RESEARCH ARTICLE

Tracking the temporal dynamics of insect defoliation by high-resolution radar satellite data

Soyeon Bae[1] | Jörg Müller[1,2] | Bernhard Förster[3] | Torben Hilmers[4] |
Sophia Hochrein[1] | Martin Jacobs[4] | Benjamin M. L. Leroy[5] | Hans Pretzsch[4] |
Wolfgang W. Weisser[5] | Oliver Mitesser[1]

[1]Department of Animal Ecology and Tropical Biology, Biocenter, University of Würzburg, Würzburg, Germany

[2]Bavarian Forest National Park, Grafenau, Germany

[3]Chair for Strategic Landscape Planning and Management, Technical University of Munich, Freising, Germany

[4]Chair for Forest Growth and Yield Science, School of Life Sciences Weihenstephan, Technical University of Munich, Freising, Germany

[5]Terrestrial Ecology Research Group, Department of Ecology and Ecosystem Management, Technical University of Munich, Freising, Germany

Correspondence
Soyeon Bae
Email: soyeon.grace.bae@gmail.com

Funding information
Bayerisches Staatsministerium für Ernährung, Landwirtschaft und Forsten, Grant/Award Number: ST357 and Z073; National Research Foundation of Korea, Grant/Award Number: 2020R1A6A3A03038391

Handling Editor: Hooman Latifi

Abstract

1. Quantifying tree defoliation by insects over large areas is a major challenge in forest management, but it is essential in ecosystem assessments of disturbance and resistance against herbivory. However, the trajectory from leaf-flush to insect defoliation to refoliation in broadleaf trees is highly variable. Its tracking requires high temporal- and spatial-resolution data, particularly in fragmented forests.

2. In a unique replicated field experiment manipulating gypsy moth *Lymantria dispar* densities in mixed-oak forests, we examined the utility of publicly accessible satellite-borne radar (Sentinel-1) to track the fine-scale temporal trajectory of defoliation. The ratio of backscatter intensity between two polarizations from radar data of the growing season constituted a canopy development index (CDI) and a normalized CDI (NCDI), which were validated by optical (Sentinel-2) and terrestrial laser scanning (TLS) data as well by intensive caterpillar sampling from canopy fogging.

3. The CDI and NCDI strongly correlated with optical and TLS data (Spearman's ρ = 0.79 and 0.84, respectively). The $\Delta NCDI_{Defoliation(A-C)}$ significantly explained caterpillar abundance (R^2 = 0.52). The NCDI at critical timesteps and $\Delta NCDI$ related to defoliation and refoliation well discriminated between heavily and lightly defoliated forests.

4. We demonstrate that the high spatial and temporal resolution and the cloud independence of Sentinel-1 radar potentially enable spatially unrestricted measurements of the highly dynamic canopy herbivory. This can help monitor insect pests, improve the prediction of outbreaks and facilitate the monitoring of forest disturbance, one of the high priority Essential Biodiversity Variables, in the near future.

KEYWORDS
canopy herbivory, defoliation severity, gypsy moth, insect disturbance, intra-annual time-series, *Lymantria dispar*, remote sensing, Sentinel-1

1 | INTRODUCTION

The eruptive population dynamics of insects have challenged humans throughout their history, with an outbreak of desert locust among the punishments meted out to the Egyptians in the Old Testament [Exodus 10:13-15(2011)]. In forest ecology, the complex dynamics of eruptive folivorous moths and tree-killing bark beetles are determined by host tree abundance, symbionts, natural enemies, competitors and climate (Alalouni et al., 2013; Liebhold et al., 1998; Yguel et al., 2011). Climate change has already been shown to further increase the likelihood of outbreaks of insect pests (Millar & Stephenson, 2015; Seidl et al., 2014). Quantitative data on the impact of insect outbreaks on trees over a large scale provide a baseline for forest management policies and guide the distribution of government-sponsored financial support for forest owners. Furthermore, ecosystem disturbance is recognized as a high-priority essential biodiversity variable (EBV) class, as it is a driver of substantial change in ecosystems (GEO BON, 2017; Pettorelli et al., 2016).

Outbreaks of tree-killing insects over large areas have been mapped using remote sensing methods (Hollaus & Vreugdenhil, 2019; Senf et al., 2015). Even the early stage of tree die-back induced by bark beetle infestation can be identified by optical satellite data (Abdullah et al., 2019). A more complex challenge is detecting defoliation by insects in deciduous trees, as many tree species subsequently refoliate. In spring, deciduous trees develop their foliage within a few weeks, with herbivorous insects feeding on the highly nutritious new leaves shortly thereafter. Because the timing of leaf-flush and herbivore attack varies among trees and sites, the timing of defoliation and refoliation events varies as well (Lovett & Tobiessen, 1993; Senf et al., 2017; Townsend et al., 2012), requiring sensors with a sufficient temporal and spatial resolution.

Among defoliating insects, the gypsy moth *Lymantria dispar* (*Lepidoptera*: *Erebidae*), native to Europe and Asia, is a serious pest in deciduous forests (Lowe et al., 2000). During the last decades, expansion of the species' range and outbreaks of increasing frequency have been reported over the world (Giese & Schneider, 1979; Liebhold et al., 1992; Orozumbekov et al., 2009), supported by rising temperatures (Régnière, 2009; Seidl et al., 2014). The ecological and economic concerns associated with defoliation by gypsy moths have led to the deployment of remote sensing in monitoring efforts. Vegetation indices from spaceborne multispectral sensors, such as MODIS and Landsat, have been used in the near-real-time monitoring of gypsy moth outbreaks in the United States (de Beurs & Townsend, 2008; Spruce et al., 2011; Townsend et al., 2012). More recently, Meng et al. (2018) used airborne laser scanning to improve the detection of defoliation severity by gypsy moths at a tree scale. In that study and in others, the canopy status between two time slots of pre-defoliation and defoliation was compared (de Beurs & Townsend, 2008; Meng et al., 2018; Townsend et al., 2012). However, such analyses are too coarse to detect insect defoliation when the trees are defoliated for only a few weeks. In addition, spectral image acquisition is regularly impeded by cloud cover so that the optimal time for detecting defoliation can easily be missed. Thus,

studies of gypsy moth defoliation require very dense intra-annual time-series data (Mitchell et al., 2017; Senf et al., 2017). Intra-annual remote sensing studies of gypsy moth infestations are lacking, and little information is available for other deciduous defoliators, even with coarse resolution only appropriate to homogeneous landscape (Chávez et al., 2019; Olsson et al., 2016).

The global coverage provided by the Sentinel-1 mission, recently launched C-band synthetic aperture radar (SAR) satellites, is facilitating ecological studies by generating remote sensing data with high spatial and temporal resolution. The open-access policy for these data opens up new possibilities for forest ecology, such as the monitoring of forest biomass (Huang et al., 2018), deforestation (Reiche et al., 2018), phenology (Frison et al., 2018), burnt area (Carreiras et al., 2020) and windthrow (Tomppo et al., 2021). As such, recent studies showed the potential of cloud-independent Sentinel-1 radar data that might also facilitate monitoring of the temporal dynamics of defoliation by gypsy moths and subsequent refoliation, by providing dense intra-annual time-series data during the tree-growing season of temperate forests. Such dense time-series analysis of SAR data for mapping insect disturbance was not studied. It is because detecting attack by tree-killing insects may less require such intra-annual change detection (Ortiz et al., 2013; Ranson et al., 2003; Tanase et al., 2018), and such dense SAR data had not been available even for studying disturbance by defoliating insects, as shown in monthly analysis of ERS-2 SAR data in Kaasalainen et al. (2010).

In this study, we made use of a well-replicated experimental manipulation of caterpillar densities in mixed oak forests (see more details of the experimental design in Leroy et al., 2021) to test the applicability of publicly available Sentinel-1 radar (henceforth, simply referred to as 'radar') data to quantify the temporally and spatially highly variable herbivory and tree response by refoliation (Figure 1). To validate our approach, we used terrestrial laser scanning (TLS) data and caterpillar counts by canopy fogging and compared radar-derived indices with Sentinel-2 optical data.

2 | MATERIALS AND METHODS

2.1 | Study site

The study was conducted in a hilly area (200–500 m a.s.l.) in central Germany (centred at N 49°37′, E 10°24′). The plots were selected based on a gypsy moth survey of egg mass conducted in 2018 by regional forest offices on the 2,802 transects (10 trees per transect). The 'defoliation risk index' (DRI) was calculated with the mean egg mass density on transects divided by the critical egg-mass density (high risk if DRI > 1 and low risk if DRI < 0.5; for details, see file S1 in Leroy et al., 2021). Candidate plots required conditions of non-recent spray history, structural homogeneity and at least 5 ha. The combination of two high-risk (H) and two low-risk (L) candidate plots, composing one spatial block, located within 10 km was selected as candidate blocks. The final blocks satisfying approval of landowner and state were selected, and the detail of the plot selection can be

FIGURE 1 Schematic temporal trajectory of the normalized canopy development index (NCDI) at a heavily defoliated forest (DEF+). The time intervals below the *x*-axis indicate the range of the days of the year (DOY) and constitute the average values of the NCDI at the corresponding timestep. Grey areas indicate the different stages during seasonal development
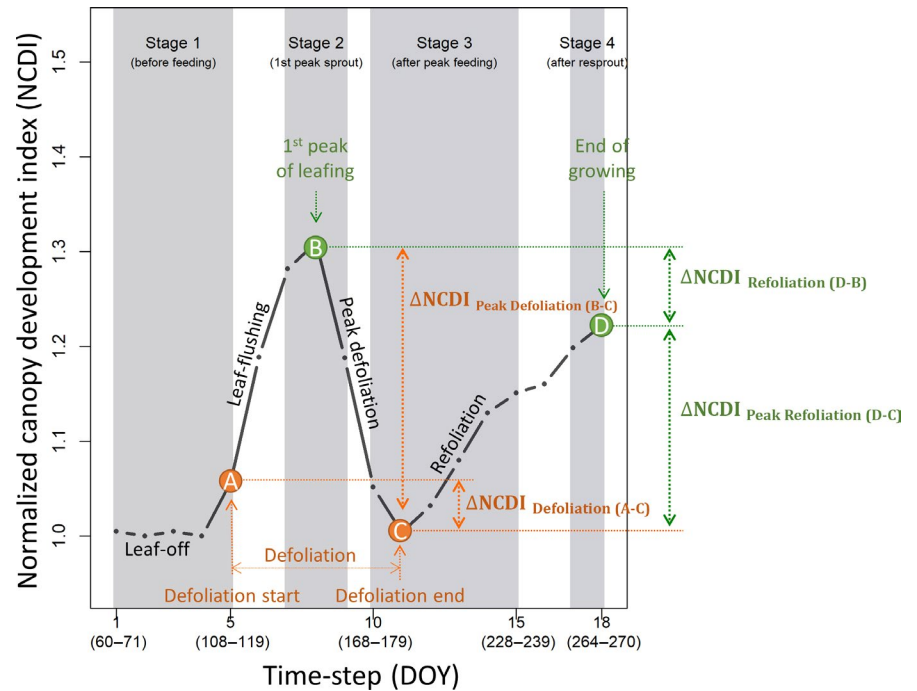
found in Leroy et al. (2021). The fully factorial experimental block design comprises two risk levels of defoliation by gypsy moths, H and L, and two pest-control treatments, sprayed with Mimic (active ingredient tebufenozide, M) and an unsprayed control (C). The Mimic plots were aerially sprayed with tebufenozide (Mimic®, Spiess-Urania Chemicals) by a helicopter between May 3rd and 23rd 2019. In each block, one of the two plots per risk level was randomly chosen for insecticide treatment; the other plot was left unsprayed and served as the control (HC and LC). The 2 × 2 factorial design thus results in four treatment combinations (HC, HM, LC and LM) that are distributed over 11 spatial blocks (A, B, D, F, G, H, J, M, N, O and S) for a total of 44 forest plots dominated by deciduous oak *Quercus robur* and *Quercus petraea* spread over about 70 km (Figure S1). Each block contained four treatment combinations, with similar forest structures, stand ages and tree species composition within a block. The mean plot area investigated was 4.6 ha (*SD*: ±2.1), which was the core area we collected the data of remote sensing and ground truth, corresponding to the subplot named in Leroy et al. (2021). To expand the dataset of highly defoliated areas for our binary a posterior classification analysis, we added another 12 plots (H2C type; high-risk and unsprayed plots) in the same study area observed in the field as highly defoliated but without TLS and caterpillar sampling. These five types constituted the a priori classification (before the onset of herbivory). Based on field observations during the 2019 growing season, the (44 + 12=) 56 plots were reclassified (a posteriori, after the onset of herbivory) according to their defoliation severity as either heavily (DEF+) or lightly (DEF−) defoliated forests. The Mimic plots (M) still had a few active caterpillars after spraying (see the caterpillar abundance from the four treatments in Figure S2), and the forest canopy of our study area was defoliated by other caterpillars as well as caterpillars of gypsy moths. Thus, we defined the Mimic plots (HM and LM) as well as LC plots as lightly defoliated rather than

no-defoliation plots, as they were defoliated at least lightly. Other than expected based on egg mass counts, the defoliation severity in some HC plots was light due to other ecological factors that limited defoliation. These three HC plots (AHC, GHC and SHC) were thus re-assigned to the DEF− group by field observation after the end of defoliation.

2.2 | Sentinel-1 data processing

2.2.1 | Sentinel-1 data acquisition and pre-processing

Sentinel-1 C-band SAR data covering our study area were obtained from the ESA Scientific Hub (https://scihub.copernicus.eu/) between March and September 2019. All available level-1 ground-range-detected high-resolution (GRDH) products acquired by the interferometric wide-swath mode were selected, including from both ascending (relative orbits of 117 and 15) and descending (relative orbits of 168 and 66) satellite passes. The GRDH products have a pixel spacing of 10 m and two polarizations, VV (vertically transmitted, vertically received radar pulse) and VH (vertically transmitted, horizontally received radar pulse). In total, 174 products over 132 days were retrieved. On average, 102 days of total 214 days between March and September were covered by GRDH products per plot. These were pre-processed using the Sentinel Application Platforms (SNAP) Sentinel-1 Toolbox software. The pre-processing steps included precise orbit correction, thermal noise removal and radiometric calibration to the radar brightness coefficient β^0. Radiometric terrain flattening and range-Doppler terrain correction were then executed based on the digital elevation model of the Shuttle Radar Topography Mission (SRTM v.4.1), with a 1-arc-second resolution (approximately 30 m), and β^0-transformed to γ^0, which

represents the backscatter coefficient normalized by the local illuminated area (see the Supplementary Note S 3.2 in Bae et al. (2019) and the batch processing graph at https://github.com/So-YeonBae/Sentinel1-Biodiversity). Lastly, the values were converted to dB as $10 \times \log_{10}\gamma^0$ (henceforth, $10 \times \log_{10}\gamma^0$ is referred to as γ^0). The mean γ^0 values of the VV and VH polarizations within each plot were extracted to characterize forest canopy status, using the function *extract* in the R package RASTER (Hijmans, 2019).

2.2.2 | Canopy development index

The time-series γ^0 values of the VV and VH polarizations were smoothed using the Gaussian window function within the function *smth.gaussian* in the R package SMOOTHER (Hamilton, 2015; see Supplementary Note 1 for parameterizing). Two smoothed products were generated, per plot, from all relative orbits by polarization, a 12-day composite and a single-day smoothed time series. For the former, the averages of the 12-day intervals (e.g. the average of DOY 60–71 for the first step) were calculated, and the Gaussian smooth function was then applied to the averaged values generating 18 timesteps between March and September (i.e. DOY 60–273). For the latter, the single-day smoothed and interpolated time series was generated.

The difference between the smoothed γ^0 values of VV and VH, defined herein as the canopy development index (CDI = $\gamma^0_{VV} - \gamma^0_{VH}$; unit: dB), was computed because the different intra-annual temporal profiles between VV and VH were found to identify phenology in deciduous forests (Frison et al., 2018). This index was called the ratio between VV and VH, as previous studies computed their ratio before log-transformation (Frison et al., 2018; Nicolau et al., 2021; Veloso et al., 2017). In this study, we named it CDI for the convenient labelling of the following indices. To reduce the impact of other, irrelevant effects on backscatter intensities, such as species composition or soil characteristics, the normalized CDI (NCDI = (CDI/CDI$_{\text{min during leafoff}}$)) was calculated by dividing the time-series CDI values at each plot by the minimum CDI value during the leaf-off season (CDI$_{\text{min during leafoff}}$; DOY 60–119; stage 1 in Figure 1) of each plot.

2.2.3 | ΔNCDI indices as surrogates of defoliation and refoliation

Four critical time-steps (A, B, C and D) indicating prominent phenological transitions were selected from the 18 timesteps of the NCDI of the 12-day composite time-series data based on the temporal trajectory of the NCDI and literature (e.g. bud-burst at DOY 109 for oak forests in Belgium in 2018 in Marchand et al. (2020) and defoliation of gypsy moth between DOY 120 and 180 with the peak at DOY 150 in de Beurs and Townsend (2008)). Step A was the start of leaf-flushing and the following defoliation (the fifth timestep; average DOY 108–119; label A in Figure 1), as the temporal trajectories of the NCDI were observed to be divided between heavily and lightly defoliated forests from the sixth timestep. Step B was the first peak of leaf-flushing

and was defined as the timestep with the maximum NCDI between the seventh and ninth timesteps (DOY 132–167; Stage 2 in Figure 1). Step C, the end of defoliation, was defined as the timestep with the minimum NCDI between the 10th and 15th timesteps (DOY 168–239; Stage 3 in Figure 1). Step D, the end of the growing season, was defined as the timestep with the maximum NCDI between the 17th and 18th timesteps (DOY 252–270; Stage 4 in Figure 1), as leaf senescence starts at the earliest DOY 270 in German deciduous forests for the last three decades (Liu et al., 2020). These four critical timesteps were used to calculate two indices related to defoliation (Equations 1–2) and two indices related to refoliation (Equations 3–4).

$$\Delta\text{NCDI}_{\text{Defoliation(A–C)}} = \text{NCDI}_{\text{Step A}} - \text{NCDI}_{\text{Step C}} \quad (1)$$

$$\Delta\text{NCDI}_{\text{Peak Defoliation(B–C)}} = \text{NCDI}_{\text{Step B}} - \text{NCDI}_{\text{Step C}} \quad (2)$$

$$\Delta\text{NCDI}_{\text{Peak Refoliation(D–C)}} = \text{NCDI}_{\text{Step D}} - \text{NCDI}_{\text{Step C}} \quad (3)$$

$$\Delta\text{NCDI}_{\text{Refoliation(D–B)}} = \text{NCDI}_{\text{Step D}} - \text{NCDI}_{\text{Step B}} \quad (4)$$

The four ΔNCDI indices of the five forest types of the a priori classification (HC, H2C, HM, LC and LM) were compared using pairwise Wilcoxon–Mann–Whitney rank-sum tests.

2.3 | Sentinel-2 data processing

Sentinel-2 optical data were also obtained from the ESA Scientific Hub on 56 plots for the same period covered by the Sentinel-1 data, between March and September 2019. All available level-1c products covering our study area in which the cloud cover was <70% were selected, resulting in products covering a total 15 days. Some were cloud-free, but most had patches of clouds and haze. The cloud-free and haze-free plots were selected from each product, and their normalized difference water index (NDWI) values were calculated, consisting of 699 values from the 15 days (on average, 12.5 day-values per plot; see the DOY and the number of plots per day in Figure S3).

$$\text{NDWI} = \frac{B_8 - B_{11}}{B_8 + B_{11}} \quad (5)$$

where B_8 is band 8 of the near-infrared wavelength (842 nm), and B_{11} is band 11 of the short-wave infrared wavelength (1,610 nm; see more explanation on NDWI in Supplementary Note 2). Spearman rank correlation coefficients using a two-sided test were calculated to compare (interpolated) single-day CDI and NDWI on the Sentinel-2 acquisition dates and plots ($n = 699$) due to the non-normal distributions of the two variables.

2.4 | Terrestrial laser scanning

Terrestrial laser scanning (TLS, RIEGL VZ-400i laser-scanning system; RIEGL Laser Measurement Systems) was conducted to track

the forest canopy structure in three dimensions before and after insect defoliation. The first campaign (TLS1) took place before the start of the 2019 growing season (DOY 79–107; 20 March–17 April; t1), the second campaign (TLS2) during peak feeding of the gypsy moth (DOY 179–200; 28 June–19 July, t2) and the third campaign (TLS3) after the end of defoliation (DOY 240–249; 27 August–5 September; t3) at 43 plots of HC, HM, LC and LM types. Since the GHM plot (HM type in a block G) missed the leaf-off time of the first campaign, it was excluded, such that the TLS data from 43 plots were used in the validation of the CDI derived from Sentinel-1. For TLS1, 20 centre-trees at each plot, that is, five centre-trees per the four cardinal directions, were scanned on an average of four times per tree. For TLS2 and TLS3, 12 centre-trees at each plot, that is, three centre-trees per the four cardinal directions, were scanned. From TLS1, TLS2 and TLS3, respectively, the point clouds within a boundary of the four circles with a radius of 20 m from the second centre-trees along the four cardinal directions (N2, S2, E2 and W2; Figure S4) were extracted to calculate the foliage index, as this area (*c.* 0.5 ha) was the best-scanned area (see Figure S4 for the scanning and extracted locations). Only points higher than 10 m above the ground were included, to exclude vegetation cover from the shrub and herb layers in the assessment of canopy defoliation. The point cloud was divided into voxels with a size of 125 cm³ ($5 \times 5 \times 5$ cm), and the number of occupied voxels (NoVoxels) was counted. NoVoxels on t1 (NoVoxels1) was used as a leaf-off baseline of forest structure. To include the points reflected from trunks and branches (sometimes hidden by foliage) in NoVoxels on t2 and t3 (NoVoxels2 and NoVoxels3), we merged each point clouds from TLS1 and TLS2 for NoVoxels2 before voxelating and TLS1 and TLS3 for NoVoxels3. The foliage index on t2 and t3 was calculated by dividing NoVoxels2 and NoVoxels3 by NoVoxels1, respectively. Spearman rank correlation coefficients were calculated to test the association between the foliage index on t2 and t3 and single-day NCDI on the same dates as t2 and t3 ($n = 86$, 43 plots × 2 campaigns).

2.5 | Caterpillar sampling

Caterpillars from the canopies were sampled using a pyrethrum-knockdown technique and four tarpaulin sheets (3×5 m) covering a total sampling area of 60 m² per forest plot. We conducted the sampling shortly after treatment (DOY 143–158; 23 May–7 June) and after the defoliation period (DOY 182–185; 1 July–4 July) with the records of fogging quality at 44 plots of HC, HM, LC and LM types (for details of sampling quality, see Supplementary Note 3). Plots with poor sampling qualities due to weather conditions, that is, mean qualities of two-times samplings <2.0, were excluded (see the comparison between all plots ($n = 44$) and plots excluding poor qualities ($n = 38$) in Figure S5). Spearman rank correlation coefficients were calculated to test the association between $\Delta NCDI_{Defoliation(A-C)}$ and the total number of the caterpillars over two-times samplings to include the non-normal distributions of variables. Before the statistical test, caterpillar abundances were natural log-transformed

to improve the right-skewed distributions. Quantile regression was applied to estimate the effect of caterpillar abundance on $\Delta NCDI_{Defoliation(A-C)}$ and $\Delta NCDI_{Peak\ Defoliation(B-C)}$. The quantile regression is appropriate if the predictor variable acts as an ecological limiting factor, and other potential limiting factors are not measured (Cade & Noon, 2003). In our study, caterpillar abundance was one of the limiting factors constraining defoliation severity, but other limiting factors such as parasite infection (Alalouni et al., 2013), tree species composition and its spatial arrangement (Yguel et al., 2011) were not reflected in the regression.

2.6 | A priori and a posteriori classification

All forest plots were classified according to the a priori categories H2C, HC, HM, LC and LM as well as with regard to the a posteriori categories DEF+ and DEF− by eight classification methods (see used methods in Supplementary Note 4), including naïve Bayes classification ('e1071'; Meyer et al., 2019), a method which originally gained broad attention when successfully utilized for spam identification and e-mail text classification (Sahami et al., 1998). We chose naïve Bayes as the reference method as it is well known, available in any popular computing environment, and a simple and efficient method for multi-class classification, specifically when the sample size is small (Ashari, 2013; Huang & Li, 2011; Sordo & Zeng, 2005). Seven algebraic metrics extracted from either the 12-day or the single-day NCDI time series were used as predictors: NCDI at steps B, C and D and four ΔNCDI indices (Equations 1–4). The classification methods were trained and evaluated based on threefold cross-validation with split datasets. The data were split into training ($n = 38$, ~68% of the complete dataset) and test ($n = 18$, ~32%) subsets for each cross-validation. The training data were chosen randomly within each class but were (approximatively) equally distributed between classes. Each classification method was parameterized with the training dataset, and its performance was then evaluated based on the test dataset. Threefold cross-validation was utilized to ensure that variation is representative both in the training and test dataset. In addition, it balances the conflicting requirements for low bias in model performance estimates and moderate variance (James et al., 2013). Six performance indices such as sensitivity and specificity (see more details in Table 1) were given as means of the threefold cross-validation for each class and arithmetic means, that is, (macro) average values, were obtained for entire classes (Herrera et al., 2016).

3 | RESULTS

The indices based on satellite radar (Sentinel-1), optical sensor (Sentinel-2) and TLS data were highly correlated. The CDI derived from the radar data correlated strongly with the NDWI derived from the optical data (Spearman's $\rho = 0.79$, $p < 0.001$) and with the amount of foliage estimated from TLS (Spearman's $\rho = 0.84$, $p < 0.001$; Figure 2). In some plots, the association between NDWI

TABLE 1 Quality indicators for the a priori and a posteriori classifications of the test dataset ($n = 18$) using the naïve Bayes classifier based on algebraic metrics calculated from the single-day time series. (a) Schematic confusion matrix for a category X (e.g. for HC) to explain the quality indicators in Table 1b. (b) Sens: sensitivity, fraction of cases correctly classified for a category X among all observations in that category: $a_x/(a_x + c_x)$; Spec: specificity, fraction of correctly classified complementary cases (e.g. not HC) among all observations in the complementary categories: $d_x/(b_x + d_x)$; PPV: positive predictive value, that is, fraction of correctly classified cases among all classifications as the specific category: $a_x/(a_x + b_x)$; NPV: negative predictive value, that is, fraction of correctly classified complementary cases among all classifications as a complementary category: $d_x/(c_x + d_x)$; F1: harmonic mean of the sensitivity and the positive predictive value: $a_x/(a_x + b_x/2 + c_x/2)$; Balanced accuracy: mean of the sensitivity and specificity: $\frac{1}{2}(a_x/(a_x + c_x) + d_x/(b_x + d_x))$. All values are givens as means of threefold cross-validation

(a)

	Observation	
	X	Not X
Prediction		
X	True positive: a_x	False positive: b_x
Not X	False negative: c_x	True negative: d_x

(b)

	Sens $a_x/(a_x + c_x)$	Spec $d_x/(b_x + d_x)$	PPV $a_x/(a_x + b_x)$	NPV $d_x/(c_x + d_x)$	F1 $2(Sens \times PPV)/(Sens + PPV)$	Balanced accuracy $\frac{1}{2}(Sens + Spec)$
A priori classification						
H2C	1.00	0.95	0.85	1.00	0.92	0.98
HC	0.64	1.00	1.00	0.92	0.77	0.82
HM	0.61	0.81	0.44	0.90	0.51	0.71
LC	0.56	0.89	0.58	0.89	0.55	0.72
LM	0.28	0.86	0.51	0.82	0.32	0.57
Mean (macro average)	0.62	0.90	0.68	0.9	0.61	0.76
A posteriori classification (binary)						
Heavily (DEF+)	0.93	0.97	0.93	0.97	0.93	0.95
Heavily (DEF−)	0.93	0.97	0.93	0.97	0.93	0.95
Mean (macro average)	0.93	0.97	0.93	0.97	0.93	0.95

and radar backscatters was low. For example, the NDWI values at plots NLM and NLC (LM and LC types in block N) in late summer were lower than expected from the CDI values, whereas at plot HHC in mid-summer and all plot types in block J in early spring the CDI values were lower than expected from the NDWI values (see Figure S3 for exact dates).

For $\tau = 0.50$ (the median) and $\tau = 0.90$, the quantile regressions of caterpillar abundance and $\Delta NCDI_{Defoliation(A-C)}$ yielded R^2 values of 0.52 and 0.71, respectively (quantile regression lines in Figure 3 and coefficient plots as a function of quantile in Figure S6). The latter indicated that the prediction that 90% of the values of defoliation severity would be less than or equal to the 90% quantile regression function (the highest dashed grey line in Figure 3) would have an accuracy of 71%. The interval between the 0.90 and 0.10 regression quantile estimates (80% prediction interval for a future observation) at forests with a high abundance of caterpillars was much wider than that at forests with fewer caterpillars. Thus, while maximum $\Delta NCDI_{Defoliation(A-C)}$ as a surrogate of defoliation severity

correlated strongly with caterpillar density, there were also forests with high numbers of caterpillars but without severe defoliation. $\Delta NCDI_{Peak\ Defoliation(B-C)}$ showed a similar result with R^2 values of 0.50 for $\tau = 0.50$ (see Figure S7).

The $\Delta NCDI_{Defoliation(A-C)}$ was significantly higher at plots H2C and HC, that is, plots with high-level damage, than at plots HM, LC and LM, as shown by pairwise Wilcoxon tests (Figure 4A, see Table S1 for exact p values). The same was determined for the $\Delta NCDI_{Peak\ Defoliation(B-C)}$ (Figure 4B). Specifically, after the herbivory period (step C), the defoliation severity at the H2C plots was >0, which meant that the H2C plots had even fewer leaves than before the start of herbivory (time-step A; Figure 4A). However, the $\Delta NCDI_{Peak\ Refoliation\ (D-C)}$, as a proxy of the amount of refoliation after defoliation, was also much higher at the H2C plots (Figure 4C). The $\Delta NCDI_{Refoliation(D-B)}$, as a proxy of the amount of refoliation after the first peak of leafing, was below zero at plots H2C and HC, indicating that after defoliation the amount of foliage did not entirely recover (Figure 4D).
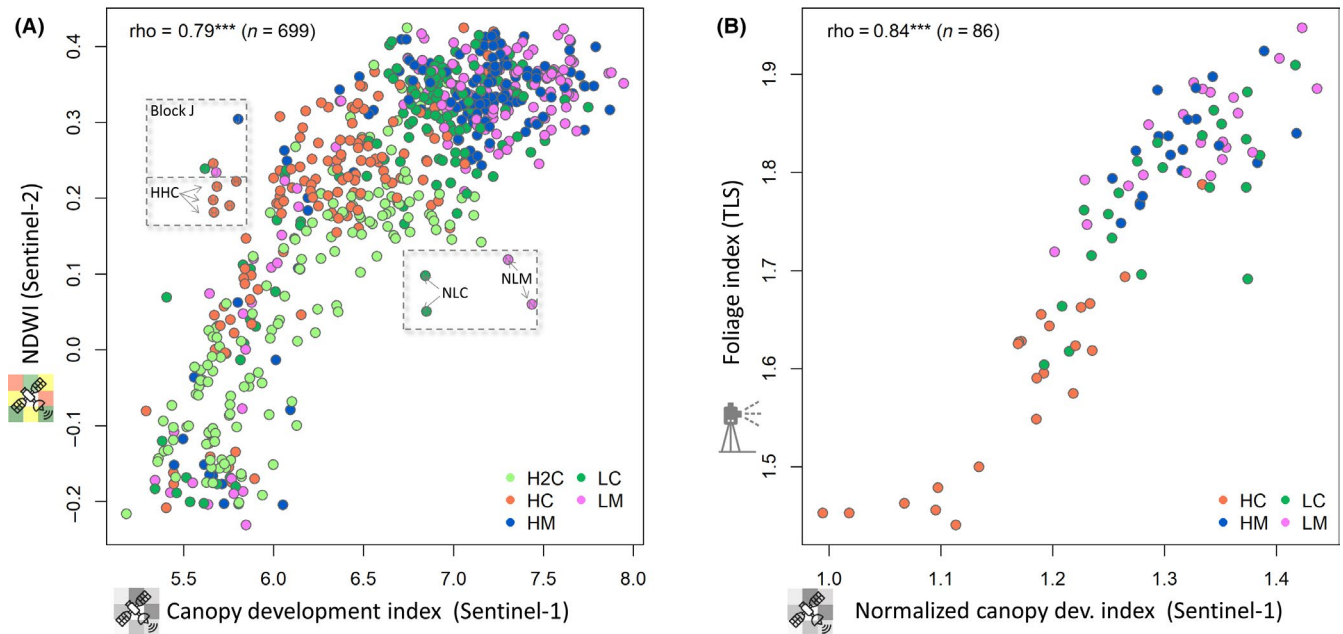
FIGURE 2 Correlations of the single-day canopy development index (CDI) and NCDI from Sentinel-1 with the vegetation indices from two other sensors. (A) with the normalized difference water index (NDWI) from Sentinel-2 (see the DOY and the number of plots per day in Figure S3). (B) with the foliage index as measured by terrestrial laser scanning (TLS; 43 plots on t2 and t3). The top left *rho* is the Spearman rank correlation coefficient, with a significance level of ***$p < 0.001$

The NCDI trajectories of DEF+ and DEF− differed (Figure 5). Most H2C and HC plots were identified as DEF+, thus characterized by a lower leafing peak and a signal intensity that decreased sharply during the peak of defoliation. Plots HM, LC and LM as well as some of the HC plots (AHC, GHC and SHC) were identified as DEF−. Compared to the plots of heavily defoliated forests, their values at the first leafing peak were higher and the decrease in signal intensity during the peak of herbivory was smaller.

A priori and a posteriori classifications yielded similar results for different classifiers (Tables S2). As an example, in the naïve Bayes classification, the macro-average of balanced accuracy for the a priori classification of the test data was 0.76, much higher than the score of a random classification (for additional performance indices, see Table 1). The sensitivity of the naïve Bayes classification, that is, the true positive rate, was highest for the H2C plot and lowest for the LM plot. In general, the specificity, that is, the true negative rate, was higher than the sensitivity. When the naïve Bayes algorithm was applied to the test dataset for the binary classification of defoliation severity (DEF+ vs. DEF−; Table 1, Tables S3), heavily or lightly defoliated forests were classified almost perfectly (see Tables S4), thus underlining the power of radar backscatter to identify highly defoliated forests.

4 | DISCUSSION

4.1 | Time-series canopy development index

Recently, Sentinel-1-based studies are applying SAR indices such as the ratio between VV and VH and Radar Vegetation Index

($4\gamma^{VH}/\gamma^{VH} + \gamma^{VV}$) for intra-annual temporal profiling of forest type and crop type (Frison et al., 2018; Holtgrave et al., 2020; Veloso et al., 2017). In temperate forests, a few studies showed seasonal differences in VV and VH intensity by forest type and used the intra-annual time-series information of those intensities in the classification of forest type (Dostálová et al., 2018; Frison et al., 2018). Our results first demonstrate that the ratio between the intensities of VV and VH (CDI) can be used to track the temporal trajectory and defoliation severity of canopy herbivory. Despite the success of MODIS time-series data in estimating forest defoliation by gypsy moths in previous studies (de Beurs & Townsend, 2008; Spruce et al., 2011), the spatial resolution of the predicted area was limited, for example, 0.6 km² or 0.25 km² of the smallest unit area (de Beurs & Townsend, 2008; Spruce et al., 2011), too coarse for management decisions for fragmented landscapes. Compared to temporally high-resolution optical sensors such as MODIS and Sentinel-2, the advantages of Sentinel-1 (pixel spacing of 10 m and repeat frequency of 2 days in our study) are its higher spatial resolution compared to MODIS (250 m) and its higher practical temporal resolution than Sentinel-2 which is strongly limited by cloud cover (12.5 day-values per plot in our study).

4.2 | CDI versus vegetation indices from two sensors

The smoothed and interpolated daily CDI (or NCDI) values from the Sentinel-1 well matched the single-day values of vegetation (or foliage) indices determined from Sentinel-2 and TLS, showing less
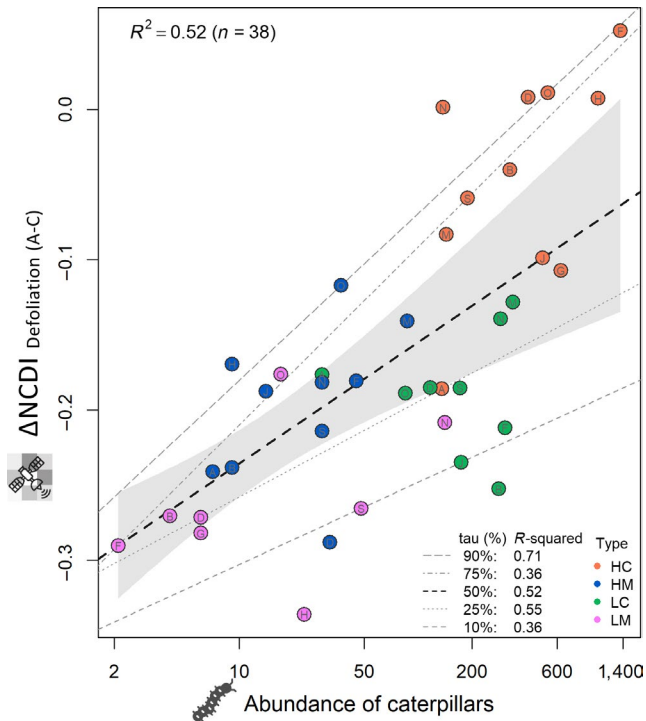
FIGURE 3 Quantile regression of the total number of caterpillars over two-times samplings and $\Delta NCDI_{Defoliation(A-C)}$ from radar. Superimposed on the plot are the $\{10, 25, 50, 75, 90\}$ quantile regression lines (grey dashed lines) and the median τ (black dashed line). The grey area represents the 90% confidence interval of the median quantile regression. The labels of the points indicate the spatial blocks. The top left R^2 is the value for $\tau = 0.50$ (the median)

signal saturation. The vegetation index using the short-wave infrared wavelength such as NDWI worked better than other optical-based indices for mapping defoliation by gypsy moths in previous studies (de Beurs & Townsend, 2008; Fraser & Latifovic, 2005; Townsend et al., 2012; see more on NDWI in Supplementary Note 2). In our study, the scatterplot of CDI versus NDWI (Figure 2A) revealed signal saturation at high NDWI but not high CDI values. It was reported that high biomass levels (lightly defoliated plots in our study design) resulted in saturation of the backscatter coefficient of C-band SAR (Tsui et al., 2012). By contrast, there was no signal saturation of CDI (or NCDI) with respect to forest type or the structure or region of the study site for both comparisons (Figure 2A,B). In few cases, the NDWI values did not match well with the CDI values (Figure 2A). The first mismatch involved the NLM and NLC plots in late summer, in which case the NDWI was lower than the CDI. This may have been due to the withered leaves that remained on the branches of drought-stricken trees, as suggested by field observations, showing CDI is less sensitive to tree drought. Conversely, at all forest plots in the J block in early spring the NDWI was higher than the CDI. This difference may have been caused by a high level of productivity by the small volume of newly flushed leaves.

Signal saturation also occurred at low values of the foliage index determined from TLS (Figure 2B). The TLS saturation at heavily defoliated plots was assumed to be the product of the difference in

canopy assessment by radar versus TLS. Our TLS assessment of canopy development excluded points lower than 10 m (i.e. vegetation cover by the understory), whereas radar data still captured the vegetation cover of the understory at heavily defoliated plots. Accordingly, the values of the foliage index were consistently lower than the NCDI values at heavily defoliated plots.

4.3 | $\Delta NCDI_{Defoliation}$ versus abundance of caterpillars

$\Delta NCDI_{Defoliation(A-C)}$ and $\Delta NCDI_{Peak\ Defoliation(B-C)}$ as surrogates of defoliation severity were closely associated with the abundance of caterpillars collected, as shown by quantile regression (Figure 3; Figure S7). Quantile regression was useful to predict the upper boundary of defoliation severity in response to an ecologically limiting factor (e.g. a higher percentile regression, such as the 0.90 quantiles), assuming that there may be other, unmeasured limiting factors. The finding of a wider prediction interval (e.g. 80% prediction interval between the 0.90 and 0.10 regression quantile estimates) in forests with a high abundance of caterpillars suggested a role for other abiotic or biotic factors that reduce canopy herbivory in forests with caterpillar infestations, such as the spatial arrangement of host tree species (Yguel et al., 2011) and pressure through parasitism and predation (Alalouni et al., 2013; Liebhold et al., 1998). In the study plots, oak was the most common tree species, and it hosted many species of moths (Brändle & Brandl, 2001), including gypsy moths (Milanović et al., 2014). According to Yguel et al. (2011), the phylogenetic isolation of oak trees from neighbouring trees decreases canopy herbivory. In future ecological studies, deviations in defoliation severity as determined from higher quantile regression lines could be used to identify the potential effects of vegetation, parasites and predators on the feeding activity of caterpillars.

4.4 | Classification between DEF+ and DEF−

The consistently high accuracy of the statistical classification of the a posteriori categories allowed almost perfect discrimination between DEF+ and DEF− (Table 1; Tables S2 and S4), based on the fine-scale time-series of Sentinel-1 (Figure 5). In other studies, the temporal trajectories of insect defoliation of broadleaf trees were analysed using time-series MODIS (Chávez et al., 2019; Olsson et al., 2016). However, the coarse spatial resolution of MODIS data is a limiting factor in studies of intensively managed and fine-grained landscapes such as those of our study area in Central Europe (Olsson et al., 2016). SAR data were not explored for defoliating insects due to the lack of open-access and high temporal-resolution data, except for Kaasalainen et al. (2010) having no success with ERS-2 data. Our results show the high potential of open-access high-resolution Sentinel-1 datasets for discriminating heavy defoliation by broadleaf defoliators. This would not be feasible with either Sentinel-2, due to cloud cover, or with TLS, due to its labour-intensive operating modus, although both have the higher spatial resolution.

FIGURE 4 The four ΔNCDI indices
related to defoliation and refoliation
according to the five a priori forest
categories. Significant differences are
indicated by letters above the figures
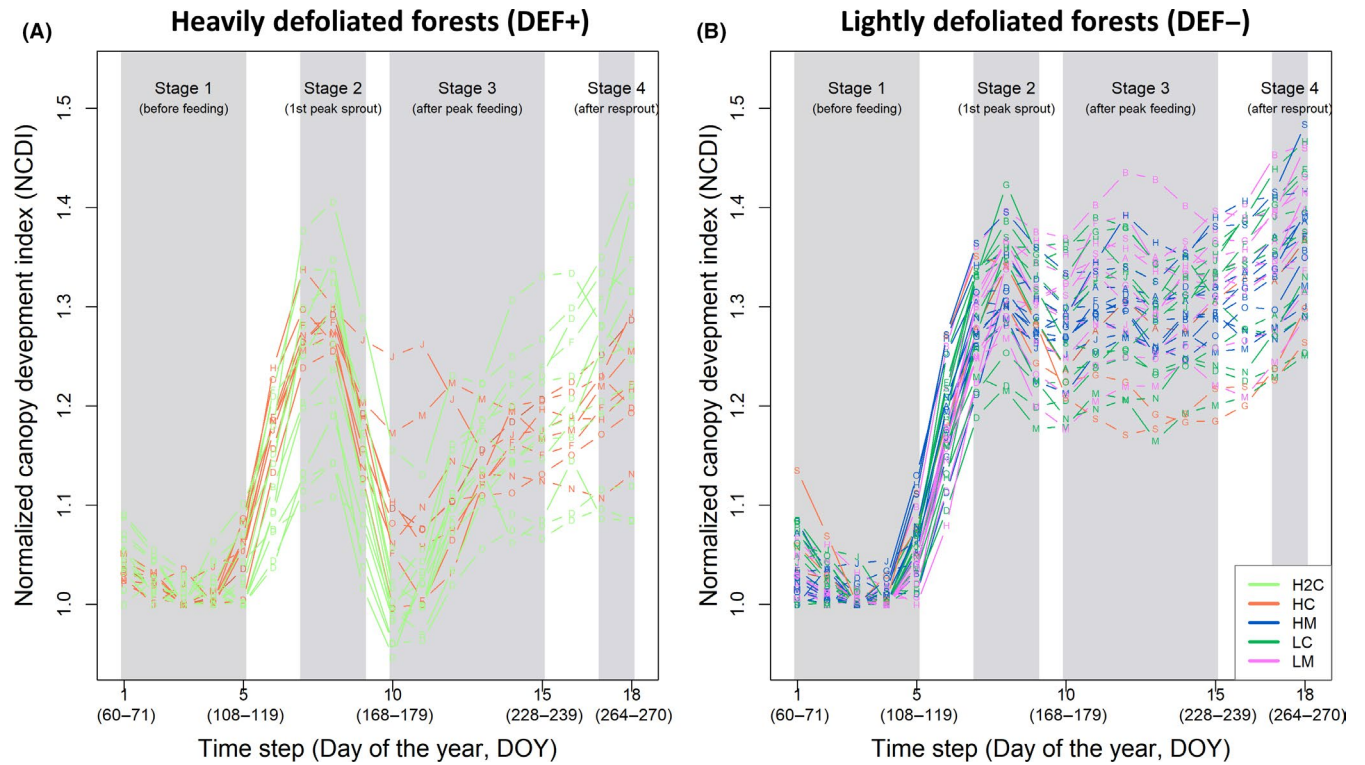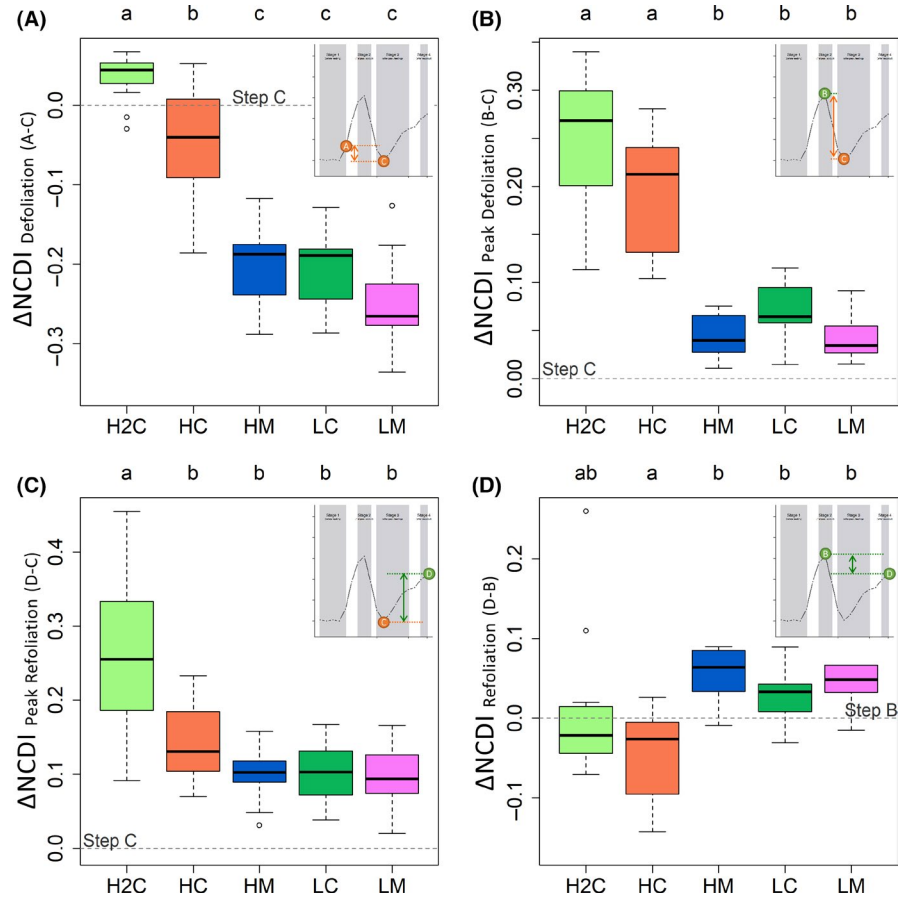and were determined using pairwise
Wilcoxon–Mann–Whitney rank-sum tests

FIGURE 5 Temporal trajectory of the normalized canopy development index (NCDI) in heavily (DEF+) and lightly (DEF−) defoliated forests. The labels of the points indicate the spatial blocks

We supplemented our analysis by discriminating the a priori classification, being aware of our limited sample size hardly sufficient for multi-parameter classification. Thus, classification quality of the five a priori categories differed. However, stands with high levels of gypsy moth egg masses and without treatment (H2C and HC) could clearly be discriminated from others while differentiation between similar treatments (e.g. HM and LM) was not successful. For further discussion, see Supplementary Note 5.

4.5 | Future application

Although our study was limited to 70 km extent, a small number of plots and a field dataset with the stand-level resolution, our results, proven by the intensively collected field dataset, opened the potential future studies of Sentinel-1 data on detecting defoliation severity gradient over a large scale. We suggested the potential use of NCDI for detecting the event of heavy defoliation by broadleaf defoliators using threshold-based classification at critical timesteps in Figure S8 and Supplementary Note 6. In addition, although perhaps more challenging, Sentinel-1 data may help investigate the causes of a defoliation event (Sommerfeld et al., 2018). For example, linking defoliation to the activity of main defoliators might be achieved by a careful examination of the specific temporal pattern of defoliation, specifically the characteristic feeding peak.

Canopy herbivory is an important variable in basic and applied ecological studies on insect herbivory (Alalouni et al., 2013; Liebhold et al., 1998; Yguel et al., 2011). Herbivory resistance, as a significant function in forest ecosystems (Pettorelli et al., 2016; Ratcliffe et al., 2017; Schuldt et al., 2018), is regularly surrogated as the inverse of canopy herbivory, which is often visually estimated by quantifying leaf damage by insects in a forest, based on a few branches. Our research has shown a way forward in quantifying canopy herbivory and herbivory resistance by remote sensing. Yet, Sentinel-1 data are needed to be tested for its applicability to future macroecological studies of herbivory resistance in broadleaf forests.

COMPETING INTEREST
The authors have no competing interests to declare.

AUTHORS' CONTRIBUTIONS
S.B., J.M., and O.M. conceived the idea of the study and wrote the first manuscript draft; B.M.L., J.M., T.H., W.W.W. and H.P. designed the field experiment; S.B., B.F., B.M.L.L., T.H., M.J. and S.H. collected the data; S.B., B.F., O.M., T.H. and M.J. analysed the data. All authors contributed to the revisions of the manuscript and gave final approval for publication.

PEER REVIEW
The peer review history for this article is available at https://publons.com/publon/10.1111/2041-210X.13726.

DATA AVAILABILITY STATEMENT
The data are available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.9w0vt4bgr (Bae et al., 2021). The batch processing configuration file for the SNAP toolbox software is available at https://github.com/So-YeonBae/Sentinel1-Biodiversity (Bae & Leutner, 2021).

ORCID
Soyeon Bae (iD) https://orcid.org/0000-0003-1961-1226
Jörg Müller (iD) https://orcid.org/0000-0002-1409-1586
Torben Hilmers (iD) https://orcid.org/0000-0002-4982-8867
Martin Jacobs (iD) https://orcid.org/0000-0002-2906-8661
Benjamin M. L. Leroy (iD) https://orcid.org/0000-0001-6007-7948
Hans Pretzsch (iD) https://orcid.org/0000-0002-4958-1868
Wolfgang W. Weisser (iD) https://orcid.org/0000-0002-2757-8959
Oliver Mitesser (iD) https://orcid.org/0000-0002-3607-877X

REFERENCES
(2011). The Holy Bible, New International Version. Biblica, Bible Gateway. Retrieved from www.biblegateway.com/versions/NewInternational-Version-NIV-Bible/

Abdullah, H., Skidmore, A. K., Darvishzadeh, R., & Heurich, M. (2019). Sentinel-2 accurately maps green-attack stage of European spruce bark beetle (*Ips typographus*, L.) compared with Landsat-8. *Remote Sensing in Ecology and Conservation*, 5, 87–106. https://doi.org/10.1002/rse2.93

Alalouni, U., Schädler, M., & Brandl, R. (2013). Natural enemies and environmental factors affecting the population dynamics of the gypsy moth. *Journal of Applied Entomology*, 137, 721–738. https://doi.org/10.1111/jen.12072

Ashari, A., Paryudi, I., & Min, A. (2013). Performance comparison between Naïve Bayes, decision tree and k-nearest neighbor in searching alternative design in an energy simulation tool. *International Journal of Advanced Computer Science and Applications*, 4. https://doi.org/10.14569/IJACSA.2013.041105

Bae, S., & Leutner, B. (2021). So-YeonBae/Sentinel1-Biodiversity: v1.0.0. *Zenodo*. https://doi.org/10.5281/zenodo.5510055

Bae, S., Levick, S. R., Heidrich, L., Magdon, P., Leutner, B. F., Wöllauer, S., Serebryanyk, A., Nauss, T., Krzystek, P., Gossner, M. M., Schall, P., Heibl, C., Bässler, C., Doerfler, I., Schulze, E.-D., Krah, F.-S., Culmsee, H., Jung, K., Heurich, M., … Müller, J. (2019). Radar vision in the mapping of forest biodiversity from space. *Nature Communications*, 10, 4757. https://doi.org/10.1038/s41467-019-12737-x

Bae, S., Müller, J., Förster, B., Hilmers, T., Hochrein, S., Jacobs, M., Leroy, B. M. L., Pretzsch, H., Weisser, W. W., Mitesser, O. (2021). Data from: Tracking the temporal dynamics of insect defoliation by high-resolution radar satellite data. *Dryad Digital Repository*, https://doi.org/10.5061/dryad.9w0vt4bgr

Brändle, M., & Brandl, R. (2001). Species richness of insects and mites on trees: Expanding Southwood. *Journal of Animal Ecology*, 70, 491–504. https://doi.org/10.1046/j.1365-2656.2001.00506.x https://doi.org/10.1046/j.1365-2656.2001.00506.x

Cade, B. S., & Noon, B. R. (2003). A gentle introduction to quantile regression for ecologists. *Frontiers in Ecology and the Environment*, 1, 412–420. https://doi.org/10.1890/1540-9295(2003)001[0412:AGITQR]2.0.CO;2

Carreiras, J. M. B., Quegan, S., Tansey, K., & Page, S. (2020). Sentinel-1 observation frequency significantly increases burnt area detectability in tropical SE Asia. *Environmental Research Letters*, 15. https://doi.org/10.1088/1748-9326/ab7765

Chávez, R. O., Rocco, R., Gutiérrez, Á. G., Dörner, M., & Estay, S. A. (2019). A self-calibrated non-parametric time series analysis approach for assessing insect defoliation of broad-leaved deciduous *Nothofagus pumilio* forests. *Remote Sensing*, 11, 204. https://doi.org/10.3390/rs11020204

de Beurs, K. M., & Townsend, P. A. (2008). Estimating the effect of gypsy moth defoliation using MODIS. *Remote Sensing of Environment*, 112, 3983–3990. https://doi.org/10.1016/j.rse.2008.07.008

Dostálová, A., Wagner, W., Milenković, M., & Hollaus, M. (2018). Annual seasonality in Sentinel-1 signal for forest mapping and forest type classification. *International Journal of Remote Sensing*, 39, 7738–7760. https://doi.org/10.1080/01431161.2018.1479788

Fraser, R. H., & Latifovic, R. (2005). Mapping insect-induced tree defoliation and mortality using coarse spatial resolution satellite imagery. *International Journal of Remote Sensing*, 26, 193–200. https://doi.org/10.1080/01431160410001716923

Frison, P.-L., Fruneau, B., Kmiha, S., Soudani, K., Dufrêne, E., Toan, T. L., Koleck, T., Villard, L., Mougin, E., & Rudant, J.-P. (2018). Potential of Sentinel-1 data for monitoring temperate mixed forest phenology. *Remote Sensing*, 10, 2049–2058. https://doi.org/10.3390/rs10122049

GEO BON. (2017). *GEO BON implementation plan 2017–2020*. Version 1.3. Group on Earth Observations Biodiversity Observation Network Secretariat, 101pp.

Giese, R. L., & Schneider, M. L. (1979). Cartographic comparisons of Eurasian gypsy moth distribution (*Lymantria dispar* L.; Lepidoptera: Lymantriidae). *Entomological News*, 90, 1–16.

Hamilton, N. (2015). *smoother: Functions relating to the smoothing of numerical data*. R package version 1.1. Retrieved from https://CRAN.R-project.org/package=smoother

Herrera, F., Charte, F., Rivera, A. J., & Del Jesus, M. J. (2016). *Multilabel classification*. Springer.

Hijmans, R. J. (2019). *raster: Geographic data analysis and modeling*. R package version 3.0-7. Retrieved from https://CRAN.R-project.org/package=raster

Hollaus, M., & Vreugdenhil, M. (2019). Radar Satellite Imagery for Detecting Bark Beetle Outbreaks in Forests. *Current Forestry Reports*, 5, 240–250. https://doi.org/10.1007/s40725-019-00098-z

Holtgrave, A.-K., Röder, N., Ackermann, A., Erasmi, S., & Kleinschmit, B. (2020). Comparing Sentinel-1 and -2 Data and Indices for Agricultural Land Use Monitoring. *Remote Sensing*, 12, 2919. https://doi.org/10.3390/rs12182919

Huang, X., Ziniti, B., Torbick, N., & Ducey, M. J. (2018). Assessment of forest above ground biomass estimation using multi-temporal C-band Sentinel-1 and Polarimetric L-band PALSAR-2 Data. *Remote Sensing*, 10, 1424. https://doi.org/10.3390/rs10091424

Huang, Y., & Li, L. (2011). Naive Bayes classification algorithm based on small sample set. *2011 IEEE International Conference on Cloud Computing and Intelligence Systems* (pp. 34–39). https://doi.org/10.1109/CCIS.2011.6045027

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning: With applications in R*. Springer.

Kaasalainen, S., Hyyppä, J., Karjalainen, M., Krooks, A., Lyytikäinen-Saarenmaa, P., Holopainen, M., & Jaakkola, A. (2010). Comparison of terrestrial laser scanner and synthetic aperture radar data in the study of forest defoliation. In W. W. A. B. Szekely (Eds.), *ISPRS Technical Commission VII Symposium: 100 Years ISPRS Advancing Remote Sensing Science* (pp. 82–87). ISPRS Council.

Leroy, B. M. L., Lemme, H., Braumiller, P., Hilmers, T., Jacobs, M., Hochrein, S., Kienlein, S., Müller, J., Pretzsch, H., Stimm, K., Seibold, S., Jaworek, J., Hahn, W. A., Müller-Kroehling, S., & Weisser, W. W. (2021). Relative impacts of gypsy moth outbreaks and insecticide treatments on forest resources and ecosystems: An experimental approach. *Ecological Solutions and Evidence*, 2, e12045.

Liebhold, A. M., Halverson, J. A., & Elmes, G. A. (1992). Gypsy moth invasion in North America: A quantitative analysis. *Journal of Biogeography*, 19, 513–520. https://doi.org/10.2307/2845770

Liebhold, A. M., Higashiura, Y., & Unno, A. (1998). Forest type affects predation on Gypsy Moth (Lepidoptera: Lymantriidae) Pupae in Japan. *Environmental Entomology*, 27, 858–862. https://doi.org/10.1093/ee/27.4.858

Liu, Q., Piao, S., Campioli, M., Gao, M., Fu, Y. H., Wang, K., He, Y., Li, X., & Janssens, I. A. (2020). Modeling leaf senescence of deciduous tree species in Europe. *Global Change Biology*, 26, 4104–4118. https://doi.org/10.1111/gcb.15132

Lovett, G., & Tobiessen, P. (1993). Carbon and nitrogen assimilation in red oaks (*Quercus rubra* L.) subject to defoliation and nitrogen stress. *Tree Physiology*, 12 3, 259–269. https://doi.org/10.1093/treephys/12.3.259

Lowe, S., Browne, M., Boudjelas, S., & De Poorter, M. (2000). *100 of the world's worst invasive alien species: A selection from the global invasive species database*. The Invasive Species Specialist Group (ISSG) a specialist group of the Species Survival Commission (SSC) of the World Conservation Union (IUCN).

Marchand, L. J., Dox, I., Gričar, J., Prislan, P., Leys, S., Van den Bulcke, J., Fonti, P., Lange, H., Matthysen, E., Peñuelas, J., Zuccarini, P., & Campioli, M. (2020). Inter-individual variability in spring phenology of temperate deciduous trees depends on species, tree size and previous year autumn phenology. *Agricultural and Forest Meteorology*, 290. https://doi.org/10.1016/j.agrformet.2020.108031

Meng, R., Dennison, P. E., Zhao, F., Shendryk, I., Rickert, A., Hanavan, R. P., Cook, B. D., & Serbin, S. P. (2018). Mapping canopy defoliation by herbivorous insects at the individual tree level using bi-temporal airborne imaging spectroscopy and LiDAR measurements. *Remote Sensing of Environment*, 215, 170–183. https://doi.org/10.1016/j.rse.2018.06.008

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., & Leisch, F. (2019). *e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien*. R package version 1.7-1. Retrieved from https://CRAN.R-project.org/package=e1071

Milanović, S., Lazarević, J., Popović, Z., Miletić, Z., Kostić, M., Radulović, Z., Karadžić, D., & Vuleta, A. (2014). Preference and performance of the larvae of Lymantria dispar (Lepidoptera: Lymantriidae) on three species of European oaks. *European Journal of Entomology*, 111, 371–378. https://doi.org/10.14411/eje.2014.039

Millar, C. I., & Stephenson, N. L. (2015). Temperate forest health in an era of emerging megadisturbance. *Science*, 349, 823–826. https://doi.org/10.1126/science.aaa9933

Mitchell, A. L., Rosenqvist, A., & Mora, B. (2017). Current remote sensing approaches to monitoring forest degradation in support of countries measurement, reporting and verification (MRV) systems for REDD. *Carbon Balance and Managment*, 12, 9. https://doi.org/10.1186/s13021-017-0078-9

Nicolau, A. P., Flores-Anderson, A., Griffin, R., Herndon, K., & Meyer, F. J. (2021). Assessing SAR C-band data to effectively distinguish modified land uses in a heavily disturbed Amazon forest. *International Journal of Applied Earth Observation and Geoinformation*, 94. https://doi.org/10.1016/j.jag.2020.102214

Olsson, P.-O., Lindström, J., & Eklundh, L. (2016). Near real-time monitoring of insect induced defoliation in subalpine birch forests with MODIS derived NDVI. *Remote Sensing of Environment*, 181, 42–53. https://doi.org/10.1016/j.rse.2016.03.040

Orozumbekov, A. A., Liebhold, A. M., Ponomarev, V. I., & Tobin, P. C. (2009). Gypsy moth (Lepidoptera: Lymantriidae) in Central Asia. *American Entomologist*, *55*, 258–265.

Ortiz, S. M., Breidenbach, J., & Kändler, G. (2013). Early detection of bark beetle green attack using TerraSAR-X and RapidEye data. *Remote Sensing*, *5*, 1912–1931. https://doi.org/10.3390/rs5041912

Pettorelli, N., Wegmann, M., Skidmore, A., Mücher, S., Dawson, T. P., Fernandez, M., Lucas, R., Schaepman, M. E., Wang, T., O'Connor, B., Jongman, R. H. G., Kempeneers, P., Sonnenschein, R., Leidner, A. K., Böhm, M., He, K. S., Nagendra, H., Dubois, G., Fatoyinbo, T., … Geller, G. N. (2016). Framing the concept of satellite remote sensing essential biodiversity variables: Challenges and future directions. *Remote Sensing in Ecology and Conservation*, *2*, 122–131. https://doi.org/10.1002/rse2.15

Ranson, K. J., Kovacs, K., Sun, G., & Kharuk, V. I. (2003). Disturbance recognition in the boreal forest using radar and Landsat-7. *Canadian Journal of Remote Sensing*, *29*, 271–285. https://doi.org/10.5589/m02-096

Ratcliffe, S., Wirth, C., Jucker, T., van der Plas, F., Scherer-Lorenzen, M., Verheyen, K., Allan, E., Benavides, R., Bruelheide, H., Ohse, B., Paquette, A., Ampoorter, E., Bastias, C. C., Bauhus, J., Bonal, D., Bouriaud, O., Bussotti, F., Carnol, M., Castagneyrol, B., … Baeten, L. (2017). Biodiversity and ecosystem functioning relations in European forests depend on environmental context. *Ecology Letters*, *20*, 1414–1426. https://doi.org/10.1111/ele.12849 https://doi.org/10.1111/ele.12849

Régnière, J. (2009). Predicting insect continental distributions from species physiology. *Unasylva*, *60*, 37–42.

Reiche, J., Hamunyela, E., Verbesselt, J., Hoekman, D., & Herold, M. (2018). Improving near-real time deforestation monitoring in tropical dry forests by combining dense Sentinel-1 time series with Landsat and ALOS-2 PALSAR-2. *Remote Sensing of Environment*, *204*, 147–161. https://doi.org/10.1016/j.rse.2017.10.034

Sahami, M., Dumais, S. T., Heckerman, D., & Horvitz, E. (1998). *A Bayesian Approach to Filtering Junk E-Mail*. AAAI 1998.

Schuldt, A., Assmann, T., Brezzi, M., Buscot, F., Eichenberg, D., Gutknecht, J., Härdtle, W., He, J.-S., Klein, A.-M., Kühn, P., Liu, X., Ma, K., Niklaus, P. A., Pietsch, K. A., Purahong, W., Scherer-Lorenzen, M., Schmid, B., Scholten, T., Staab, M., … Bruelheide, H. (2018). Biodiversity across trophic levels drives multifunctionality in highly diverse forests. *Nature Communications*, *9*, 2989. https://doi.org/10.1038/s41467-018-05421-z https://doi.org/10.1038/s41467-018-05421-z

Seidl, R., Schelhaas, M.-J., Rammer, W., & Verkerk, P. J. (2014). Increasing forest disturbances in Europe and their impact on carbon storage. *Nature Climate Change*, *4*, 806–810. https://doi.org/10.1038/nclimate2318

Senf, C., Pflugmacher, D., Wulder, M. A., & Hostert, P. (2015). Characterizing spectral–temporal patterns of defoliator and bark beetle disturbances using Landsat time series. *Remote Sensing of Environment*, *170*, 166–177. https://doi.org/10.1016/j.rse.2015.09.019

Senf, C., Seidl, R., & Hostert, P. (2017). Remote sensing of forest insect disturbances: Current state and future directions. *International Journal of Applied Earth Observation and Geoinformation*, *60*, 49–60. https://doi.org/10.1016/j.jag.2017.04.004

Sommerfeld, A., Senf, C., Buma, B., D'Amato, A. W., Després, T., Díaz-Hormazábal, I., Fraver, S., Frelich, L. E., Gutiérrez, Á. G., Hart, S. J., Harvey, B. J., He, H. S., Hlásny, T., Holz, A., Kitzberger, T., Kulakowski, D., Lindenmayer, D., Mori, A. S., Müller, J., … Seidl, R. (2018). Patterns and drivers of recent disturbances across the temperate forest biome. *Nature Communications*, *9*, 4355. https://doi.org/10.1038/s41467-018-06788-9 https://doi.org/10.1038/s41467-018-06788-9

Sordo, M., & Zeng, Q. (2005). On sample size and classification accuracy: A performance comparison. In J. L. Oliveira, V. Maojo, F. Martín-Sánchez, & A. S. Pereira (Eds.), *Biological and Medical data analysis* (pp. 193–201). Springer.

Spruce, J. P., Sader, S., Ryan, R. E., Smoot, J., Kuper, P., Ross, K., Prados, D., Russell, J., Gasser, G., McKellip, R., & Hargrove, W. (2011). Assessment of MODIS NDVI time series data products for detecting forest defoliation by gypsy moth outbreaks. *Remote Sensing of Environment*, *115*, 427–437. https://doi.org/10.1016/j.rse.2010.09.013

Tanase, M. A., Aponte, C., Mermoz, S., Bouvet, A., Le Toan, T., & Heurich, M. (2018). Detection of windthrows and insect outbreaks by L-band SAR: A case study in the Bavarian Forest National Park. *Remote Sensing of Environment*, *209*, 700–711. https://doi.org/10.1016/j.rse.2018.03.009

Tomppo, E., Ronoud, G., Antropov, O., Hytönen, H., & Praks, J. (2021). Detection of forest windstorm damages with multitemporal SAR data – A case study: Finland. *Remote Sensing*, *13*, 383. https://doi.org/10.3390/rs13030383

Townsend, P. A., Singh, A., Foster, J. R., Rehberg, N. J., Kingdon, C. C., Eshleman, K. N., & Seagle, S. W. (2012). A general Landsat model to predict canopy defoliation in broadleaf deciduous forests. *Remote Sensing of Environment*, *119*, 255–265. https://doi.org/10.1016/j.rse.2011.12.023

Tsui, O. W., Coops, N. C., Wulder, M. A., Marshall, P. L., & McCardle, A. (2012). Using multi-frequency radar and discrete-return LiDAR measurements to estimate above-ground biomass and biomass components in a coastal temperate forest. *ISPRS Journal of Photogrammetry and Remote Sensing*, *69*, 121–133. https://doi.org/10.1016/j.isprsjprs.2012.02.009

Veloso, A., Mermoz, S., Bouvet, A., Le Toan, T., Planells, M., Dejoux, J.-F., & Ceschia, E. (2017). Understanding the temporal behavior of crops using Sentinel-1 and Sentinel-2-like data for agricultural applications. *Remote Sensing of Environment*, *199*, 415–426. https://doi.org/10.1016/j.rse.2017.07.015

Yguel, B., Bailey, R., Tosh, N. D., Vialatte, A., Vasseur, C., Vitrac, X., Jean, F., & Prinzing, A. (2011). Phytophagy on phylogenetically isolated trees: Why hosts should escape their relatives. *Ecology Letters*, *14*, 1117–1124. https://doi.org/10.1111/j.1461-0248.2011.01680.x https://doi.org/10.1111/j.1461-0248.2011.01680.x

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.