.016054 \cdot SOC[\%]$	Batjes (1996) in Medeiros et al. (2014)
7	vol	x	x		x		$0.0841 - 0.00159 \cdot Sa + 0.0021 \cdot Cl + 0.0779 \cdot p_d$	Botula (2013) in Rustanto et al. (2017) <sup>1</sup> ; Qiao et al. (2019)
8	vol		x				$\frac{0.008 + 0.00367 \cdot Cl}{0.471 + 0.00411 \cdot Cl}$	Bruand et al. (1994) in Dai et al. (2013)
9	grav	x	x	x	(x)		$0.20695 + 0.00021 \cdot Cl - 0.00028 \cdot Si - 0.00179 \cdot Sa$	Chakraborty et al. (2011) in Santra et al. (2018)
11	grav		x		(x)		$0.0074 + 0.0039 \cdot Cl$	Dijkerman (1988) in van den Berg et al. (1997); Tomasella & M. Hodnett (2004); Medeiros et al. (2014); Ostovari et al. (2015) <sup>1</sup> ; Rustanto et al. (2017); Qiao et al. (2019) <sup>1</sup>
13	grav		x		(x)		$0.006 + 0.003 \cdot Cl$	Lal (1979) in Tomasella & M. Hodnett (2004) <sup>1</sup> ; Botula et al. (2012); Medeiros et al. (2014) <sup>1</sup> ; Ostovari et al. (2015) <sup>1</sup> ; Rustanto et al. (2017)
14	grav	x	x	x	x		$0.00038 \cdot Sa + 0.00153 \cdot Si + 0.00341 \cdot Cl - 0.030681 \cdot p_d$	Oliveira et al. (2002) in Tomasella & M. Hodnett (2004); Reichert et al. (2009) <sup>2</sup> ; Botula et al. (2012); Medeiros et al. (2014); Ostovari et al. (2015) <sup>1</sup> ; Rustanto et al. (2017)
16	grav		x	x	(x)	x	$-0.0419 + 0.0019 \cdot Si + 0.0039 \cdot Cl + 0.009 \cdot SOM[\%]$	Pidgeon (1972) in van den Berg et al. (1997) <sup>2</sup> ; Tomasella & M. Hodnett (2004) <sup>2</sup> ; Botula et al. (2012) <sup>2</sup> ; Medeiros et al. (2014) <sup>2</sup> ; Rustanto et al. (2017)
17	vol		x			x	$0.026 + 0.005 \cdot Cl + 0.00158 \cdot SOM[\%]$	Rawls & Brakensiek (1982) in Reichert et al. (2009) <sup>2</sup> ; Botula et al. (2012) <sup>2</sup> ; Dai et al. (2013) <sup>2</sup> ; Ostovari et al. (2015)
18	vol		x	x			$0.0091 + 0.00396 \cdot Cl + 0.0015 \cdot Si$	Tomasella & M. G. Hodnett (1998) in Dai et al. (2013); Santra et al. (2018)

# J Liste der Variablen mit signifikanten Unterschieden zwischen Modellläufen des Experimentplan 1 im GER-11

Dargestellt ist das Ergebnis des Hypothesentests pro Variable und Modelllauf des 1. Experimentplans. Als Datengrundlage dienen tägliche Anomaliedaten, die auf räumlicher Basis für jede der 115 Testvariablen erstellt wurden (genauere Informationen sind der Beschreibung von Tabelle 15 zu entnehmen). Die Bedeutung der Modelllaufkürzel sind Abbildung 44 zu entnehmen. ✓ steht für einen signifikanten Unterschied zwischen Modelllauf id01 und den anderen Modellläufen bei einer Irrtumswahrscheinlichkeit von  $\alpha = 5\%$ .

Variable	id01_GFT	id01_GFF	id01_GHT	id01_GHF	id01_GST	id01_GSF	id01_AFT	id01_AST	id01_ASF
surface u-stress (land)			✓	✓		✓	✓	✓	
surface u-stress (water)	✓	✓	✓	✓	✓	✓	✓	✓	✓
surface u-stress (ice)									
surface v-stress (land)	✓	✓	✓	✓	✓	✓	✓	✓	✓
surface v-stress (water)	✓	✓	✓	✓	✓	✓	✓	✓	✓
surface v-stress (ice)									
surface sensible heat flux (land) (= 69)	✓	✓	✓	✓	✓	✓		✓	✓
surface sensible heat flux (water) (= 70)	✓	✓	✓	✓	✓	✓		✓	✓
surface sensible heat flux (ice) (= 71)									
surface latent heat flux (land) (= 66)	✓			✓	✓	✓	✓	✓	✓
surface latent heat flux (water) (= 67)	✓	✓	✓	✓	✓	✓	✓	✓	✓
surface latent heat flux (ice) (= 68)									
sea-ice: conductive heat									
residual heat flux for melting sea ice									
boundary layer dissipation	✓	✓	✓	✓	✓	✓	✓	✓	✓
surface sensible heat flux	✓	✓	✓	✓	✓	✓		✓	✓
surface latent heat flux	✓			✓	✓	✓	✓	✓	✓
total cloud cover	✓	✓	✓	✓	✓	✓	✓	✓	✓
10m u-velocity	✓	✓	✓	✓	✓		✓	✓	✓
10m v-velocity	✓	✓	✓	✓	✓	✓	✓	✓	✓
10m u-velocity (reg)	✓	✓	✓	✓			✓	✓	✓
10m v-velocity (reg)	✓	✓	✓	✓	✓	✓	✓	✓	✓
2m temperature	✓	✓	✓	✓	✓	✓	✓	✓	✓
2m dew point temperature	✓	✓	✓	✓	✓	✓	✓	✓	✓
surface temperature (land) (see also 54)		✓	✓	✓	✓	✓	✓	✓	✓
10m windspeed	✓	✓	✓	✓	✓		✓	✓	✓
net surface solar radiation	✓	✓	✓	✓	✓	✓	✓	✓	✓
net surface thermal radiation		✓	✓	✓	✓	✓	✓	✓	✓
net top solar radiation	✓	✓	✓	✓		✓	✓	✓	✓

Fortsetzung auf der nächsten Seite



Fortsetzung der vorherigen Seite

Variable	id01_GFT	id01_GFF	id01_GHT	id01_GHF	id01_GST	id01_GSF	id01_AFT	id01_AST	id01_ASF
top thermal radiation (OLR)	✓	✓	✓	✓	✓	✓	✓	✓	
surface u-stress		✓	✓	✓		✓	✓		
surface v-stress	✓	✓	✓	✓	✓	✓	✓	✓	
net surf. thermal radiation (clear sky)	✓	✓	✓	✓	✓	✓	✓	✓	✓
net top thermal radiation (clear sky)	✓	✓	✓	✓	✓	✓	✓	✓	✓
surface thermal cloud forcing	✓	✓		✓	✓	✓	✓	✓	✓
top thermal cloud forcing	✓	✓		✓	✓	✓	✓	✓	✓
top solar radiation upward	✓	✓	✓	✓		✓	✓	✓	✓
surface solar radiation upward	✓	✓	✓	✓		✓	✓	✓	✓
surface thermal radiation upward	✓	✓	✓	✓	✓	✓		✓	✓
(effective) sea-ice skin temperature	✓	✓	✓	✓	✓	✓		✓	✓
land: residual surface heat budget / sea-ice: conductive heat flux	✓	✓	✓	✓	✓	✓	✓	✓	✓
vertically integrated specific humidity	✓	✓	✓	✓	✓	✓	✓		
vertically integrated liquid water cont.	✓	✓	✓		✓	✓	✓	✓	✓
convective available potential energy (CAPE)	✓	✓	✓		✓		✓		✓
vertically integrated cloud ice	✓	✓	✓	✓		✓	✓	✓	✓
surface latent heat flux (=147)	✓			✓	✓	✓	✓	✓	✓
minimum 2m-temperature	✓	✓	✓	✓	✓	✓		✓	✓
minimum surface temperature	✓	✓	✓	✓	✓	✓		✓	✓
maximum 2m-temperature	✓	✓	✓	✓	✓	✓	✓	✓	✓
maximum surface temperature		✓	✓	✓	✓	✓		✓	✓
maximum 10m-wind speed (w/o gusts)		✓	✓	✓	✓	✓	✓	✓	✓
maximum height of convective cloud tops	✓	✓		✓	✓	✓	✓	✓	
maximum of the expected gust velocity near the surface	✓	✓	✓		✓	✓	✓		
surface temperature (land)		✓	✓	✓	✓	✓	✓	✓	✓
surface temperature (water)									
surface temperature (ice)									
surface roughness length (land)							✓	✓	✓
surface roughness length (water)	✓	✓	✓		✓	✓	✓	✓	✓
surface roughness length (ice)							✓	✓	✓
surface albedo (land)	✓	✓		✓	✓	✓	✓	✓	✓
surface albedo (water)									
surface albedo (ice)									
specific humidity surface (land)	✓	✓	✓	✓	✓	✓	✓	✓	✓
specific humidity surface (water)	✓	✓	✓	✓		✓	✓	✓	✓
specific humidity surface (ice)									
heat conductivity of dry soil	✓	✓	✓	✓	✓	✓		✓	✓
parameter for increasing the heat conductivity of the soil due to soil moisture	✓	✓	✓	✓	✓	✓		✓	✓
pore volume	✓	✓	✓	✓	✓	✓		✓	✓
specific humidity surface	✓	✓	✓	✓	✓	✓	✓	✓	✓
surface pressure	✓	✓	✓	✓	✓	✓	✓	✓	✓
surface temperature (mean over gridbox)	✓	✓	✓	✓	✓	✓		✓	✓
mean sea level pressure	✓	✓	✓	✓	✓	✓	✓	✓	✓
total cloud cover (see also 164)	✓	✓	✓	✓	✓	✓	✓	✓	✓

Fortsetzung auf der nächsten Seite

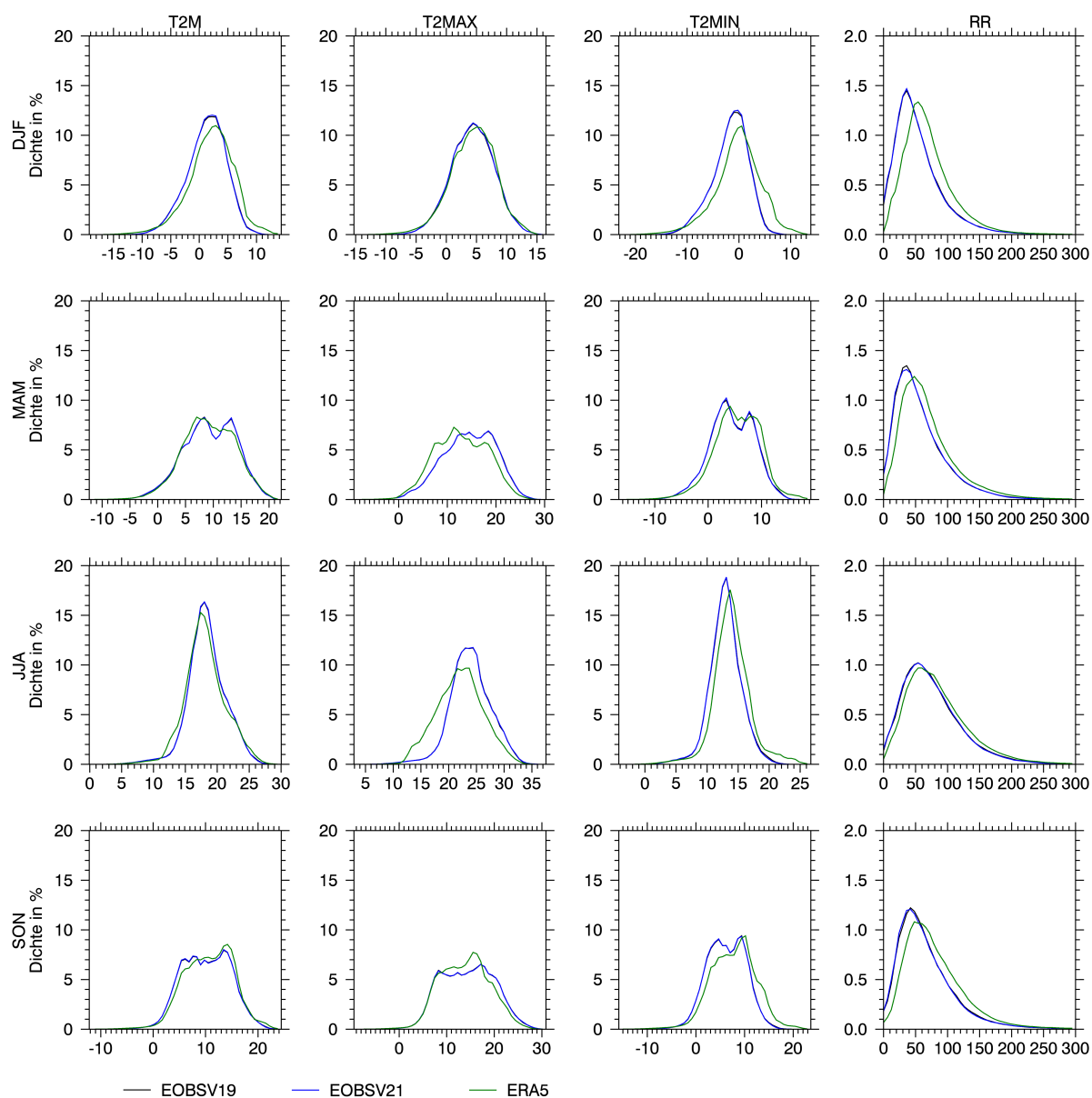
*J Liste der Variablen mit signifikanten Unterschieden zwischen Modellläufen des Experimentplan 1 im GER-11*

Fortsetzung der vorherigen Seite

Variable	id01_GFT	id01_GFF	id01_GHT	id01_GHF	id01_GST	id01_GSF	id01_AFT	id01_AST	id01_ASF
surface roughness length	✓	✓	✓	✓		✓	✓	✓	✓
surface background albedo									
surface albedo	✓	✓		✓	✓	✓	✓	✓	✓
vegetation ratio									
leaf area index									
snow temperature		✓	✓	✓	✓	✓	✓	✓	✓
sea ice cover									
sea ice depth									
height of the planetary boundary layer	✓	✓	✓		✓	✓		✓	
skin reservoir content	✓	✓	✓	✓	✓	✓	✓	✓	✓
soil wetness		✓		✓	✓	✓	✓	✓	✓
drainage (part of 160)	✓	✓	✓	✓	✓	✓	✓	✓	✓
total runoff (surface runoff + drainage)			✓				✓		✓
surface evaporation (land)	✓			✓	✓	✓	✓	✓	✓
surface evaporation (water)	✓	✓	✓	✓	✓	✓	✓	✓	✓
surface evaporation (ice)									
total precipitation	✓		✓	✓	✓	✓	✓		✓
large scale precipitation	✓	✓			✓	✓	✓	✓	
convective precipitation			✓	✓	✓	✓			✓
snow fall	✓	✓	✓	✓	✓	✓	✓		
snow melt	✓	✓	✓	✓	✓	✓	✓	✓	✓
snow depth		✓	✓	✓	✓	✓	✓	✓	✓
snow depth change	✓	✓	✓	✓	✓	✓	✓	✓	✓
transpiration at the surface	✓		✓	✓	✓	✓	✓	✓	✓
surface evaporation	✓			✓	✓	✓	✓	✓	✓
bare soil evaporation at the surface			✓	✓	✓	✓		✓	✓
evaporation over snow at the surface	✓	✓	✓	✓	✓	✓	✓	✓	✓
evaporation from skin reservoir at the surface	✓		✓			✓		✓	✓
ERES									
ustar	✓	✓	✓		✓		✓	✓	✓
surface sensible heat flux (=146)	✓	✓	✓	✓	✓	✓		✓	✓
fraction of frozen soil (layer 1)	✓	✓	✓	✓	✓	✓		✓	✓
fraction of frozen soil (layer 2)	✓	✓	✓		✓	✓		✓	✓
fraction of frozen soil (layer 3)	✓	✓	✓	✓	✓	✓	✓	✓	✓
fraction of frozen soil (layer 4)	✓	✓	✓	✓	✓	✓	✓	✓	✓
fraction of frozen soil (layer 5)	✓		✓	✓	✓	✓	✓	✓	✓
soil temperature (layer 1)	✓	✓	✓	✓	✓	✓	✓	✓	✓
soil temperature (layer 2)	✓	✓	✓	✓	✓	✓	✓	✓	✓
soil temperature (layer 3)	✓	✓	✓	✓	✓	✓		✓	✓
soil temperature (layer 4)	✓	✓	✓	✓	✓	✓		✓	✓
soil temperature (layer 5)	✓	✓	✓	✓	✓	✓		✓	✓
Gesamtanzahl	80	81	82	85	82	88	74	89	86

# K Kernfunktionsdichteschätzung der klimatologischen Validationsdatensätze pro Jahreszeit für GER-11

Die Datengrundlage bilden die Monatsmittel (T2M, T2MAX und T2MIN) bzw. -summen (RR) aller Landpunkte des GER-11 Gebiets von 2000-2018. Die jeweils an die Jahreszeit und Variable angepassten Werte der x-Achse sind für die Temperaturen (1.-3. Spalte) in °C und für den Niederschlag (4. Spalte) in mm angegeben (analog zu Abb. 73).



# L Vergleich von Modellläufen mit EOBSV21

Die Werte entsprechen dem jeweiligen Median von RMSE und r der Modellläufe des Experimentplan 1 (Kaltstart des Jahres 2000) im Vergleich zu EOBSV21 und sind Teil von Abbildung 74. Die Einfärbung orientiert sich an den Werten pro jeweiliger Variable, Jahreszeit und Qualitätsmerkmal (Rang 1 = dunkelgrün, Rang 10 = dunkelrot).

		T2M		T2MAX		T2MIN		RR	
		RMSE	r	RMSE	r	RMSE	r	RMSE	r
ann	id01	1.3935	0.9252	5.1129	0.8860	4.2846	0.7761	4.2143	0.4921
	id01_AFT	1.3918	0.9250	5.1062	0.8858	4.2917	0.7762	4.2336	0.4933
	id01_AST	1.3944	0.9260	5.1398	0.8856	4.2487	0.7782	4.2403	0.4914
	id01_ASF	1.3943	0.9257	5.1359	0.8854	4.2594	0.7777	4.2526	0.4892
	id01_GFT	1.3983	0.9249	5.1232	0.8855	4.2656	0.7777	4.2430	0.4928
	id01_GFF	1.3874	0.9252	5.0996	0.8858	4.2933	0.7758	4.2321	0.4925
	id01_GHT	1.3923	0.9245	5.1017	0.8857	4.2899	0.7761	4.2260	0.4910
	id01_GHF	1.3968	0.9253	5.1301	0.8855	4.2554	0.7778	4.2272	0.4927
	id01_GST	1.3954	0.9261	5.1437	0.8853	4.2473	0.7786	4.2315	0.4936
id01_GSF	1.3953	0.9258	5.1409	0.8850	4.2536	0.7780	4.2085	0.4911	
DJF	id01	1.7347	0.9207	4.1504	0.8740	2.8355	0.8688	2.6036	0.6117
	id01_AFT	1.7335	0.9207	4.1479	0.8739	2.8408	0.8682	2.6081	0.6106
	id01_AST	1.7391	0.9226	4.1827	0.8740	2.7999	0.8699	2.6068	0.6109
	id01_ASF	1.7377	0.9223	4.1740	0.8741	2.8119	0.8700	2.6083	0.6112
	id01_GFT	1.7535	0.9208	4.1743	0.8737	2.8155	0.8692	2.6044	0.6117
	id01_GFF	1.7316	0.9214	4.1428	0.8738	2.8379	0.8685	2.6061	0.6120
	id01_GHT	1.7462	0.9203	4.1515	0.8740	2.8369	0.8684	2.6032	0.6126
	id01_GHF	1.7346	0.9218	4.1683	0.8746	2.8138	0.8692	2.6085	0.6116
	id01_GST	1.7358	0.9228	4.1797	0.8744	2.7991	0.8703	2.6032	0.6132
id01_GSF	1.7408	0.9220	4.1759	0.8742	2.8075	0.8702	2.6046	0.6115	
MAM	id01	1.2090	0.9358	5.8879	0.9148	5.0213	0.6965	5.3743	0.4264
	id01_AFT	1.2073	0.9361	5.8841	0.9155	5.0308	0.6967	5.3714	0.4279
	id01_AST	1.2196	0.9362	5.9278	0.9145	4.9770	0.6978	5.4079	0.4284
	id01_ASF	1.2197	0.9357	5.9179	0.9141	4.9912	0.6980	5.3981	0.4233
	id01_GFT	1.2184	0.9358	5.9196	0.9138	4.9916	0.6967	5.3422	0.4306
	id01_GFF	1.2065	0.9357	5.8800	0.9151	5.0397	0.6960	5.3505	0.4286
	id01_GHT	1.2165	0.9354	5.8845	0.9146	5.0218	0.6963	5.3820	0.4258
	id01_GHF	1.2199	0.9356	5.9199	0.9141	4.9811	0.6973	5.3900	0.4259
	id01_GST	1.2226	0.9355	5.9308	0.9140	4.9719	0.6968	5.3796	0.4299
id01_GSF	1.2193	0.9360	5.9292	0.9139	4.9803	0.6965	5.3737	0.4272	
JJA	id01	1.2061	0.9464	5.4308	0.9097	5.0125	0.7751	3.9000	0.5299
	id01_AFT	1.2062	0.9458	5.4312	0.9092	5.0092	0.7758	3.9079	0.5307
	id01_AST	1.2112	0.9463	5.4568	0.9094	4.9742	0.7766	3.9439	0.5245
	id01_ASF	1.2080	0.9465	5.4518	0.9096	4.9800	0.7758	3.9279	0.5273
	id01_GFT	1.2098	0.9466	5.4442	0.9098	4.9809	0.7753	3.9323	0.5269
	id01_GFF	1.2061	0.9460	5.4253	0.9103	5.0144	0.7752	3.9138	0.5298
	id01_GHT	1.2044	0.9460	5.4193	0.9097	5.0186	0.7753	3.9288	0.5257
	id01_GHF	1.2149	0.9462	5.4544	0.9092	4.9845	0.7760	3.9034	0.5282
	id01_GST	1.2087	0.9469	5.4623	0.9090	4.9692	0.7776	3.9230	0.5304
id01_GSF	1.2106	0.9467	5.4569	0.9086	4.9780	0.7770	3.9130	0.5271	
SON	id01	1.2229	0.8809	4.3700	0.8254	3.6570	0.7127	3.9779	0.5731
	id01_AFT	1.2252	0.8805	4.3669	0.8254	3.6590	0.7123	4.0048	0.5736
	id01_AST	1.2334	0.8817	4.4041	0.8250	3.6109	0.7154	4.0192	0.5721
	id01_ASF	1.2317	0.8811	4.3943	0.8244	3.6263	0.7142	3.9948	0.5703
	id01_GFT	1.2368	0.8806	4.3813	0.8259	3.6418	0.7129	3.9611	0.5708
	id01_GFF	1.2245	0.8802	4.3532	0.8254	3.6678	0.7119	3.9928	0.5704
	id01_GHT	1.2238	0.8802	4.3627	0.8259	3.6712	0.7120	3.9853	0.5711
	id01_GHF	1.2318	0.8809	4.3866	0.8247	3.6250	0.7143	3.9864	0.5693
	id01_GST	1.2350	0.8818	4.4048	0.8241	3.6099	0.7149	3.9766	0.5747
id01_GSF	1.2323	0.8810	4.3959	0.8244	3.6238	0.7140	3.9865	0.5723	

# Glossar

- a-files** aufbereitete Antriebsdaten für einen REMO-Modelllauf
- boundary conditions** konstante Oberflächeneigenschaften, die im Modell definiert werden müssen
- DJF** Dezember-Januar-Februar (Winter)
- Downscaling** durch unterschiedliche Methoden erreichte Erhöhung der Modellauflösung
- e-files** Ergebnisdateien für jede Ausgabevariabel eines REMO-Modelllauf
- Endian** Byte-Reihenfolge, in der Einträge gelesen werden
- Equilibrium** Gleichgewichtszustand verschiedener Modellphysiken
- f/g-files** REMO-Dateien, die alle Variablen zum Zeitpunkt  $t-1$  und  $t-2$  enthalten, um einen Restart zum Zeitpunkt  $t$  durchführen zu können
- g-files** globale Antriebsdaten aus Reanalyseedaten oder Globalmodellen die für einen REMO-Modelllauf benötigt werden
- Hintcasting** Simulation des Klimas in der Vergangenheit
- in situ Messungen** Messungen unmittelbar an einer bestimmten Position
- JJA** Juni-Juli-August (Sommer)
- Kaltstart** Lauf eines Modells ohne sich im Gleichgewicht (Equilibrium) befindlichen Variablen
- Kinder** alle Unterprogramme, die von einem Programm aufgerufen werden
- LSPII** von Hagemann (2002) aus LSPI weiterentwickelte Tabelle, die Landnutzungsklassen die jeweiligen Landoberflächenparameter zuweist
- m-files** Ergebnisdateien der Monatsmittel eines REMO-Modelllauf für bestimmte Variablen
- MAM** März-April-Mai (Frühling)
- n-files** Ergebnisdateien der Tagesmittel eines REMO-Modelllauf für bestimmte Variablen
- Namelist** Konfigurations-Datei für REMO, in welcher die Modelleinstellungen und Schalter festgelegt sind
- Nesting** Einbettung eines höher aufgelösten Modelllaufs in einen niedriger aufgelösten
- Parallelisieren** wenn Rechenoperationen parallel auf mehreren Haupt- oder Grafikprozessoren durchgeführt werden, sodass die Rechenzeit abhängig von der Anzahl der Prozessoren verkürzt werden kann
- Peaks** markante Ausschläge einer Kurve
- Relaxationszone** Randbereich vom Modelluntersuchungsgebiet, der von den gröber aufgelösten Antriebsdaten beeinflusst wird
- REMOLAND** Projektname
- s-files** Ergebnisdateien der Standardabweichung eines REMO-Modelllauf für bestimmte Variablen
- SON** September-Oktober-November (Herbst)
- Spin-Up** Zeitraum, die ein Modell zum Erreichen des Equilibriums braucht
- t-files** Ergebnisdateien eines REMO-Modelllauf, die zum erneuten Nesten verwendet werden können
- Thermoisoplethen** kombinierte Darstellung der Temperatur in Abhängigkeit von zwei Variablen
- UTC** koordinierte Weltzeit (engl. Coordinated Universal Time, franz. Temps universel coordonné)
- Warmstart** Lauf eines Modells mit sich im Gleichgewicht (Equilibrium) befindlichen Variablen

**xe-Daten** entspricht den e-, n-, m- und s-files eines REMO-Modelllaufs

**xf-Daten** entspricht den f/g-files eines REMO-Modelllaufs

**xt-Daten** entspricht den t-files eines REMO-Modelllaufs