Energy Informatics

RESEARCH Open Access

A digital twin of a local energy system based on real smart meter data

Daniel Bayer* and Marco Pruckner

*Correspondence: daniel.bayer@uni-wuerzburg.de

Modeling and Simulation, University of Würzburg, Würzburg, Germany

Abstract

The steadily increasing usage of smart meters generates a valuable amount of high-resolution data about the individual energy consumption and production of local energy systems. Private households install more and more photovoltaic systems, battery storage and big consumers like heat pumps. Thus, our vision is to augment these collected smart meter time series of a complete system (e.g., a city, town or complex institutions like airports) with simulatively added previously named components. We, therefore, propose a novel digital twin of such an energy system based solely on a complete set of smart meter data including additional building data. Based on the additional geospatial data, the twin is intended to represent the addition of the abovementioned components as realistically as possible. Outputs of the twin can be used as a decision support for either system operators where to strengthen the system or for individual households where and how to install photovoltaic systems and batteries. Meanwhile, the first local energy system operators had such smart meter data of almost all residential consumers for several years. We acquire those of an exemplary operator and discuss a case study presenting some features of our digital twin and highlighting the value of the combination of smart meter and geospatial data.

Keywords: Digital twin, Simulation, Local energy system, Decision support system, Smart meter data utilization, Future energy grid exploration

Introduction

In this and the next decade, local energy systems that supply residential households and small companies face with major challenges. For instance, the number of rooftop and open-space Photovoltaic (PV) installations has to increase to fulfill the ambitious expansion targets for PV systems. The Federal Network Agency of Germany (2022) expects the installed rooftop PV power to be 172.7 GW in 2037 resulting in 4.8 times the value of 2021. Furthermore, the coupling of the electricity and heat sectors will increase, mainly driven by the expansion of heat pumps to reduce fossil fuel consumption (Sterchele and Palzer 2017). For example, the share of heat pumps among the heating systems of newly constructed buildings broke through the 50% mark for the first time in 2021 according to the Federal Statistical Office of Germany (2022), displacing traditionally used gas boilers and oil heating. Such future energy systems also involve the intelligent and sector-coupled control of all components (Gharavi and Ghafurian 2011).

Another development in current and future energy systems is the expansion of smart metering infrastructure. According to Kerai (2022), over 50% of all residential electricity meters are smart meters in the United Kingdom at the end of 2021. This development leads to a large amount of data about electricity demand and feed-in with a high spatial resolution, which gets very valuable in the context of energy system decarbonization. For instance, this data can be used for load profiling, forecasting or demand response tasks on individual building level (Wang et al. 2019). Suppose this data is collected for an entire town or city. In this case, the question arises of how this data can be used in conjunction with geospatial data for simulating future grid states and how the data can be utilized in decision support systems for analyzing where residential PV installations or community batteries are profitable, or where grid expansion or a retrofit of substations are required.

Energy systems research is a vivid field of research. There are many publications on single aspects of future energy systems, e.g., optimal component sizing based on smart meter data, sector coupling using heat pumps, or optimal control of heat pumps and battery storage systems. We identify a lack of theoretical approaches focusing on a full combination of these aspects. As highlighted in the related work, existing approaches do not integrate smart meter data in combination with corresponding building conditions while performing individual expansion decisions model endogenous.

An upcoming trend in the application of computer science is the creation of digital twins, which originates in Product Lifecycle Management (Grieves 2005). A digital twin is a virtual model of an existing product or system with an information exchange both from the existing system to the virtual model and vice versa (Singh et al. 2021; Jones et al. 2020). A good overview of the digital twin concept can be found in VanDerHorn and Mahadevan (2021). As publications like Petrova-Antonova and Ilieva (2020) promote digital twins for decision support on a general level, we contribute a digital twin of a complete local energy system for exploring future grid states, such as higher PV or heat pump penetration rates, especially with the focus on the residential sector. The twin is based on the complete smart meter data of a residential energy system and additional data about roof sections and the heat demand of the existing buildings. The goal is that additional simulated components imitate the expected behavior as they would have in reality. So, the digital twin is used as a decision support system to assist

1) residential customers where a PV installation (and potentially combined with a battery storage system) is profitable,
2) local system operators with the evaluation of grid related issues like where to expand the grid and the assessment of community batteries, and
3) researchers for the evaluation of different control strategies concerning the charging and discharging of the batteries as well as the exploitation of flexibility potentials resulting from sector coupling.

The described methodology can be used to model a digital twin of an arbitrary local energy system, from a larger company with many buildings and individually metered consumers like airports or harbors over small towns to big cities or counties. The only requirement is smart meter time series availability and additional information about roof areas and building heat demand. We demonstrate our digital twin using data from our project partner Stadtwerk Haßfurt GmbH.

The remaining part of this publication has the following structure. First, we review the literature about the usage of smart meter data and simulation methods in energy systems research in the Related work section. Thereupon, we describe the modeling of our digital twin in the Methodology section. In the subsequent main part, we present an exemplary case study that exemplifies some of the capabilities of the digital twin and demonstrates the value of geospatial information. Finally, we end with a Conclusion and Outlook on future work.

Related work

Literature pertaining to energy systems based on a large amount of smart meter data presents a wide range of use cases. For the use case of data analysis, Iyengar et al. (2016) utilize data from over 14,000 smart meters widely spread across one city and analyzes impacts on energy consumption, for example hot summer or cold winter days. Other authors like McLoughlin et al. (2015), Khan et al. (2018) or Haben et al. (2016) use smart meter data for characterizing residential customers according to their daily energy usage patterns utilizing clustering techniques. In Miyasawa et al. (2021), the smart meter data of a major city is used for forecasting the spatial demand in the next hours.

Fewer papers address the simulative enrichment of smart meter data with PV installations or battery storage systems. Nyholm et al. (2016) investigate how battery storage can increase self-sufficiency and self-consumption of more than 2000 households distributed throughout Sweden. For every investigated household the authors obtained the smart meter profiles from the Distribution Network Operator (DNO). They prove that a residential battery storage can increase self-sufficiency by 12.5–30% compared to the case without storage. Based on the smart meter data of more than 4000 households Schopfer et al. (2018) present a predictor giving the optimal PV and battery configuration. In both examples, the effect of PV and battery storage on the local energy system remains an open question. The effect of an expansion of residential PV installations and heat pumps based on a large amount of smart meter data is investigated by Nigmatulina et al. (2020), where the authors analyze the increase in peak loads at the household and local substation level caused by higher PV penetration rates. Therefore, they use real data from around 15,000 buildings. The authors conclude that peak reduction can be more profitable in some cases than only increasing self-consumption. In Klonari et al. (2016) smart meter data is used to predict the maximum installable rooftop PV power that can be installed to prevent substation overloading. Douglass et al. (2019) use smart meter data from 300 households with PV and battery storage to predict system load if tariff regulations are changed. We notice that all these approaches do not respect individual roof conditions. Nevertheless, Khan et al. (2019) emphasize that roof orientation should be considered when simulating new PV systems.

An early approach of modeling a digital twin imitating the electrical behavior of households, especially with a focus on whiteware and electric mobility is presented by Karnouskos and de Holanda (2009). As there was neither smart meter data nor good data about mobility behavior available, the authors present a way to simulate the usage of whiteware and mobility demands, including a control for maximizing the usage of renewable energy generation. Moreover, aspects like the realistic planning of PV installations for rooftops have not been considered. Francisco et al. (2020) present how smart meter data can be used for almost real-time building energy benchmarking. The authors propose an energy management system on city-level for detecting building renovation potential and real-time energy efficiency strategies. In contrast to our approach, the energy management system does not consider the possible integration of renewable energies. Moreover, the effects of possible renovations, for example by heat pumps, are not in the focus of their publication. Ruohomaki et al. (2018) present a digital twin of the city of Helsinki where sensor data about air quality and building energy consumption is combined with a 3D model of all buildings. One goal is to compare similar buildings among each other and discover savings opportunities. Moreover, urban designers use a digital twin to analyze the solar energy potential in a part of Helsinki (Hämäläinen 2020). A methodological framework for building universal digital twins of smart cities is presented briefly by Petrova-Antonova and Ilieva (2020) on a general level without concrete implementations. Shahat et al. (2021) present a review of city digital twins and highlight gaps in knowledge, like the combination of energy-related data with geospatial data.

In energy systems research, simulation methods are commonly used to support component planning and sizing decisions or control strategy evaluation. Pruckner et al. (2014) present a combination of an optimization and simulation model to support the decision of when to build and how to size new gas-fired power plants to replace traditional coal or nuclear power plants. Blasi et al. (2019) evaluate the effect of higher PV penetration rates on substation level using a power flow calculation. Von Appen et al. (2015) analyze the grid impact of residential PV installations in Australia and Germany and probable changes if a battery storage system is added to individual households. Lazzeroni et al. (2015) present a study on the impact of general PV penetration on a distribution grid in the Middle East.

A first analysis of the effects of higher heat pump penetration rates in exemplary distribution systems is presented by Akmal et al. (2014), noting the threat of overloads at substations. Overloads due to higher heat pump penetration rates can also occur at the transmission level (Waite and Modi 2020). As a higher penetration of heat pumps might impact local energy systems, Ali et al. (2022) compare different strategies for planning local energy system strengthening, especially on power line level. Edmunds et al. (2021) evaluate the maximal hosting capacity in local energy systems for added heat pumps and optimized electric vehicles. A combination of PV and a heat pump including battery or thermal storage is evaluated in different climate zones for one exemplary residential house by Bee et al. (2019). Table 1 lists the most important references focusing on local energy systems and categorizes them according to the criteria considered in this publication.

Gaps in knowledge and research questions

Summarizing, we can say that the combination of real smart meter data, PV systems and additional heat pumps including individual component sizing, different control strategies and grid topology is the essence for building a digital twin of a local energy system. This combination has not yet been fully worked out in the references. In this context,

Table 1 Classification of the related work in the energy systems research

Reference	Level	PV	Network structure included	Smart meter data	Different control strategies	HPs considered	Indiv. comp. sizing
Blasi et al. (2019)	D	✓	✓	No	No	No	No
Nigmatulina et al. (2020)	D	✓	✓	No	No	No	No
Schopfer et al. (2018)	H	✓	-	✓	No	No	✓
Nyholm et al. (2016)	H	✓	-	✓	No	No	✓
Paolo et al. (2015)	D	✓	✓	No	No	No	No
Von Appen et al. (2015)	D+T	✓	No	No	✓	No	✓
Klonari et al. (2016)	D	✓	✓	✓	No	No	No
Douglass et al. (2019)	D	✓	No	✓	✓	No	No
Ali et al. (2022)	D	No	✓	No	No	✓	No
Edmunds et al. (2021)	D	✓	✓	No	No	✓	No
Bee et al. (2019)	H	✓	-	No	No	✓	✓
Our proposed digital twin	H+D	✓	✓	✓	✓	✓	✓

Column 'level': considered system level, i.e. household (H), local energy system (D) or transmission system (T). The other columns are self-explanatory

we highlight that all references in Table 1 are not able to integrate constructional conditions, namely roof orientation and size. Moreover, we find that most of the considered references focusing on distribution system level add simulated components to random households, which does not represent the system or customer behavior realistically (Palm 2020; Schopfer et al. 2018).

The concept of digital twins seems to be a promising approach to build a model of a local energy system for the joint answering of individual component sizing and expansion recommending, grid analysis and the potential to integrate different control strategies. The need for research on digital twins based on smart meter data and geospatial data is reinforced by the fact that first reviews on the use of digital twins by cities like Onile et al. (2021) do not yet list approaches that integrate smart meter data and geospatial data, but do highlight the potential benefits of such approaches (Shahat et al. 2021).

Hence, our research question is how to model a digital twin of a local energy system based on smart meter data and additional building information to be able to act as a decision support system for the aspects highlighted in the introduction. Our particular focus is on model-endogenous individual decision-making for PV and storage addition based on metrics such as Self-Sufficiency Rate (SSR) in conjunction with system analysis capabilities. The latter includes, for example, an analysis of the grid load or different storage configurations and the possibility to evaluate different control strategies.

Methodology

Modeling of the digital twin

We take an existing local energy system as a starting point for modeling the digital twin. Our granularity level is the system level, i.e., we model individual households, but the household consumption is modeled as one entity given by the data. This approach includes the consideration of power flows only at a balanced level. We explicitly mirror the spatial system structure, including the allocation of the metering points to their corresponding substations besides the individual roof and heating demand conditions. A metering point has one smart meter that measures an individual household or special components such as wind farms or open-space PV installations. Moreover, some residential rooftop PV installations and heat pumps have an individual meter and are not connected to the households' meter. So, we model a metering point and the information about the metered components as one unit, called a *measurement unit*. For every measurement unit we have a time series. Thus, a more detailed look at the components and stuff metered behind is not required. In general, the components we consider in this paper are:

- PV installations
- battery storage systems
- heat pumps

We model an intermediate layer between the substations and the measurement units, the so-called *control units*. These units aim to enrich the collected smart meter data stored in the measurement units with added simulated components, i.e., PV installations, battery storage systems and heat pumps. So, a control unit reflects the level where decisions regarding the simulative addition and the control of such simulated, added components are made. Regardless of the building type, i.e., a single-family detached home or a multi-family residential, a rooftop PV installation and a heat pump will be installed on building level, not on apartment unit level. Moreover, we can assume that all components inside a building could potentially communicate with each other. Thus, every building is modeled as a single control unit in our digital twin. Only multiple buildings on the same property belonging to each other are merged into one single control unit, like a detached house with a separate garage. The correspondence of the real energy system and the mapping in the digital twin are presented in Fig. 1. The concept of adding components on building (or property) level, but not on a flat level, follows the idea of Steber et al. (2017), where PV or battery is added once per building only.

We notice that a control unit can have multiple measurement units due to different reasons. First, an existing PV installation or heat pump can have an individual meter. Or second, a multi-family residential holds a meter for every apartment unit. Figure 1 visualizes some examples. The definition of the control unit also allows the representation of open-space PV installations and local wind farms inside the digital twin. A real energy system always shows a tree structure on local system level (Lakervi and Holmes 1995, ch. 10). Each consumption point is connected to exactly one substation. So, a control unit is connected to exactly one substation in our digital twin.

Simulated addition of components and virtual smart meters

The addition of new, simulated components takes place inside the control units. Since the measurement units know which components they are measuring, the control unit can use this information united over all connected measurement units to know which simulated components can be added, or in the case of PV installations the size of an existing system. As simulatively added components will change the demand from or feed-in to the energy system, we introduce a so-called *virtual smart meter* that measures

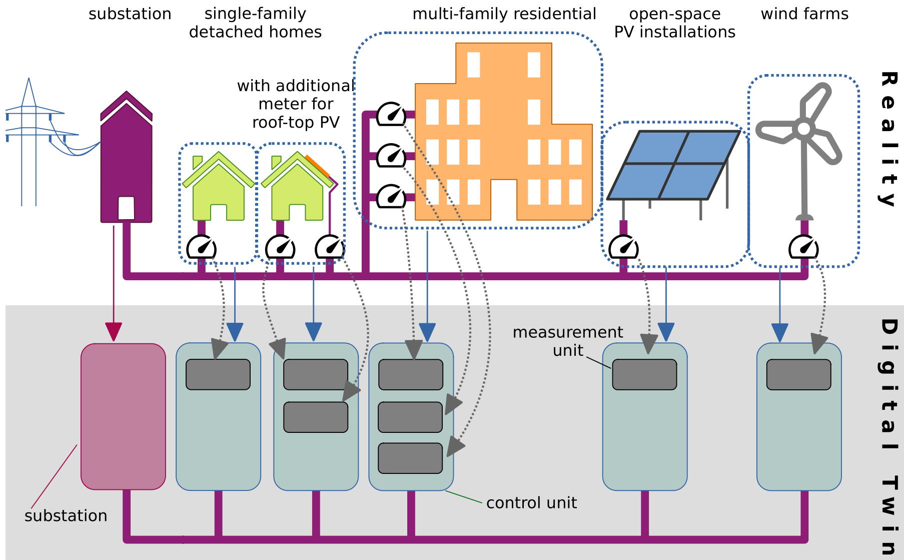

Fig. 1 The correspondence of the real energy system and the abstraction in the digital twin. The real modeled system is shown above. Below, the correspondence to the concepts of our digital twin are depicted: Turquoise boxes represent control units, and gray boxes represent measurement units. Additionally, the possible representation of local wind farms and open-space photovoltaic installations is illustrated on the right. Grey arrows represent the transformation of smart meters to measurement units, blue arrows represent the transformation of real entities to control units, crimson arrows depict the transformation of substations

the energy balance inside the control unit. Figure 2 illustrates the connection between the real smart meter data inside a measurement unit and the virtual smart meter of the complete control unit. The control unit can be expanded with a PV installation, a battery storage system, a heat pump or arbitrary combinations of these three components.

Simulation concept inside the digital twin

All variables and sets used in the further course are listed in Table 2. To simulate the system behavior given an arbitrary addition of simulated components over the complete time span for which data is available, the digital twin incorporates a discrete simulation with fixed time steps. The step size must be equidistant and has to fit the temporal resolution of the smart meter data. The simulation repeats the same procedure in each time step $t \in \mathcal{T}$, where $\mathcal{T}$ is the ordered set of all time steps. In principle, the load at every virtual smart meter $P_c^{vSM}(t)$ is first calculated for each control unit $c \in \mathcal{C}$, where $\mathcal{C}$ is the set of all control units. Subsequently, the active power is calculated for each substation by summing the readings of the virtual smart meter of all control units connected. By summing the power of all substations, we get the total active power of the system.

For computing the power of the virtual smart meters $P_c^{vSM}(t)$ as illustrated in Fig. 3, the simulation does the following steps for every control unit c: First, every measurement unit connected to c reads the value of the real smart meter from the data provided for the current time step. Thereupon, the demand or feed-in of all simulatively added components of c is computed giving a fixed order according to their flexibility potential, starting with the PV installation as an uncontrollable component and ending with the battery storage system as the one with most flexible control capabilities. The individual

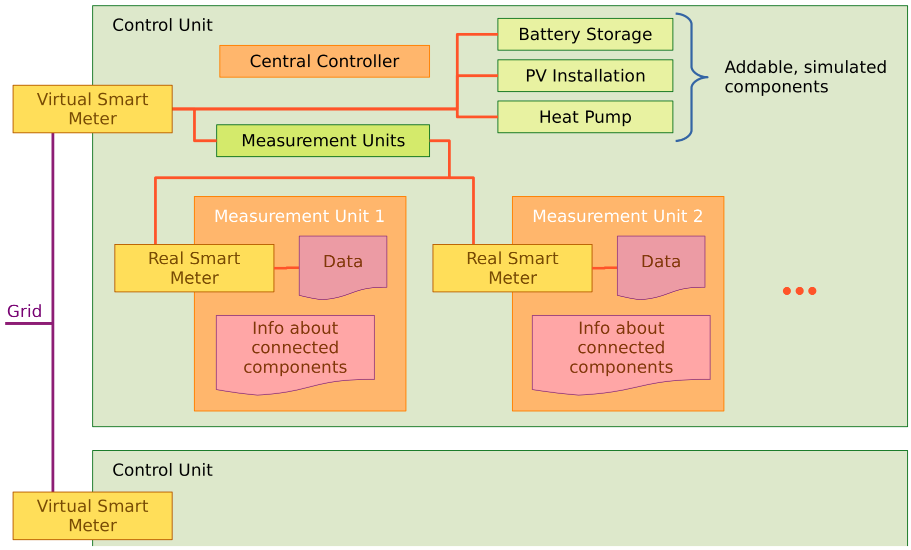

Fig. 2 Concept of a virtual smart meter: Distinction of measurement units as smart meter data input and control units that combine all smart meters in one house and potentially add simulated components. Red lines depict control unit internal energy flows. Crimson lines represent the local energy system

Table 2 List of frequently used sets and variables

Set name	Set description
$\mathcal{C}$	Set of all control units
$\mathcal{M}_c$	Set of all measurement units per control unit *c*
$\mathcal{S}$	Set of all substations
$\mathcal{T}$	Set of all time steps
Variable name	**Variable description**
$P_{c,j}^{rSM}(t)$	Real smart meter reading at time step *t* for control unit *c* and measurement unit *j*
$P_c^{vSM}(t)$	Virtual smart meter reading at time step *t* for control unit *c*
$P_c^{HP}(t)$	Power of the heat pump at time step *t* for control unit *c*
$P_c^{PV}(t)$	Power generation of the PV installation at time step *t* for control unit *c*
$P_{act,c}^{Bat}(t)$	Power of the battery at time step *t* at control unit *c*
$E_c^{\text{self cons.}}(t)$	Self-consumed energy at time step *t* for control unit *c*

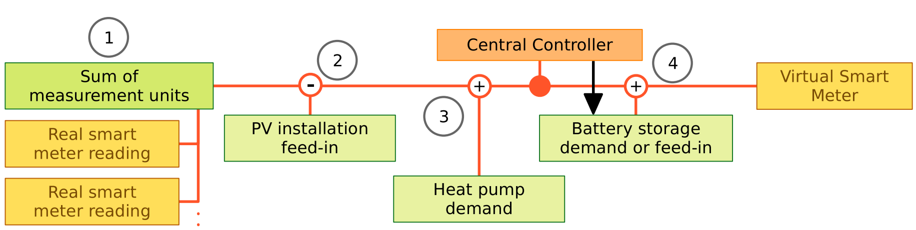

Fig. 3 Visualization of the virtual smart meter load computation and the ordering of component calculations inside of a control unit. The battery action is determined by the controller inside the unit

computations inside the simulated components will be presented later. First, we sum the load of all connected measurement units per control unit. Thereupon, we compute the power generation $P_c^{PV}(t)$ of the connected PV installation. If no PV installation is simulated, we set $P_c^{PV}(t) = 0$.

Next, we compute the electricity demand of the heat pump $P_c^{HP}(t)$ or set $P_c^{HP}(t) = 0$ if no heat pump is simulated. Finally, the demand or feed-in $P_{act,c}^{Bat}(t)$ of the battery storage system is computed if present. The concrete action of the battery storage system is given by a control strategy, for example a rule-based one maximizing PV self-consumption, inside the control unit. So, the power at the virtual smart meter of a control unit c at time step t is given by

$$P_c^{vSM}(t) = \sum_{j \in \mathcal{M}_c} P_{c,j}^{rSM}(t) - P_c^{PV}(t) + P_c^{HP}(t) + P_{act,c}^{Bat}(t) \tag{1}$$

where $\mathcal{M}_c$ indicates the set of all measurement units connected to control unit c. Please note, that the value of $P_{act,c}^{Bat}(t)$ is negative while discharging.

Furthermore, the power $P_c^{\text{self cons.}}(t)$, that is produced and also consumed on site c, is computed following the idea of Luthander et al. (2015). Per time step $P_c^{\text{self cons.}}(t)$ is the minimum of the power of first all local consumers and second the PV installation including the battery, formally

$$P_c^{\text{self cons.}}(t) = \min\left\{ \underbrace{\Psi_c(t) + P_c^{\text{HP}}(t)}_{\text{Local demand}},\ \underbrace{P_c^{\text{PV}}(t) - P_{act,c}^{\text{Bat}}(t)}_{\text{Local production}} \right\} \tag{2}$$

where $\Psi_c(t) = \sum_{j \in \mathcal{M}_c} P_{c,j}^{rSM}(t)$ is the sum of all real smart meter readings at time step t for control unit c. Please note, that $P_c^{\text{PV}}(t) - P_{act,c}^{\text{Bat}}(t) \geq 0$ holds always following the above arguments. The directly consumed energy per time step is the power multiplied with the time step duration, i.e. $E_c^{\text{self cons.}}(t) = P_c^{\text{self cons.}}(t) \cdot \triangle t$

After the last time step has been computed, metrics such as SSR (see Eq. 5) are calculated on control unit level. The simulation outputs detailed time series on the load of all substations, virtual smart meters and the simulated components. In the following subsections, we describe how the components are modeled and how control units are selected for the simulated addition of components.

Data concept of the digital twin

Our modeled digital twin requires multiple data sources. Besides the smart meter time series, the model requires the address of the meters to combine multiple meters installed at the same address to one control unit. Moreover, information about existing roof sections including their orientation has to be given per address. This roof data is taken from cadastral data or 3D geographical data in the CityGML format with a Level of Detail (LoD) of two. Further information about CityGML and the possible LoD can be found in Löwner et al. (2013). Additionally, the digital twin requires the knowledge which components are measured by a smart meter. This information should be available to the system operator. Otherwise, we can conclude the existence

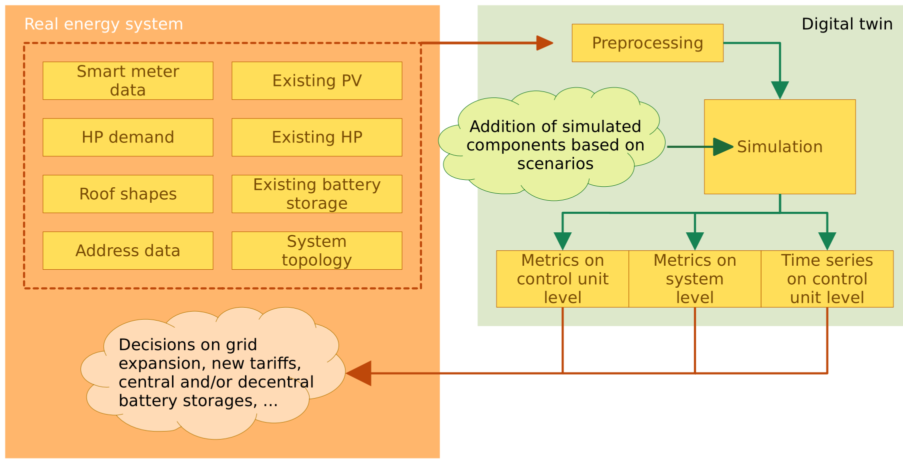

Fig. 4 Required data used by the digital twin and conceptual data flow including the input and output for the simulation

of components from the smart meter time series. For example, Weigert et al. (2020) present an approach to detect heat pumps in smart meter data, and Neubert et al. (2022) detect private charging stations and PV installations. Finally, the estimated annual electricity demand of a heat pump is required per building. The data is processed inside the digital twin. Details on preprocessing are presented in the next section. All required data is depicted schematically in Fig. 4.

Simulation of the individual components

PV installations

We use existing feed-in profiles for simulating the feed-in of a residential, rooftop PV installation. The profiles are taken from the meters where we know they exclusively measure a rooftop PV installation. The metered data is turned into a profile by dividing every element of the time series through the nominal power of the installation. The orientation of the measured installation is taken from the roof data. The advantage of real measured PV is that special effects, such as snow covering the PV installation in winter, or cloud cover reducing PV feed-in, are already included in the profiles without simulating the PV feed-in from historic sun radiation profiles. Literature like Klonari et al. (2016) already exists where PV profiles are generated from smart meter data. The PV installation consists of one or more subcomponents, each representing one roof section where a covering with PV modules is simulated. Information about existing roof sections and their orientation must be given in the data.

For computing the maximum installable power per control unit c, we iterate over all roof sections attributed to c taking the area of every section. Roof windows and dormers reduce the available area of a roof section for installing PV panels. Moreover, roofs have different shapes that rectangular PV modules cannot fully cover. Therefore, we introduce a hyper-parameter r measured in kWp/m^2 saying which nominal PV power can be installed per roof area. The nominal power $P^N_{c,i}$ per roof section i per control unit $c \in \mathcal{C}$ with area $A_{c,i}$ is thus computed by $P^N_{c,i} = A_{c,i} \cdot r$. Consequently, the nominal power of the entire PV installation is the sum of all nominal powers of the subsections. Generally, we can exclude small roof sections or sections with an inappropriate orientation, like north roofs in Europe.

During the simulation run at time step t, the PV component computes the current generation per roof section as the product of the profile with the correct orientation $S_{c,i}$ and the nominal power of the section. The sum over all sections i per control unit c

$$P_c^{PV}(t) = \sum_i S_{c,i}(t) \cdot P_{c,i}^N = \sum_i S_{c,i}(t) A_{c,i} \cdot r \tag{3}$$

gives us the total generation of the component for unit c.

Battery energy storage systems

The battery storage system is implemented as an input–output-model with the capacity E_{max} in kWh, the maximum available power P_{max} in kW and the initial state of charge as parameters. In such models, the charge level $E(t+1)$ of the battery at time step $t+1$ is calculated by adding the energy fed into or taken from the battery per time step $E(t+1) = E(t) + P_{act,c}^{Bat}(t) \cdot \triangle t$. They are vividly used for simulating battery storage at a systematic level, e.g. in Nyholm et al. (2016) or Martins et al. (2016). Some authors also refer to these models as so-called simple storage models, like Lyden et al. (2018). During the simulation, the controller of the control unit can set a charging (or discharging) request P_{req}. If $P_{req} > 0$, the battery should be charged, if $P_{req} < 0$ it should be discharged. An internal logic inside the battery model checks if this request can be fulfilled. That is the case if there is enough free or available capacity and the requested power does not exceed the maximum power. If the request cannot be fulfilled, the internal logic calculates how much energy can be taken from the battery until it is empty or fed in until reaching the maximum. The flowchart in Fig. 5 illustrates this logic. A similar idea for deciding when to charge or discharge the battery can also be found in Li et al. (2019).

Heat pumps

The simulation requires a set of normalized heat pump profiles, at least one profile. The profile is expected to be normalized to an annual electricity consumption of 1000 kWh. Besides that, the expected annual energy demand of the heat pump must be included in the data. Each expanded control unit gets a randomly selected heat pump profile, which is scaled with its individual annual heat pump energy demand divided by 1000 kWh. An advantage of simulating the heat pump using real measured profiles is that these profiles fit temporarily with the used PV feed-in profiles. Thus, weather conditions such as cold winter days with a thick blanket of snow consistently affect all profiles.

Remarks on component control strategies

Even though the battery storage system currently has a rule-based control strategy, the design is not limited to those. It is possible to merge existing cellular approaches like those presented by Dengler et al. (2022) into our model, where a control unit would represent a so-called *energy cell* following their wording.

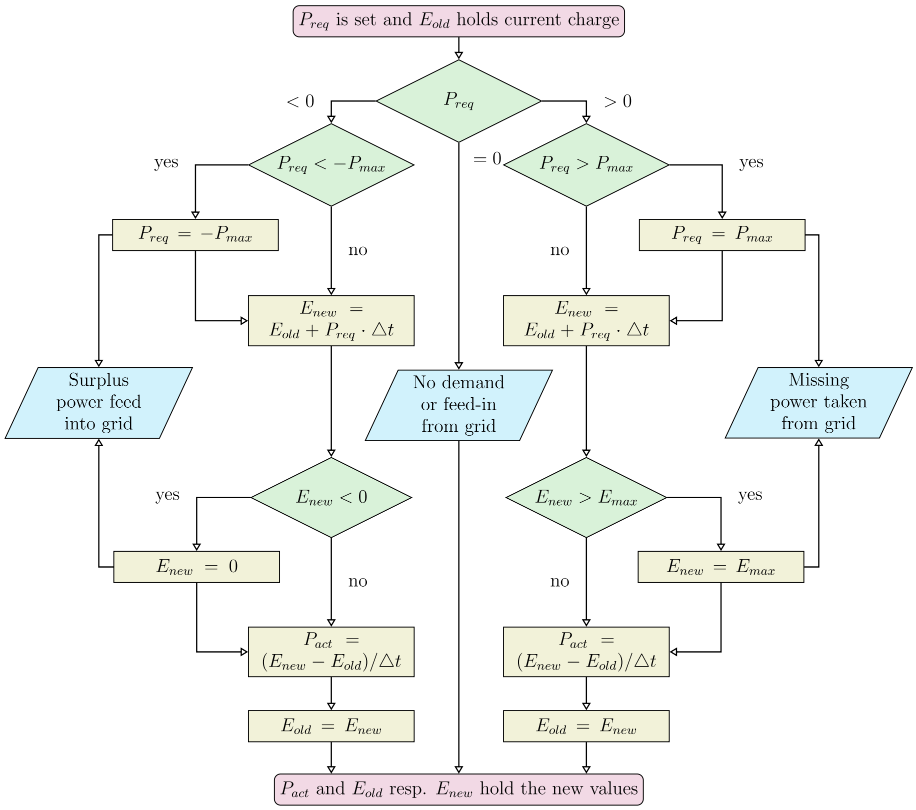

Fig. 5 Internal logic of the battery for charging or discharging. Δt is the step size between two time steps in hours. E_{old} is the current battery charge at the end of the last step, E_{new} the battery charge at the end of the step, P_{req} is the charge request and $P^{Bat}_{act,c}$ the actual power that can be delivered from or fed into the battery. Above, P_{act} is also noted as $P^{Bat}_{act,c}(t)$, but for the sake of clarity indices are not shown in this figure

Computed metrics

To measure the effectiveness of a PV installation, we use two measures, the Self-Consumption Rate (SCR) and the Self-Sufficiency Rate (SSR) following the definition of Steber (2018). The SCR of an arbitrary control unit c is defined as the quotient of the total self-consumed energy of c and the total PV generation over the period considered, formally noted following the definitions of Eq. 2:

$$\mathrm{SCR}(c) = \frac{\sum\limits_{t \in \mathcal{T}} E_c^{\text{self cons.}}(t)}{\sum\limits_{t \in \mathcal{T}} P_c^{PV}(t) \cdot \triangle t} \tag{4}$$

Analogously, the SSR of an arbitrary control unit c is defined as the quotient of the total self-consumed energy of c and the total consumption of c over the period considered:

$$\mathrm{SSR}(c) = \frac{\sum\limits_{t \in \mathcal{T}} E_c^{\text{self cons.}}(t)}{\sum\limits_{t \in \mathcal{T}} \left(\Psi_c(t) + P_c^{HP}(t) \right) \cdot \triangle t} \tag{5}$$

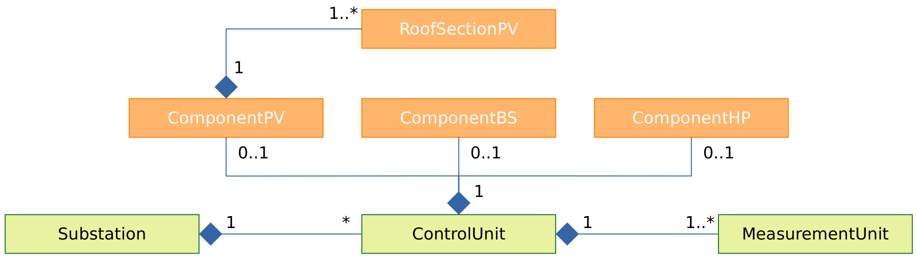

Fig. 6 The UML class diagram of the implementation with the cardinalities given. Orange boxes depict simulated components, green boxes depict units representing real objects

Planning of the addition of simulated components

As illustrated in Fig. 4, we have to define a scenario of how simulated components should be added and where this should happen. We notice, that all considered components (i.e. residential PV installations, battery storage systems and heat pumps) might already exist in any combination in existing buildings. For the further course, we denote the set of all possible component combinations as $\mathcal{Q} \subset \mathbb{P}(\{PV, BS, HP\})$ assuming a fixed order of the combinations. Now, we define a matrix M_{exp}, where each row denotes the existing component combinations and each column represents the target combinations. In each cell (i, j) we can define the portion of control units with current combination i that get extended to combination j. The extension happens by adding the missing components. We note, that some matrix combinations are impossible, as components would get removed. These combinations have to be zero. The absolute number of control units expanded from combination i to j, denoted as $N_{exp}(i,j)$, is the product of the $M_{exp}(i,j)$ and the total number of control units having combination i in reality.

Control unit selection modes

For selecting concrete instances, several selection modes are possible. For instance, a simple approach could be a random selection of the control units.

Some publications cited in Related work section like Schopfer et al. (2018) report that the addition of new PV installations is more profitable for some households than for others. Thus, we present a more sophisticated mode of selecting the control units for new component addition using the SSR defined in Eq. 5 as a metric. In this mode, we expand all units with all possible combinations and run the simulation for the complete time span, noting the results of the metric per unit and expansion combination. Finally, we select the $N_{exp}(i,j)$ control units per $(i,j)^{\text{th}}$ combination achieving the best value of the metric. We call this mode *Best-SSR-Selection-Mode*.

Implementation and verification of the simulation

We use an object-oriented design, where all components and units are designed as an individual class. This is a common approach that is also used by Bazan and German (2012). The class diagram in UML notation is illustrated in Fig. 6.

To verify our simulation implementation, we use a very small, hand-crafted data set with artificially generated data for the period of one day. To do so, we define three sequential scenarios for which we verified the output by hand. The first one is a scenario without any expansion. The test can be accepted as the input equals the output. The second one is a scenario where we equip the control units with a PV installation. The last scenario finally includes all active components, i.e., we equip the control units not only with PV installations but also with a battery storage system. All outputs are checked via hand, which is still possible for 24 time steps. On top of that, we verified the results of an exemplary case during an expert call with our project partner.

As the digital twin simulates new PV installations and heat pumps based on existing profiles, we have to check the plausibility of the profiles before using them. Since normalized PV feed-in time series are expected to show a temporarily similar profile, we compare the time series of all PV installations with the same orientation among each other using mean squared error and the Pearson correlation coefficient. Profiles showing a low correlation coefficient or a mean squared error high above the mean will be excluded. Heat pump profiles may diverge in their temporal patterns, thus we compute the energy consumed per month and remove all series that do not show a typical seasonal profile.

Details on preprocessing

In this section, we define the way the data is preprocessed by the digital twin to pass them into the simulation. Before processing the existing smart meter data, the start, the end and the temporal resolution of the recorded data are noted. With this information, the preprocessing of the digital twin creates a list of all time steps for which data is available and numbers them consecutively. The temporal alignment of the smart meter data is checked against this list of time steps.

This means that missing values are interpolated, and duplicated entries are removed. Later during the simulation, the load value of a measurement unit can easily be read from these files with the help of the time steps. The simulation requires additional information on which measurement units are at one place and thus form a control unit. Furthermore, the preprocessing also requires information about existing PV installations, battery storage systems and heat pumps per measurement unit to assign a value to $R_E(m)$ per measurement unit m. Based on this information, the preprocessing of the digital twin creates a list $\mathcal{S}$ of all substations, a list $\mathcal{C}$ of all control units and a list $\mathcal{M}_c$ of all existing measurement units per control unit c. It also outputs the interconnections between these components. Subsequently, all existing PV installations and heat pumps measured by an individual smart meter exclusively are used to generate a set of normalized PV feed-in profiles and heat pump profiles. Finally, we compute the difference between the recorded smart meter data and the real measured system load per time step. A difference can appear as it could happen that some meters are not smart meters and we thus lack this data, but for the calculation of the total system load it is still of interest. The whole preprocessing procedure is illustrated in Fig. 7. Except for the large-sized smart meter time series, the preprocessing stores the output in one single place, an SQL database. This idea can be found in the literature, for example, in Karnouskos and de Holanda (2009). We, therefore, use SQLite (Kreibich 2010).

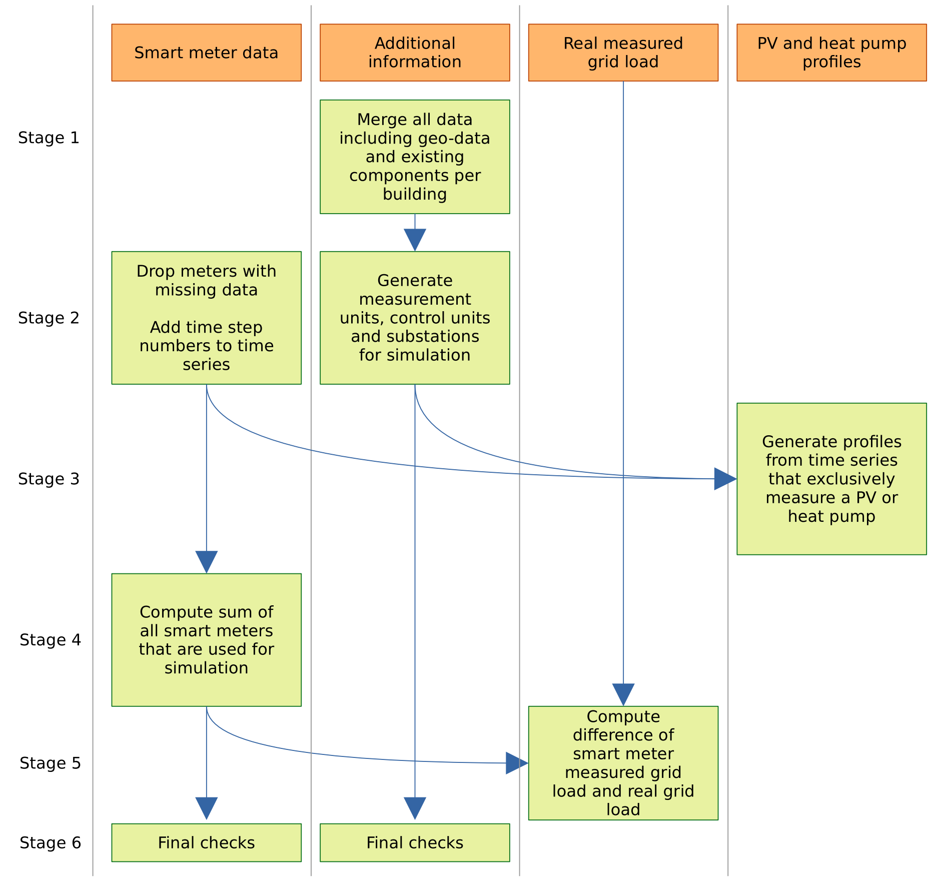

Fig. 7 Overview of the preprocessing pipeline. The orange boxes in the top line depict the different classes of input or output data - i.e., smart meter data from residential households, additional information about addresses, connected components, roof areas etc. of these households, the real measured grid load and PV or heat pump profiles. Green boxes represent processing steps

Case study

In this case study, we present the capabilities of the above-defined digital twin. We, therefore, take real data from our project partner, which is fed into the digital twin. The case study investigates two exemplary aspects of our digital twin: First, we investigate how the selection of the control units for adding a PV installation works. Therefore, we study which control units are expanded with a simulated PV or PV-battery-combination using different computation modes for the control unit expansion selection. In this context, we also demonstrate the value of geospatial data. Second, we analyze the effect of the added PV(-battery)-installations and heat pumps on the total grid load of the energy system.

Used smart-meter data and preprocessing

We acquired real measured smart meter data for all residential households from Haßfurt and Theres, two small towns located next to each other in northern Bavaria, Germany. This unique data set contains the data of over 8000 smart meters over a period of three years, i.e., 2019 until 2021, with a time resolution of one hour, and including the system topology. Thus, the time step size for the simulation is one hour. Moreover, we have

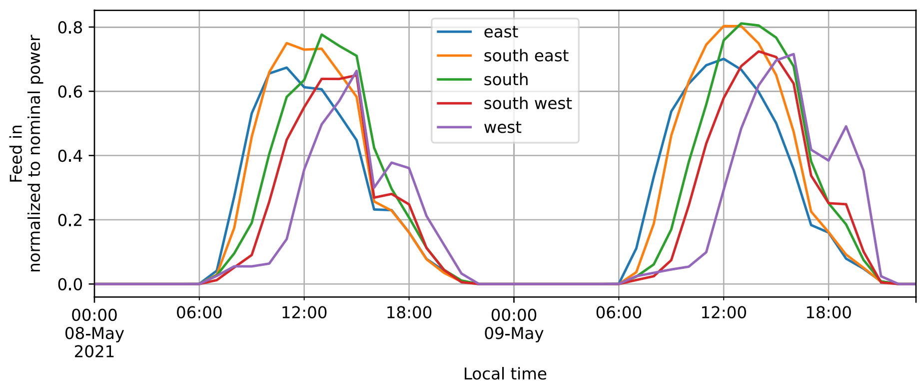

Fig. 8 Plot of exemplary feed-in time series per roof orientation for two days in May 2021

all information as illustrated in Fig. 4. The data about the existing properties, buildings and roof sections are delivered from the Bavarian state land surveying office with a LoD of 2 (Aringer and Hümmer 2011). During preprocessing, we remove all control units at an unknown location or a location without roof information. Removing households with unknown locations or missing data points is a common approach that has already been done, for example, by Nyholm et al. (2016). Finally, we have 3903 control units each located at a different location (i.e., address). Almost all considered locations, exactly 3642, have at least one residential building. 441 control units already possess a PV installation without battery storage, and 86 units have both components.

Profile filtration

As defined in Section Simulation of the individual components, we search all measured time series for exclusively measured heat pumps or PV installations for generating PV and heat pump profiles. The data shows 49 meters that exclusively measure a PV installation. For all of those, we take a look at a satellite photo given the address of the meter to identify the orientation of the PV installation using two different sources, Bing Maps (2022) and Google Maps (2022). In the end, we obtain 15 meters where we can determine the orientation. The real measured time series are normalized by dividing all measured values by the nominal power of the installation to obtain proper profiles. Figure 8 depicts the exemplary feed-in time series per orientation for two consecutive days in May 2021. We can see that the different roof section orientations show their peak accordingly to the solar trajectory for both days and that the peak value changes for different orientations.

Scenario definition

In order to present the capabilities of the digital twin we define the following five scenarios. Our scenarios aim to reflect the expected expansion state at the end of 2029.

Progress of PV expansion in Haßfurt und Theres up to now and future plans

We set the nominal power installable per roof area r (see Eq. 3) to the median obtained from the data, which is 66.2 Wp/m^2. For the sake of clarity, we will not evaluate different values. The interested reader may conduct Wiginton et al. (2010) or Hopf et al. (2017) for a more detailed discussion of the usable area.

We furthermore need to make an assumption about the expansion state at the end of 2029. In 2021, new rooftop PV installations with a total power of 1.2 MWp were installed resulting in a total nominal power of 9 MWp. Assuming that the 2021 addition rate will continue, by the end of 2029 we will have nearly 19 MWp of installed rooftop PV power. Between 2013 and 2020, many newly installed rooftop PV installations had a power lower or equal to 10 kWp. The reason for this past development is a massive tax benefit for residential installations below 10 kWp, as highlighted in Truong et al. (2016). From the beginning of 2023, the limit will be raised to 30 kWp (Federal Gazette of Germany 2022). Our projected target of 19 MWp installed rooftop PV capacity goes along with the plans of the Federal Network Agency of Germany (2022).

Scenarios in detail

Scenario A represents the addition of PV power from 2021 linearly extrapolated to 2029. So we simulatively add a PV installation in combination with a battery storage system to 45% of all control units that neither have a PV installation nor a heat pump. The control units used for expansion are selected using the *Best-SSR-Selection-Mode* as defined in Section Planning of the addition of simulated components. Per simulatively added PV installation we limit the added power upwards to 10 kWp, even though PV installations with more power could be installed on a given roof, representing the situation in 2021.

Scenario B represents the addition of PV power from 2021 linearly extrapolated to 2029. However, in contrast to Scenario A, we increase the maximal power per PV installation to 30 kWp which seems to be better for estimating the future development in Germany as noted above. Still, we limit the addition of PV systems to 10 MWp in order to achieve the target of an installed PV power of 19 MWp in 2029. This leads to a reduction of expanded control units to 908 (about 33% expansion).

Scenario C equals Scenario B, but we use no battery storage system. It is used to answer the first question of the case study, i.e., to investigate how the choice of control units selected for addition changes when only PV systems are added instead of PV-storage combinations.

Scenario D expands Scenario A by adding additional heat pumps to every control unit which gets an added PV installation. The total annual electricity consumption of all heat pumps is 7.1 GWh. Following an idea presented by Edmunds et al. (2021), we estimate the annual consumption for a simulated heat pump based on the building volume using existing annual consumption values as ground truth.

Scenario E equals Scenario A except that we replace the roof data to understand the value of the geospatial data. This means that all installed PV systems face south with a power of 10 kWp. The added PV power is limited to 10 MWp, this equals the added PV power in Scenario A.

Table 3 Overview of the defined scenarios and some results: the number of Control Units (CUs) with added components and the total sum of added PV power

Scenario	Max. kWp per installation	battery added	heat pumps added	Number of expanded CUs	Total added PV power
Scenario A	10 kWp	yes	no	1 313	10 007 kWp
Scenario B	30 kWp	yes	no	908	10 009 kWp
Scenario C	30 kWp	no	no	632	10 001 kWp
Scenario D	10 kWp	yes	yes	1 185	10 000 kWp
Scenario E	10 kWp	yes	no	1 000	10 000 kWp

Settings for all scenarios

For all of the above-defined scenarios, we use the following parameters for the PV installation and the battery:

PV installations The size of the installation is determined by the individual roof conditions but might be limited to a maximal power depending on the scenario definition. Roof sections where less than 2 kWp can be installed, are ignored.

Battery storage system For ease of comparison, we have set the capacity for all simulatively added battery storage systems to a fixed value of 7.5 kWh and use a C-rate of 1.

An overview of the defined scenarios can be found in Table 3. It also contains the absolute number of expanded control units per scenario.

Metrics

To measure the first investigation goal of this case study, we inspect the installed PV power and the mean annual energy consumption of an individual control unit in combination with the gain of the PV installation measured by using the above-defined SSR (see Eq. 5). For the second investigation goal, the metric to investigate is the grid load.

Results

The last two columns of Table 3 show how many control units are expanded by a PV-(battery)-combination. In the further course, we compare Scenario A, B and C in terms of the selection of control units for PV addition. We also compare Scenario A and E to highlight the value of geospatial information. For the analysis of the grid load, we use all defined scenarios.

Selection of control units

The number of expanded control units declines in Scenario B and C compared to Scenario A, because the upper power limit for the newly created PV installations is much higher than in Scenario A. Moreover, a statistical analysis of the total power of the added PV units per control unit reveals that many roof topologies allow larger PV installations than only 10 kWp in total, which are cut off in Scenario A. Whereas the mean power over all added PV installations is 7.62 kWp in Scenario A, this value rises up to

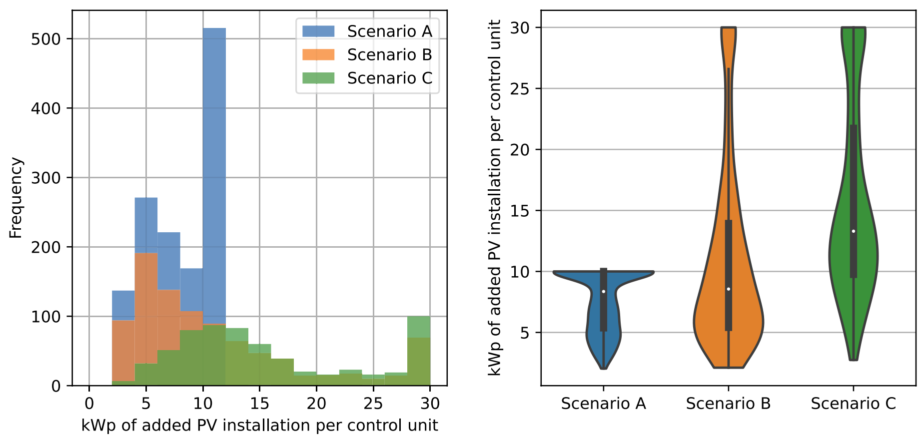

Fig. 9 Visualization of the distribution of simulatively added PV installations grouped by the defined scenarios. Left: The distributions visualized in a histogram. Right: Violin plots of the distributions. Scenario D is not evaluated in this context and Scenario E has a fixed PV power of 10 kWp

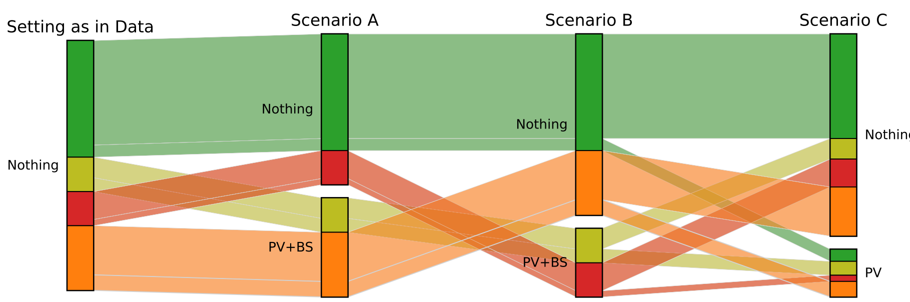

Fig. 10 Change in expansion group membership (i.e. if expansion took place using a PV- / PV-battery-installation or not) for the different evaluated scenarios. The different trajectories are colored differently. We can see that the selected control units for adding components changes significantly depending on the scenario. Scenario D and Scenario E are not shown for the sake of clarity

11.02 kWp for Scenario B. The added PV peak power distributions of Scenario A and B are very similar with values between 0 kWp and 9 kWp. For higher values, we can see that the modus of this distribution for Scenario A is at the 0.75-quantile, which is at a level of 10 kWp. A histogram and a violin plot of the added peak power distributions for Scenario A, B and C is depicted in Fig. 9. A comparison between Scenario A and E is given in the next subsection.

When comparing the control units selected between Scenario A and Scenario B, 456 equal units get expanded in both scenarios. Whereas 452 units are expanded in Scenario B but not in A, and 857 are expanded in Scenario A but not in B. As the simulation selects those units with best the SSR, we can conclude that allowing bigger PV installations has a huge impact on which units are selected for expansion and which are not. We see the same situation when comparing Scenario B and C. While only 266 control units are expanded in B and C, 642 units are only expanded in B but not in C. Moreover, 366 units are expanded in Scenario C but not in B. Figure 10 shows a parallel plot of the combination of how control units are expanded across the different evaluated scenarios.

Value of geospatial data

When replacing information about roof shapes and orientations so that they all face south with a capacity of 10 kWp, we first notice that the number of expanded control units reduces from 1313 units in Scenario A to 1000 units in Scenario E. Moreover, 884 units get a simulated PV installation in Scenario A that are not selected in Scenario E. Vice versa, 490 units get a simulated PV installation in Scenario E, but not in Scenario A. This means that roughly half of the units selected for expansion would not be selected if roof information was available. In addition, the number of PV expanded units required to meet a specified expansion target is underestimated by 24% in our case.

Effects on grid load

As we have data for three complete years, we analyze the result on grid load level per year that is assumed to be projected in 2029. For clarity, we will only use data from 2021 as a baseline scenario for analyzing the effect of more PV penetration on grid load. The annual energy import from the upper system level is 22.0 GWh in 2021, while the annual feed-in to the upper system level is 26.0 GWh in 2021.

In Scenario A, the annual import is reduced by 3.1 GWh or 14% to 19.0 GWh. In Scenario B we have an annual import of 19.1 GWh (a reduction of 13%) and in Scenario C it is 19.2 GWh (a reduction of 12.5%). The feed-in to the upper grid level increases by 6.2 GWh or 24% to 32.2 GWh in Scenario A and B or 6.3 GWh in Scenario C. Summarizing all scenarios, we see a reduction of the annual electricity import by about 13% in our case study, where the installed residential rooftop PV power is approximately doubled. The feed-in to the upper grid level increase by 24%. Between Scenario A, B and C we cannot identify significant differences over the complete time span. We shall notice, that the feed-in to the upper grid level as present in the baseline is mainly driven by local wind farms and big-sized, open-space PV installations. Thus, comparing our results with other publications should be based on the absolute changes with respect to the added total PV capacity.

There are 5040 hours in the baseline where the local energy system requires import from the upper system. The hours with import decrease to 4459 hours or 88% in Scenario A, or 4435 and 4399 in Scenario B and Scenario C, respectively. The inverse situation occurs for hours with feed-in to upper grid levels, where the hours with feed-in increase from 3719 hours in baseline to 4300 or 16% more in Scenario A. Even though the sum of electricity feed-in to the upper grid level increases by 24%, the number of hours where this happens only increases by 16%. This indicates that feed-in to the upper level of the grid is slightly more frequent and, more importantly, stronger. Or in other words, if there is a feed-in to the upper grid level already in the baseline, this will be more intense in future grid states. But the times when the import is replaced by export is much less increased. An overview of the import from and export to the upper system level for the different scenarios can be found in Fig. 11. Moreover, in Scenario C, with no added batteries, the maximum import cannot be reduced compared to the baseline, as the maximum import appears in the evening on March 23. In Scenario A, the maximum import happens at the same time, but the peak can be reduced by 1.0%. We can conclude that system states with more residential PV-battery-installations do not help reduce import peaks of the complete local system.

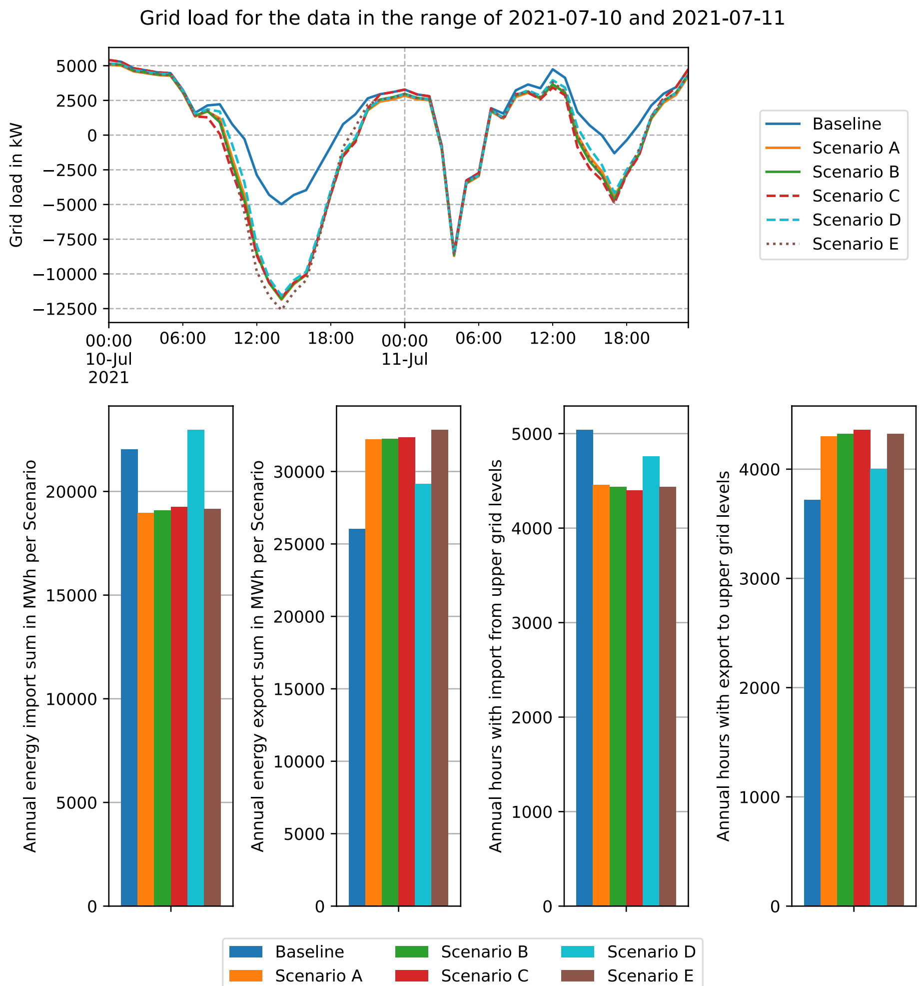

Fig. 11 Analytical plots of the total grid load for different scenarios. Top: Grid load in hourly resolution for two exemplary days with data taken between 10th until 12th of July 2021. Bottom: The four graphs show the annual electricity import sum from and export (i.e. feed-in) sum to the upper system level in MWh besides the hours with import from or feed-in to the upper system level per Scenario

Another feature of the simulation is the fact that the grid load can be investigated on an hourly level. For example, Fig. 11 shows the grid load for different scenarios using the data from two days in July 2021. The effect of the PV feed-in is visible. At a closer look, we can detect that the grid load in Scenario A and B is slightly higher compared to Scenario C during the morning hours (i.e. between 9:00 and 12:00) and slightly higher in the late evening (i.e., between 20:00 and 02:00). This effect can be attributed to the residential battery storage systems, that are not added in Scenario C.

Effect of heat pumps on grid load

For analyzing the effect of heat pumps on the grid load, we compare Scenario D with our baseline. On an annual level, the imported energy rises by 4% compared to the baseline or even 21% compared to Scenario A (which equals Scenario D except for the presence of heat pumps). At the same time, the annual energy feed-in to the upper grid level

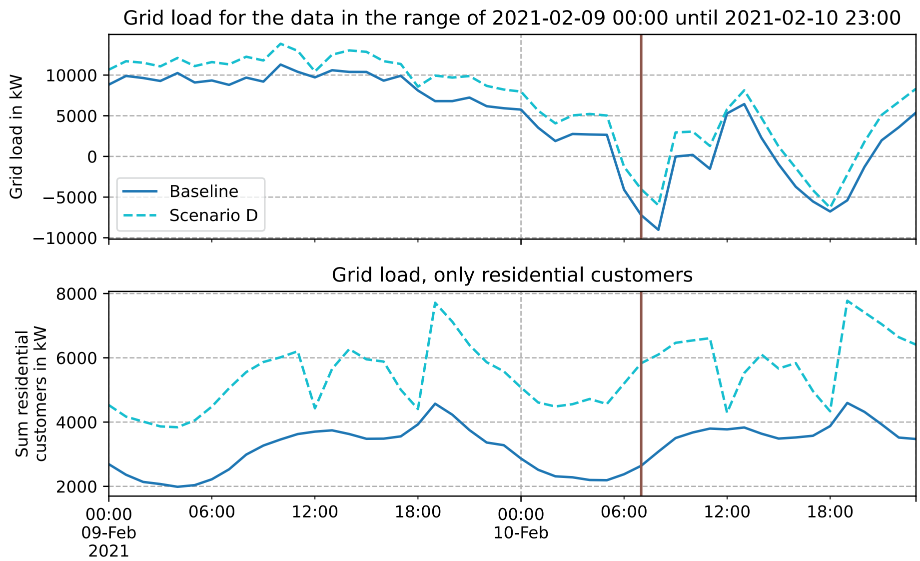

Fig. 12 Effect of the heat pumps (Scenario D) on grid load for two exemplary days in winter of 2021. The vertical line is placed at 9th February 2021, 07:00. Top: Gird load in hourly resolution. Bottom: Grid load without wind farms and large enterprises, i.e. only residential buildings

increases by 12% compared to the baseline. The hours with energy import decrease by 6% compared to the baseline, while the hours with energy export increase by 8%.

As noted above, wind farms have a notable impact on grid load. Based on the results, we can analyze this effect using data from 9th February 2021. In the morning, at 7 am, we see a feed-in to the upper grid level of 7.2 MW, even though the individual households consume 2.6 MW. In Scenario D, the consumption of the households will rise by 121% to 5.8 MW due to the added heat pumps. Individual PV installations or battery storage systems cannot buffer this demand as they are neither producing nor storing energy on this winter morning. This situation is depicted in Fig. 12.

Value of geospatial data

The missing geospatial data in Scenario E compared to Scenario A also affects the simulated grid load. While there is a feed-in to the upper grid level of 11.9 MW using the data of 10th July 2021 at 14:00 (Scenario A), the feed-in increases to 12.6 MW when ignoring the roof orientations (Scenario E) (see Fig. 11, dotted graph in the upper plot). This is an overestimation of approx. 7%. During the summer, this exemplary situation occurs often. The mean difference of the daily peak feed-in to the upper grid level for Scenario E compared to A is higher than 500 kW for 30 days between June and August.

Conclusion and outlook

In the past ten years, renewable energy sources have become increasingly popular due to the decarbonization of energy production. Private households can nowadays buy a PV installation and a battery storage systems at an affordable price. Besides that, there are increasingly more big energy consumers present, like heat pumps. Due to the ongoing development of the internet of things, a lot of data is available about the current production and demand of energy on household level that is recorded by smart meters.

In this paper, we built a novel digital twin of an existing local energy system based on smart meter data and supplementary building data, including roof shapes and heat demand information. We present an abstraction of the actual system containing control units and measurement units. The center of the twin forms a data-driven simulation of the system, where different expansion scenarios of local renewable energy sources (like residential PV / PV and battery installations) and higher penetration rates of big energy consumers like heat pumps can be evaluated. A highlight is the model-endogenous selection of control units for adding a PV(-battery)-installation based on metrics like the SSR.

We finally present a small case study to present some capabilities of our digital twin. There, we first investigate how many control units have to install a PV installation to achieve expansion targets where the units are ordered according to their self-sufficiency increase. Second, we analyze the effect of more PV and heat pump penetration on import from and export to upper system levels over a year. In the context of energy systems research, our case study justifies papers like Khan et al. (2019) which emphasize the importance of roof orientation for PV simulations. More importantly, we extend this information to the grid level, showing that there are days in summer when the peak feed-in to the upper grid level is overestimated by 7% if no geospatial data is available. In summary, geospatial data combined with smart meter data provides more detailed information on future energy systems, as if this data is missing, as Shahat et al. (2021) expected.

Up to now, electric mobility has yet to be considered. Since literature such as Strobel et al. (2022) clearly show that a large electric fleet significantly impacts peak load, especially on the local distribution level we are concerned with, the next step is integrating electric mobility into the digital twin. Moreover, advanced analysis will be performed to evaluate different control strategies for the battery storage system charging, especially those that include demand and PV production forecasting. Also, more economic metrics like the net present value will be integrated to decide where to add new simulated components.

Abbreviations

CU	Control unit
DNO	Distribution network operator
LoD	Level of detail
PV	Photovoltaic
SCR	Self-consumption rate
SOC	State of charge
SSR	Self-sufficiency rate

Acknowledgements
Not applicable.

Author contributions
DB: design and implementation of the digital twin, formal analysis, writing of this manuscript. MP: reviewing and commenting of this manuscript, funding acquistion. Both authors read and approved the final manuscript.

Funding
Open Access funding enabled and organized by Projekt DEAL. This paper is part of the funded project 'DigiSWM - Artificial intelligence and advanced data analytics for an interplay of power, heat, and mobility in the private sector' with funding reference number DIK-2103-0017//DIK0298/02 by the funding agency ***Bavarian Collaborative Research Program (BayVFP) of the Free State of Bavaria - funding line 'Digitization'***. This publication was supported by the Open Access Publication Fund of the University of Wuerzburg.

Availability of data and materials
Not applicable.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Received: 16 January 2023 Accepted: 5 March 2023
Published online: 28 March 2023

References

Akmal M, Fox B, Morrow JD, Littler T (2014) Impact of heat pump load on distribution networks. IET Gener Transm Distrib. 8(12):2065–73. https://doi.org/10.1049/iet-gtd.2014.0056

Ali S, Wintzek P, Zdrallek M, Monscheidt J, Gemsjäger B, Slupinski A (2022) Deduction of strategic planning guidelines for urban medium voltage grids with consideration of electromobility and heat pumps. Electricity 3(4):505–541. https://doi.org/10.3390/electricity3040026

Aringer K, Hümmer F (2011) Die dritte dimension im kataster–aufbau eines landesweiten gebäudemodells am beispiel bayerns. ZfV 4:210–218

Bazan P, German R (2012) Hybrid Simulation of Renewable Energy Generation and Storage Grids. In: Laroque C, Himmelspach J, Pasupathy R, Rose O, Uhrmacher AM, eds, Proceedings of the 2012 Winter Simulation Conference, Piscataway, NJ, USA. IEEE Press, pp 1–12https://doi.org/10.1109/WSC.2012.6465134

Bee Elena, Prada Alessandro, Baggio Paolo, Psimopoulos Emmanouil (2019) Air-source heat pump and photovoltaic systems for residential heating and cooling: potential of self-consumption in different european climates. Build Simul 12(3):453–463. https://doi.org/10.1007/s12273-018-0501-5

Blasi TM, Fernandes TS, Gevers DN, Aoki AR, Impinnisi PR, Pinto CL, Damasceno PM (2019) Evaluation of the impacts of renewables sources and battery systems in distribution feeders with different penetration levels. In: 2019 IEEE PES Innovative Smart Grid Technologies Conference - Latin America (ISGT Latin America), p 1–6. https://doi.org/10.1109/ISGT-LA.2019.8895349

Dengler G, Bazan P, German R (2022) Simulation of a cellular energy system including hierarchies and neighborhoods. Energy Inform. https://doi.org/10.1186/s42162-022-00243-2

Douglass P, Hou P, Martens S, Yang G (2019) Technical and economic impact of residential bess on distribution systems under alternative tariff regimes. In: Proceedings of the 25th International Conference on Electricity Distribution (CIRED 2019), number 942 in CIRED Conference Proceedings, pp 942–948. AIM. https://doi.org/10.34890/406

Edmunds C, Galloway S, Dixon J, Bukhsh W, Elders I (2021) Hosting capacity assessment of heat pumps and optimised electric vehicle charging on low voltage networks. Appl Energy 298:117093

Federal Gazette of Germany (2022) Annual Tax Act 2022. Federal Law Gazette of Germany 51.

Federal Network Agency of Germany (2022) Approval of the scenario framework 2023-2037/2045, July 2022. [Online] https://www.netzentwicklungsplan.de/. Accessed 13 Oct

Federal Statistical Office of Germany (Destatis) (2022) Press release No. 226 of 2 June 2022: More than half of the residential buildings constructed in 2021 are heated by heat pumps, 2022. [Online]. https://www.destatis.de/EN/Press/2022/06/PE22_226_31121.html. Accessed 2 Jun

Francisco A, Mohammadi N, Taylor JE (2020) Smart city digital twin-enabled energy management: toward real-time urban building energy benchmarking. J Manag Eng. https://doi.org/10.1061/(asce)me.1943-5479.0000741

Gharavi H, Ghafurian R (2011) Smart grid: the electric energy system of the future. Proc IEEE 99(6):917–921. https://doi.org/10.1109/jproc.2011.2124210

Google Maps (2022) The city of Hassfurt, Germany, 2022. [Online]. https://www.google.com/maps/@50.0374073,10.5192559,3142m. Accessed 26 July 2022

Grieves Michael W (2005) Product lifecycle management: the new paradigm for enterprises. Int J Prod Dev 2(1/2):71. https://doi.org/10.1504/ijpd.2005.006669

Haben Stephen, Singleton Colin, Grindrod Peter (2016) Analysis and clustering of residential customers energy behavioral demand using smart meter data. IEEE Trans Smart Grid 7(1):136–144. https://doi.org/10.1109/TSG.2015.2409786

Hämäläinen M (2020) Smart city development with digital twin technology. In: 33rd Bled eConference-Enabling Technology for a Sustainable Society: June 28–29, 2020, Online Conference Proceedings. University of Maribor

Hopf K, Kormann M, Sodenkamp M, Staake T (2017) A decision support system for photovoltaic potential estimation. In: Proceedings of the 1st International Conference on Internet of Things and Machine Learning, IML '17, New York, NY, USA. Association for Computing Machinery, pp 1–10, ISBN 9781450352437. https://doi.org/10.1145/3109761.3109764.

Iyengar S, Lee S, Irwin D, Shenoy P (2016) Analyzing energy usage on a city-scale using utility smart meters. In: Proceedings of the 3rd ACM International Conference on Systems for Energy-Efficient Built Environments. ACM, pp 51–60https://doi.org/10.1145/2993422.2993425

Jones D, Snider C, Nassehi A, Yon J, Hicks B (2020) Characterising the digital twin: a systematic literature review. CIRP J Manuf Sci Technol 29:36–52. https://doi.org/10.1016/j.cirpj.2020.02.002
Karnouskos S, De Holanda TN (2009) Simulation of a smart grid city with software agents. In: 2009 Third UKSim European Symposium on Computer Modeling and Simulation, IEEE, pp 424–429. https://doi.org/10.1109/EMS.2009.53
Kerai M (2022) Smart meter statistics in great britain: quarterly report to end december 2021. Technical report, Department for Business, Energy and Industrial Strategy, London, UK
Khan Zafar A, Jayaweera Dilan, Alvarez-Alvarado Manuel S (2018) A novel approach for load profiling in smart power grids using smart meter data. Electr Power Syst Res 165:191–198. https://doi.org/10.1016/j.epsr.2018.09.013. (**ISSN 0378-7796**)
Klonari V, Toubeau JF, LobryJ, VallÃƒÂ'Âe F (2016) Photovoltaic integration in smart city power distribution: a probabilistic photovoltaic hosting capacity assessment based on smart metering data. In: 2016 5th International Conference on Smart Cities and Green ICT Systems (SMARTGREENS), IEEE, pp 1–13
Kreibich J (2010) Using SQLite. O'Reilly Media, Inc., Sebastopol. ISBN 978-0-596-52118-9
Lakervi E, Holmes EJ (1995) Electricity distribution network design. In: Johns A, Platts J (eds) IET power and energy series, vol 21, 2nd ed. Institution of Engineering and Technology, London. https://doi.org/10.1049/PBPO021E
Lazzeroni P, Olivero S, Repetto M, Stirano F (2015) Impact of PV penetration in a distribution grid: a middle-east study case. In: 2015 IEEE 1st International Forum on Research and Technologies for Society and Industry Leveraging a better tomorrow (RTSI). IEEE. https://doi.org/10.1109/RTSI.2015.7325123
Li Hong Xian, Horan Peter, Luther Mark B, Ahmed Tarek M F (2019) Informed decision making of battery storage for solar-PV homes using smart meter data. Energy Build 198:491–502. https://doi.org/10.1016/j.enbuild.2019.06.036
Luthander Rasmus, Widén Joakim, Nilsson Daniel, Palm Jenny (2015) Photovoltaic self-consumption in buildings: a review. Appl Energy 142:80–94. https://doi.org/10.1016/j.apenergy.2014.12.028
Löwner MO, Benner J, Gröger G, Häfele KH (2013) New concepts for structuring 3d city models – an extended level of detail concept for CityGML buildings. In: Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp 466–480
Lyden Andrew, Pepper Russell, Tuohy Paul G (2018) A modelling tool selection process for planning of community scale energy systems including storage and demand side management. Sustain Cities Soc 39:674–688. https://doi.org/10.1016/j.scs.2018.02.003
Bing Maps (2022) The city of Hassfurt, Germany, 2022. [Online] https://www.bing.com/maps. Accessed 26 July 2022
Martins R, Musilek P, Hesse HC (2016) Optimization of photovoltaic power self-consumption using linear programming. In: 2016 IEEE 16th International Conference on Environment and Electrical Engineering (EEEIC), IEEE, pp 6–12. . https://doi.org/10.1109/eeeic.2016.7555581
McLoughlin F, Duffy A, Conlon M (2015) A clustering approach to domestic electricity load profile characterisation using smart metering data. Appl Energy 141:190–199
Miyasawa A, Akira S, Fujimoto Y, Hayashi Y (2021) Spatial demand forecasting based on smart meter data for improving local energy self-sufficiency in smart cities. IET Smart Cities 3(2):107–120. https://doi.org/10.1049/smc2.12011
Neubert M, Gnepper O, Mey O, Schneider A (2022) Detection of electric vehicles and photovoltaic systems in smart meter data. Energies 15(13):4922. https://doi.org/10.3390/en15134922
Nigmatulina N, Mashlakov A, Belonogova N, Honkapuro S (2020) Techno-economic impact of solar power system integration on a DSO. In: 2020 17th International Conference on the European Energy Market (EEM), pp 1–6. https://doi.org/10.1109/EEM49802.2020.9221951
Nyholm Emil, Goop Joel, Odenberger Mikael, Johnsson Filip (2016) Solar photovoltaic-battery systems in swedish households – self-consumption and self-sufficiency. Appl Energy 183:148–159. https://doi.org/10.1016/j.apenergy.2016.08.172
Onile Abiodun E, Machlev Ram, Petlenkov Eduard, Levron Yoash, Belikov Juri (2021) Uses of the digital twins concept for energy services, intelligent recommendation systems, and demand side management: A review. Energy Rep 7:997–1015. https://doi.org/10.1016/j.egyr.2021.01.090
Palm Alvar (2020) Early adopters and their motives: differences between earlier and later adopters of residential solar photovoltaics. Renew Sustain Energy Rev 133:110142
Petrova-Antonova D, Ilieva S (2020) Digital twin modeling of smart cities. In: Human Interaction, Emerging Technologies and Future Applications III, Springer International Publishing, pp 384–390
Pruckner M, Thurner C, Martin A, German R (2014) A coupled optimization and simulation model for the energy transition in Bavaria. In: Proceedings of the International Workshop on Demand Modeling and Quantitative Analysis of Future Generation Energy Networks and Energy Efficient Systems, Bamberg, University of Bamberg Press, pp 97–104. ISBN 978-3-86309-208-5
Ruohomaki T, Airaksinen E, Huuska P, KesÃƒÂ¤niemi O, Martikka M, Suomisto J (2018) Smart city platform enabling digital twin. In: 2018 International Conference on Intelligent Systems (IS), pages 155–161. IEEE, pp 155–161. https://doi.org/10.1109/IS.2018.8710517
Schopfer S, Tiefenbeck V, Staake T (2018) Economic assessment of photovoltaic battery systems based on household load profiles. Appl Energy 223:229–248
Shahat E, Hyun CT, Yeom C (2021) City digital twin potentials: a review and research agenda. Sustainability 13(6):3386. https://doi.org/10.3390/su13063386
Singh M, Fuenmayor E, Hinchy E, Qiao Y, Murray N, Devine D (2021) Digital twin: origin to future. Appl Syst Innov 4(2):36. https://doi.org/10.3390/asi4020036
Steber DB (2018) Integration of decentralized battery energy storage systems into the german electrical power system. PhD thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)
Steber D, Pruckner M, Bazan P, German R (2017) SWARM —providing 1 MW FCR power with residential PV-battery energy storage — simulation and empiric validation. In: 2017 IEEE Manchester PowerTech, Manchester, UK, IEEE; p 1–6. ISBN 978-1-5090-4238-8. https://doi.org/10.1109/PTC.2017.7981091
Sterchele P, Palzer A, Henning HM (2017) The role of heat pumps in the transformation of national energy systems-example germany. In: Proceedings of 12th IEA Heat Pump Conference

Strobel L, Schlund J, Pruckner M (2022) Joint analysis of regional and national power system impacts of electric vehicles-a case study for Germany on the county level in 2030. Appl Energy 315:118945

Truong C, Naumann M, Karl R, Müller M, Jossen A, Hess H (2016) Economics of residential photovoltaic battery systems in Germany: the case of tesla's powerwall. Batteries 2(2):14. https://doi.org/10.3390/batteries2020014

VanDerHorn E, Mahadevan S (2021) Digital twin: generalization, characterization and implementation. Decis Support Syst 145

von Appen J, Braslavsky JH, Ward JK, Braun M (2015) Sizing and grid impact of PV battery systems - a comparative analysis for australia and germany. In: 2015 International Symposium on Smart Electric Distribution Systems and Technologies (EDST). IEEE. https://doi.org/10.1109/SEDST.2015.7315280

Waite M, Modi V (2020) Electricity load implications of space heating decarbonization pathways. Joule 4(2):376–394. https://doi.org/10.1016/j.joule.2019.11.011

Wang Y, Chen Q, Hong T, Kang C (2019) Review of smart meter data analytics: applications, methodologies, and challenges. IEEE Trans Smart Grid 10(3):3125–3148. https://doi.org/10.1109/TSG.2018.2818167

Weigert A, Hopf K, Weinig N, Staake T (2020) Detection of heat pumps from smart meter and open data. Energy Inform. https://doi.org/10.1186/s42162-020-00124-6

Wiginton Lewis K, Nguyen Ha T, Pearce Joshua M (2010) Quantifying rooftop solar photovoltaic potential for regional renewable energy policy. Comput Environ Urban Syst 34(4):345–357. https://doi.org/10.1016/j.compenvurbsys.2010.01.001. (**ISSN 0198-9715**)

Yunus Khan TM, Soudagar ME, Kanchan M, Afzal A, Banapurmath NR, Akram N, Mane SD, Shahapurkar K (2019) Optimum location and influence of tilt angle on performance of solar PV panels. J Therm Anal Calorim. 141(1):511–532. https://doi.org/10.1007/s10973-019-09089-5

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.