Time Series Aggregation for ESMs

Subjects: Energy & Fuels

Contributor: Maximilian Hoffmann

Due to the high degree of intermittency of renewable energy sources and the growing interdependences amongst formerly separated energy pathways, the modeling of adequate energy systems is crucial to evaluate existing energy systems and to forecast viable future ones. However, this corresponds to the rising complexity of energy system models (ESMs) and often results in computationally intractable programs. To overcome this problem, time series aggregation is frequently used to reduce ESM complexity. As these methods aim at the reduction of input data and preserving the main information about the time series, but are not based on mathematically equivalent transformations, the performance of each method depends on the justifiability of its assumptions.

time series aggregation   complexity reduction   energy system model   temporal resolution

renewable energy systems   clustering   typical days   time slices   system states

snapshots

1. Introduction

Due to the climate change caused by anthropogenic CO_2 emissions resulting from the burning of fossil fuels, a major turnaround in the fields of energy supply and consumption is an increasing necessity. Key aspects of addressing this challenge are the integration of renewable energy sources (RES) into existing energy systems, as well as a closer coupling of energy forms and sectors [1].

The evolution of the energy sector has been accompanied by a consistent effort to model and predict its development. Early attempts to forecast future energy demands can be traced back to the 1950s and constitute simple, assumption-based scenarios [2]. Another theoretical foundation for modern energy system models (ESMs) is the principle of peak-load-pricing first described by Boiteux in 1949 [3] (English translation in 1960 [4]) and Steiner in 1957 [5]. This approach distinguishes between capacity and the operating costs of facilities producing non-storable goods. Thus, it applies to many simple energy systems with a single good (commodity), e.g., electricity which has led to the development of approaches to solve simple generation expansion planning (GEP) problems with the annual load duration curve as shown by Sherali et al. [6]. As key drivers for the early progress of energy system modelling, three factors are of particular significance:

- The need for security of supply to be provided for the growing demand by governmental quasi-monopolistic institutions throughout the 1970s, as well as the building of a reliable connection between a more competitive energy sector and later public interest [7][8].

- The progress of computational resources that enabled more complex models [9].

- The integration of non-dispatchable technologies, such as most of the RES and their impact on energy pricing; an effect that has constantly gained importance and which was first described in 1982 [10].

Ever since the first ESMs, based on optimizations rather than just simulations, were developed during the 1970s and 1980s [9][11], two major options arose for their developers, namely whether to focus on economic mechanisms, sometimes described as a top-down approach, or on the technical dimension, usually known as a bottom-up one [12][13]. Amongst the bottom-up frameworks, this review focuses on a vast number of approaches to modeling the different dimensions of energy systems, including methodologies such as optimizations and simulations [9][12][13]. With respect to the scope of ESMs, two fundamental dimensions can be delineated: namely spatiotemporal and techno-economic. The spatiotemporal dimension comprises the setting of input data that a model is intended to incorporate. The spatial sub-dimension is focused on the number of regions and their connections to each other, as energy systems on a national or even larger scale usually face the challenge of taking energy transmission between different regions into account. The temporal sub-dimension is divided into two aspects, namely temporal resolution (TR), often referred to as time steps, and the overall time horizon [9], which also concerns questions of storage modeling [14][15][16][17][18][19][20] as well as the linking of dynamic processes [21][22][23] and investment dynamics [24][25]. In contrast, the techno-economic dimension deals with how the components are represented in the model and whether their design and/or operation are optimized or, if their operational behavior is simply simulated, how they are mathematically represented and if the impact of supply and demand on energy prices is dynamically modeled or not [13]. Each of the dimensions listed above drives the overall complexity of ESMs, while the spatiotemporal resolution also affects the techno-economical dimension directly, e.g., the TR also limits the (technical) operational exactness of components in the energy system.

Figure 1 illustrates the classification into top-down and bottom-up models [13][21][26][27][28], top-down model types [29][30] and bottom-up model dimensions [13].
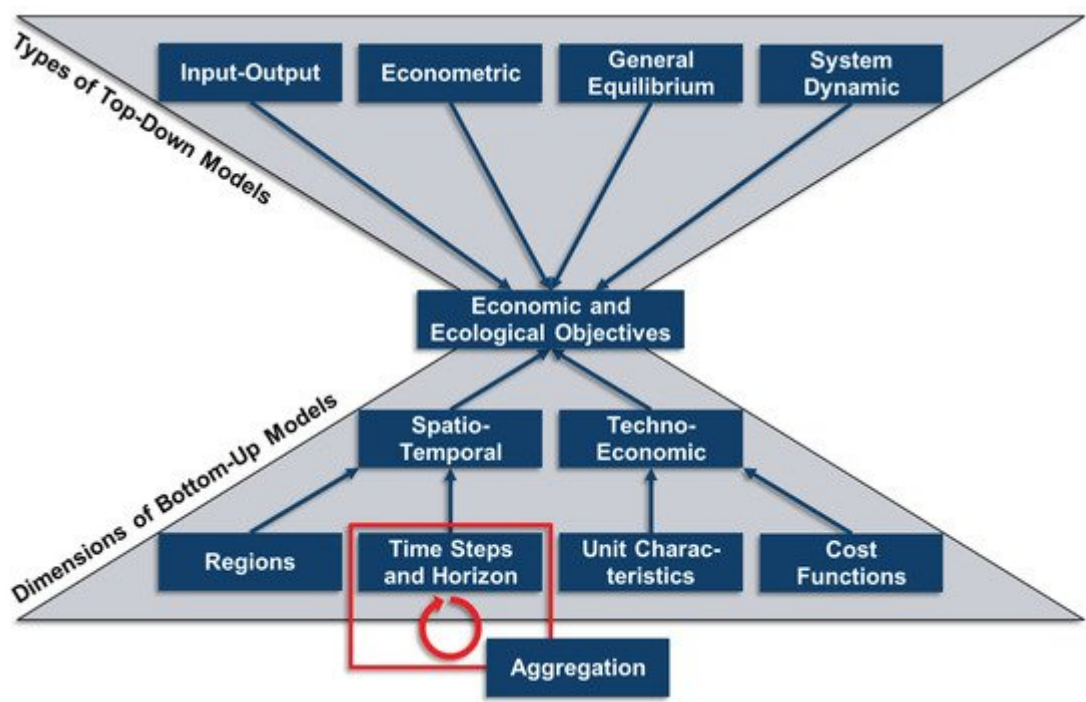
Figure 1. Classification of energy system models (ESMs), the sub-dimensions of bottom-up models and the scope of the review on time series aggregation (TSA).

2. Time Series Aggregation

The following section deals with the general concept of TSA. For the mathematical examinations of the following section, the nomenclature of Table 1 is used.

Table 1. Nomenclature for the mathematical examinations in the following section.

Sets & Indices	
Set	$\mathbb{D}$
Set/Index of Attributes	$\mathbb{A}, a$
Set/Index of Time Steps	$\mathbb{S}, s$
Set/Index of Periods	$\mathbb{P}, p$
Set/Index of Inner-Period Time Steps	$\mathbb{T}, t$
Set/Index of Cluster Members	$\mathbb{C}, k$
Cluster Center (as Defined)	c
Discrete Value of Time Series (Normed)	$x'(x)$
Number of Items in a Set	N
Matrix Path for Dynamic Time Warping	w
Minkowski Exponent	γ

The input data D usually consists of one time series for each attribute, i.e., D=A×S. The set of attributes A describes all types of parameters that are ex-ante known for the energy system, such as the capacity factors of certain technologies at certain locations or demands for heat and electricity that must be satisfied. The set of time

steps describes the shape of the time series itself, i.e., sets of discrete values that represent finite time intervals, e.g., 8760 time steps of hourly data to describe a year. For all methods presented in the following, it is crucial that the time series of all attributes have identical lengths and TR. The possible shape of this highly resolved input data is shown in the left upmost field in Figure 2.
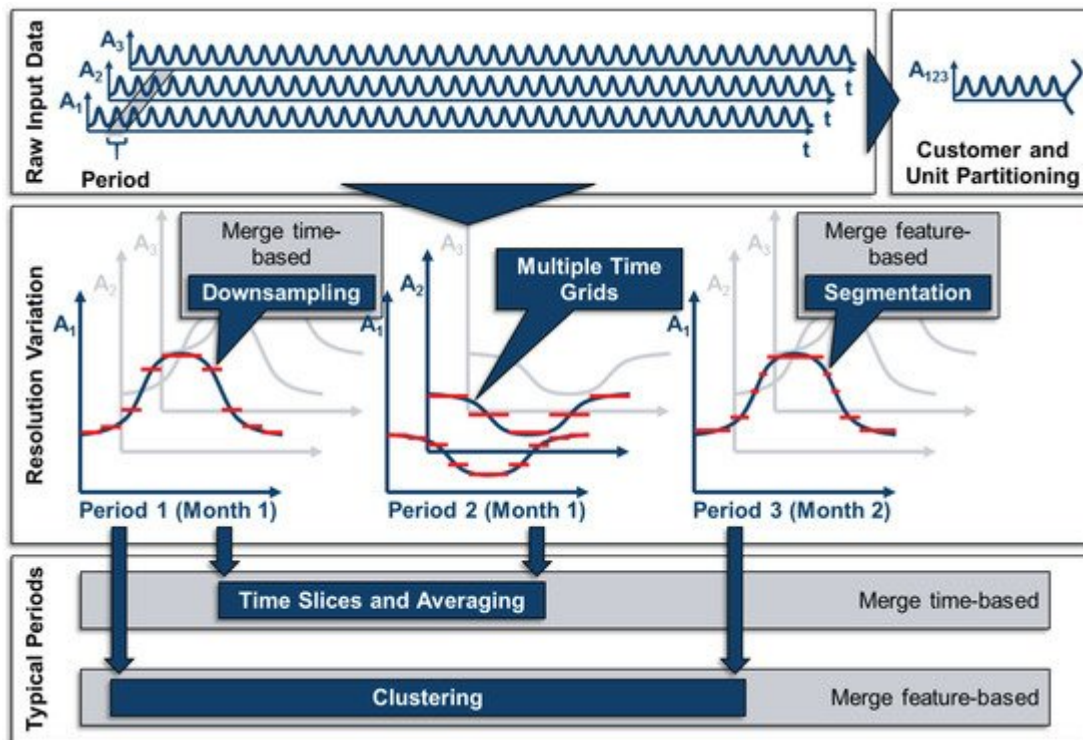

Figure 2. Methods of time series aggregation (TSA) for energy system models (ESMs).

One approach for aggregating the input time series is to merge multiple time series of attributes with a similar pattern. However, this can only be performed for attributes describing similar units (e.g., the capacity factors of similar wind turbines) or similar customer profiles (i.e., the electricity demand profiles of residential buildings). As this approach is often chosen to merge spatially distributed but similar technologies, it is not considered as TSA in the narrow sense, but as spatial or technological aggregation, as the number of time steps is not reduced in these cases. This is illustrated in the right upmost field in Figure 2.

TSA, as it is understood in this review, is the aggregation of redundant information within each time series, i.e., in the case of discrete time steps, the reduction of the overall number of time steps. This can be done in several ways. One way of reducing the number of time steps, as is shown in the central field of Figure 2, is the merging of adjacent time steps. Here, it needs to be highlighted that the periods shown in this field are for illustrative purposes only: The merging of adjacent time steps can be performed for full-length time series or time periods of time series only. Moreover, the merging of adjacent time steps can either be done in a regular manner, e.g., every two time steps are represented by one larger time step (downsampling) or in an irregular manner according to, e.g., the gradients of the time series (segmentation). A third possible approach is to individually variate the temporal

resolution for each attribute, i.e., using multiple time grids, which could also be done in an irregular manner, as pointed out by Renaldi et al. [31].

Another approach for TSA is based on the fact that many time series exhibit a fairly periodic pattern, i.e., time series for solar irradiance have a strong daily pattern. In the case of perfect periodicity, a time series could thus be represented by one period and its cardinality without the loss of any information. Based on this idea, time series are often divided into periods as already shown in the middle of Figure 2. As the periods are usually not constant throughout a year (e.g., the solar irradiance is higher in the summer than in the winter), the periods can either be merged based on their position in the calendar (time slices and averaging) or based on their similarity (clustering), as shown at the bottom of Figure 2.

In the following, methods that merge time steps or periods in a regular manner, i.e., based on their position in the time series only, will be referred to as time-based methods, whereas aggregation based on the time steps' and periods' values will be called feature-based. In this context, features refer not only to statistical features as defined by Nanopoulos et al. [32], but in a broader sense to information inherent to the time series, regardless of whether the values or the extreme values of the time series themselves or their statistical moments are used [33].

2.1. Resolution Variation

The simplest and most intuitive method for reducing the data volume of time series for ESMs is the variation of the TR. Here, three different procedures can be distinguished that have been commonly used in the literature:

2.1.1. Downsampling

Downsampling is a straightforward method for reducing the TR by representing a number of consecutive discrete time steps by only one (longer) time step, e.g., a time series for one year of hourly data is sampled down to a time series consisting of 6 h time steps. Thus, the number of time steps that must be considered in the optimization is reduced to one sixth, as demonstrated by Pfenninger et al. [34]. As the averaging of consecutive time steps leads to an underestimation of the intra-time step variability, capacities for RES tend to be underestimated because their intermittency is especially weakly represented [34]. **Figure 3** shows the impact of downsampling the PV profile from hourly resolution to 6-h time steps, resulting in one sixth of the number of time steps. In comparison to the original time series, the underestimation of extreme periods is remarkable. This phenomenon also holds true for sub-hourly time steps [35][36][37] and, for instance, in the case of an ESM containing a PV cell and a battery for a residential building, this not only has an impact on the built capacities, but also on the self-consumption rate [35][37]. For wind, the impact is comparable [36]. As highlighted by **Figure 2**, downsampling can also be applied to typical periods. To the best of our knowledge, this was initially evaluated by Yokoyama et al. [38] with the result that it could be a crucial step to resolve a highly complex problem, at least close to optimality. The general tendency of downsampling to underestimate the objective function was shown in a subsequent work by Yokoyama et al. [39] and the fact that this is not necessarily the case when combined with other methods in a third publication [40].
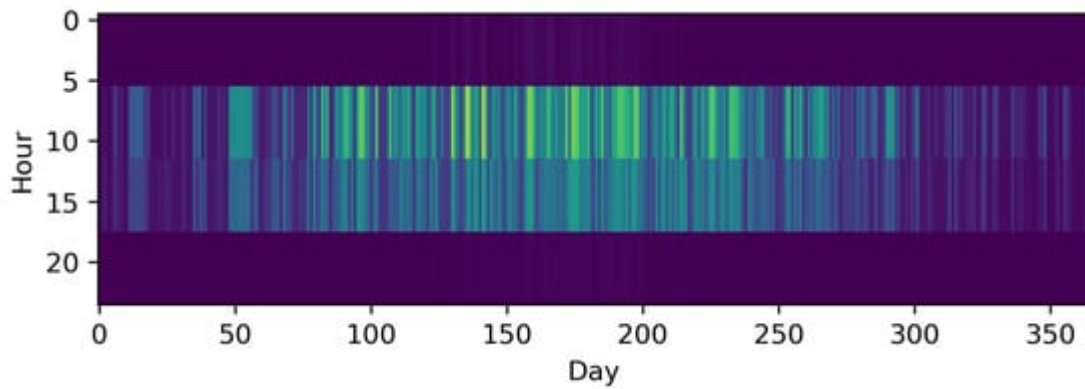
Figure 3. The time series of photovoltaic capacity factors downsampled to 1460 6 h time steps.

2.1.2. Segmentation

In contrast to downsampling, segmentation is a feature-based method of decreasing the TR of time series with arbitrary time step lengths. To the best of our knowledge, Mavrotas et al. [41] were the first to present an algorithm for segmenting time series to coarser time steps based on ordering the gradients between time steps and merging the smallest ones. Fazlollahi et al. [42] then introduced a segmentation algorithm based on k-means clustering in which extreme time steps were added in a second step. In both works, the segmentation methods were applied to typical periods, which will be explained in the following chapters. Bungener et al. [43] used evolutionary algorithms to iteratively merge the heat profiles of different units in an industrial cluster and evaluated the different solutions obtained by the algorithm with the preserved variance of the time series and the sum of zero-flow rate time steps, which indicated that a unit was not active. Deml et al. [44] used a similar, but not feature-based approach, as Mavrotas et al. and Fazlollahi et al. [41][42] for the optimization of a dispatch model. In this approach, the TR of the economic dispatch model was more reduced the further time steps lay in the future, following a discretized exponential function. Moreover, they compared the results of this approach to those of a perfect foresight approach for the fully resolved time horizon and a model-predictive control and proved the superiority of the approach, as it preserved the chronology of time steps. This was also pointed out in comparison to a typical periods approach by Pineda et al. [45], who used the centroid-based hierarchical Ward's algorithm [46] with the side condition to only merge adjacent time steps. Bahl et al. [47], meanwhile, introduced a similar algorithm as Fazlollahi et al. [42] inspired by Lloyd's algorithm and the partitioning around medoids algorithm [48][49] with multiple initializations. This approach was also utilized in succeeding publications [50][51]. In contrast to the approach of Bahl et al. [47], Stein et al. [52] did not use a hierarchical approach, but formulated an MILP in which not only extreme periods could be excluded beforehand, but also so that the grouping of too many adjacent time steps with a relatively small but monotone gradient could be avoided. The objective function relies on the minimization of the gradient error, similar to the method of Mavrotas et al. [41]. Recently, Savvidis et al. [53] investigated the effect of increasing the TR at times of the zero-crossing effect, i.e., at times when the energy system switches from the filling of storages to withdrawing and vice versa. This was compared to the opposite approach, which increased resolution at times without zero crossing. They also arrived at the conclusion that the use of irregular time steps is effective for decreasing the computational load without losing substantial information. **Figure 4** shows advantages of the hierarchical method proposed by Pineda et al. [45] compared to the simple downsampling in **Figure 3**. The inter-daily variations of the

PV profile are much more accurately preserved choosing 1460 irregular time steps compared to simple downsampling with the same number of time steps.
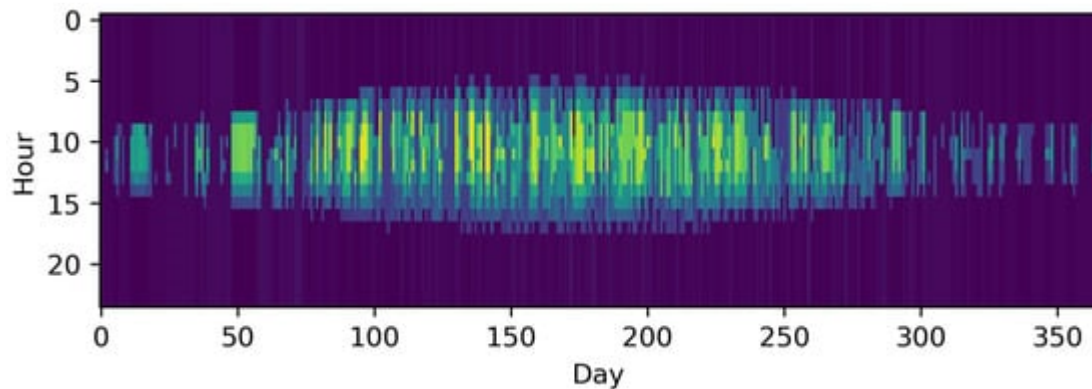

Figure 4. The time series of photovoltaic capacity factors segmented to 1460 time intervals using hierarchical merging of adjacent time steps based on centroids as proposed by Pineda et al. [45].

2.1.3. Multiple Time Grids

The idea of using multiple time grids takes into account that different components that link different time steps to each other, such as storage systems, have different time scales on which they operate [14][15][54]. For instance, batteries often exhibit daily storage behavior, whereas hydrogen technologies [14][15] or some thermal storage units [54][55] have seasonal behavior. Because of this, seasonal storage is expected to be accurately modeled with a smaller number of coarser time steps. Renaldi et al. [31] applied this principle to a solar district heating model consisting of a solar thermal collector, a backup heat boiler, and a long- and a short-term thermal storage system to achieve the optimal tradeoff between the computational load and accuracy for modeling the long-term thermal storage with 6 h time steps and the remaining components with hourly time steps. It is important to highlight that the linking of the different time grids was achieved by applying the operational state of the long-term storage to each time step of the other components if they lay within the larger time steps of the long-term storage. This especially reduced the number of binary variables of the long-term storage (because it could not charge and discharge at the same time). However, increasing the step size led to an even further increase in calculation time, as the operational flexibility of the long-term storage became too stiff and the benefit from reducing the number of variables of the long-term storage decreased. Thus, this method requires knowledge about the characteristics of each technology beforehand. Reducing the TR of single components is a highly demanding task and is left to future research.

2.2. Typical Periods

The aggregation of time series into typical periods is based on the idea that energy systems behave similarly under similar external conditions, e.g., similar energy demands and capacity factors of RES [56]. Typical periods can consist of single time steps, which are called "system states" [19][56][57][58][59], "snapshots" [33][60], or "external operation conditions" [61] in the literature, or periods containing more than one time step, e.g., "typical days" (TDs)

or "representative days", which were used by the majority of authors. In the context of control engineering, the term "system states" is especially misleading, as the state of a system not only depends on external parameters such as capacity factors and demands to be fulfilled, but also on storage levels and other endogenous state variables. Therefore, the term "system state" in discrete ESMs is only equivalent to time steps if the system is not temporally coupled, i.e., neither state variables, nor intertemporal constraints linking them with each other exist. The following will refer to "typical time steps" (TTSs) if the typical period consists of only one time step. If not stated differently in the following, the authors used TDs. However, longer periods such as typical (also called representative) weeks ([62][63][64][65] ("typical weeks"), [66][67][68][69] ("representative weeks") also exist. This work only makes further use of the word "representative" in the context of clustering, as the representative of each cluster [70] is then interpreted as the new typical period.

Analogous to the previous chapter, a number of time-based and feature-based methods exist that will be explained in the following.

2.2.1. Time-Based Merging

Time-based approaches of selecting typical periods rely on the modeler's knowledge of the model. This means that characteristics are included that are expected to have an impact on the overall design and operation of the ESM. As will be shown in the following, this was most frequently done for TDs, although similar approaches for typical weeks [63] or typical hours (i.e., TTSs) [71] exist. As pointed out by Schütz et al. [72], the time-based selection of typical periods can be divided into month-based and season-based methods, i.e., selecting a number of typical periods from either each month or from each season. However, we divide the time-based methods in consecutive typical periods and non-consecutive typical periods that are repeated as a subset with a fixed order in a pre-defined time interval.

2.2.1.1. Averaging

The method that is referred to as averaging in the following, as per Kotzur et al. [62], focuses on aggregating consecutive periods into one period. To the best of our knowledge, this idea was first introduced by Marton et al. [73], who also introduced a clustering algorithm that indicated whether a period of consecutive typical periods of Ontario's electricity demand had ended or not. In this way, the method was capable of preserving information about the order of TDs. However, it was not applied to a specific ESM. In contrast to that method, one TD for each month at hourly resolution, resulting in 288 time steps, was used by Mavrotas et al. [41], Lozano et al. [74], Schütz et al. [75], and Harb et al. [76]. Although thermal storage systems have been considered in the literature [74][75][76] (as well as a battery storage by Schütz et al. [75]), they were constrained to the same state of charge at the beginning and end of each day. The same holds true in the work of Kotzur et al. [62]. Here, thermal storage, batteries and hydrogen storage were considered and the evaluation was repeated for different numbers of averaged days. Buoro et al. [63] used one typical week per month to simulate operation cycles on a longer time scale. Kools et al. [77], in turn, clustered eight consecutive weeks in each season to one TD with 10 min resolution, which was then further down-sampled to 1 h time steps. The same was done by Harb et al. [76], who compared twelve TDs of hourly resolution to

time series with 10 min. time steps and time series down-sampled to 1 h time steps. This illustrates that both methods, downsampling and averaging, can be combined. Voll et al. [78] aggregated the energy profiles even further with only one time step per month, which can also be interpreted as one TD per month down-sampled to one time step. To account for the significant underestimation of peak loads, the winter and summer peak loads were included as additional time steps. **Figure 5** illustrates the impact of representing the original series by twelve monthly averaged consecutive typical days, i.e., 288 time steps instead of 8760.
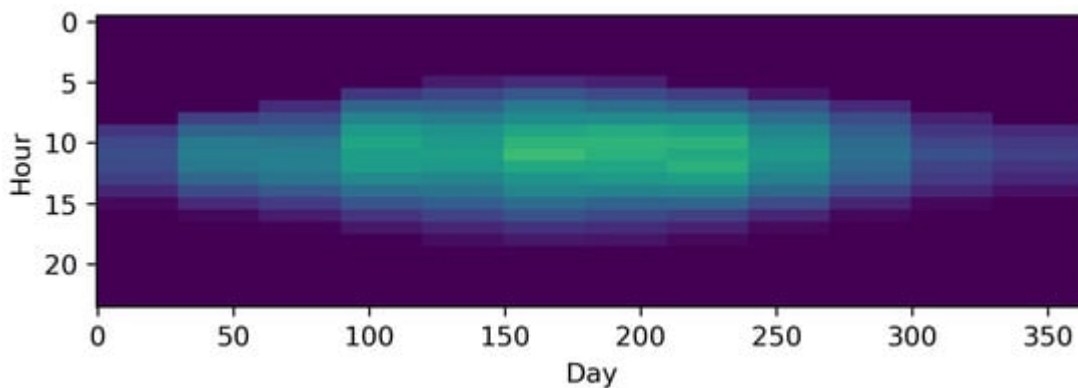

Figure 5. The time series of photovoltaic capacity factors represented by twelve monthly averaged periods as used in other studies [41][74][75] and reproduced by Kotzur et al. [62] using the python package tsam [62] (i.e., 288 different time steps).

2.2.1.2. Time Slices

To the best of our knowledge, the idea of time slices (TSs) was first introduced by the MESSAGE model [9][79] and the expression was reused for other models, such as THEA [80], LEAP [81], OSeMOSYS [82], Syn-E-Sys [83], and TIMES [84][85]. The basic idea is comparable to that of averaging, but not based on aggregating consecutive periods. Instead, TSs can be interpreted as the general case of time-based grouping of periods. Given the fact that electricity demand in particular not only depends on the season, but also on the weekday, numerous publications have used the TS method for differentiating between seasons and amongst days. In the following, this approach is referred to as time slicing, although not all of the cited publications explicitly refer to the method thus. Instead, the method is sometimes simply called "representative day" [38][39][86][87][88][89][90][91], "TD" [41][92][93][94][95][96][97][98][99], "typical daily profiles" [16][17], "typical segment" [100] "time slot" [101], or "time band" [102]. Accordingly, the term "TS" is used by the majority of authors [103][104][79][80][81][82][83][105][106][107][108][109]. The most frequent distinction is made between the four seasons [16][17][103][104][80][81][82][83][93][99][105][107][108][109] or between summer, winter. and mid-season [110][79][41][38][39][65][86][87][88][90][95][96][98][102][111][112], but other distinctions such as monthly, bi-monthly, or bi-weekly among others [110][79][41][89][91][92][94][97][100][106] can also be found. Within this macro distinction, a subordinate distinction between weekdays and weekend days [16][17][79][82][83][89][91][94][99][100][102], weekdays, Saturdays, and Sundays [93][105][107], Wednesdays, Saturdays, and Sundays [80][81], or others, such as seasonal, median, and peak [110] can be found. In contrast to the normal averaging, each TS does not follow the previous one, but is repeated in a certain order a certain number of times (e.g., five spring workdays are followed 13 times by two weekend spring days before the summer periods follow). As a visual inspection of Figure 5 and Figure 6

shows, the TS method relying on the distinction between weekdays and seasons is not always superior to a monthly distinction. The reason for this is that some input data such as the PV profile from the example have no weekly pattern and spacing the typical periods equidistantly is the better choice in this case if no other input time series (such as, e.g., electricity profiles) must be taken into account. Thus, the choice of the aggregation method should refer to the pattern of the time series considered to be especially important for the ESM. For instance, the differences between week- and weekend days is likely more important to an electricity system based on fossil fuels and without storage technologies, whereas an energy system based on a high share of RES, combined heat and power technologies, and storage units is more affected by seasonality.
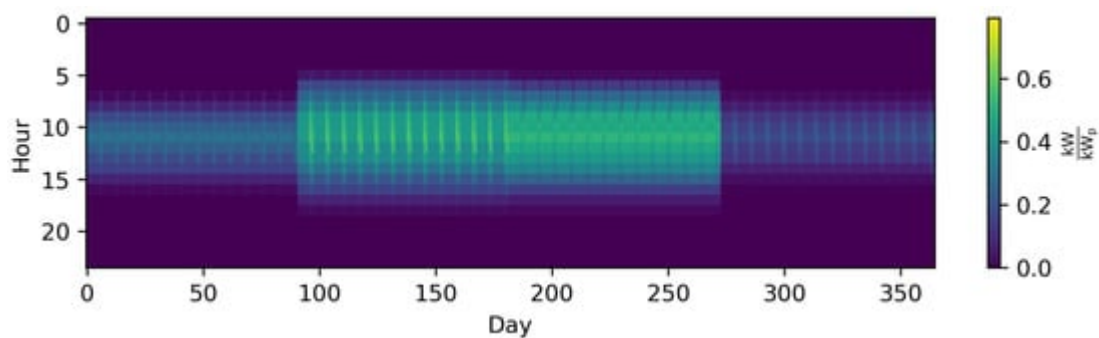

Figure 6. The time series of photovoltaic capacity factors represented by twelve time slices (TSs) (average Wednesday, Saturday and Sunday for each season) as used by Nicolosi et al. and Haydt et al. [80][81] (i.e., 288 different time steps).

2.2.1.3. Time Slices/Averaging + Downsampling/Segmentation

Like the simple averaging of consecutive time periods that can be further sampled down, e.g., as done by Harb et al. [76], the typical periods in the TS method can also be further sampled down. This can be done, for instance, by downsampling to 2 h TSs [94][96][100], 4 h TSs [110], or a number of different time step sizes to investigate the downsampling impact [38][39][40]. Moreover, day and night cycles (two diurnal TSs) [103][80][107][108][109], optionally including the peak hour of the day [103][108][109] or other TSs of irregular length [104][41][82][83][90][102][111], were also used. Mavrotas et al. [41] also implemented an algorithm for segmenting the chosen TDs to coarser TSs based on ordering of the gradients between time steps and merging the smallest ones.

The extreme case of both the downsampling method and averaging/TS method is the representation of the total time series by its mean, which was performed by Merrick et al. [110]. As this approach is unable to consider any dynamic effects, it only served as a benchmark.

2.2.2. Feature-Based Merging

In contrast to representing time series with typical periods based on a time-based method, typical periods can also be chosen on the basis of features. In this section, the clustering procedure is explained both conceptually and mathematically. To the best of our knowledge, one of the first and most frequently cited works by Domínguez-

Muñoz et al. [113] used this approach to determine typical demand days for a CHP optimization, i.e., an energy system optimization model with discrete time steps, even though it was not applied to a concrete model in this work. For this purpose, all time series are first normalized to encounter the problem of diverse attribute scales. Then, all time series are split into periods , which are compared to each other by transforming them for each value of each attribute at each time step within the period to a hyper-dimensional data point. Those data points with low distances to each other are grouped into clusters and represented by a (synthesized or existing) point within that cluster considered to be a "typical" or "representative" period. Additionally, a number of clustering algorithms are not centroid-based, i.e., they do not preserve the average value of the time series [72] which could, e.g., lead to a wrong assumption of the overall energy amount provided by an energy system across a year. To overcome this problem, time series are commonly rescaled in an additional step. This means that time series clustering includes five fundamental aspects:

- A normalization (and sometimes a dimensionality reduction).
- A distance metric.
- A clustering algorithm.
- A method to choose representatives [114].
- A rescaling step in the case of non-centroid based clustering algorithms.

As the clustered data are usually relatively sparse, while the number of dimensions increases with the number of attributes, the curse of dimensionality may lead to unintuitive results incorporating distance metrics [115], such as the Euclidean distance [114][116][117][118]. Therefore, a dimensionality reduction might be used in advance [119][120] [121], but is not further investigated in this work for the sake of brevity. In the following, each of the bullet points named above will be explained with respect to their application in TSA for ESMs. Further, the distance metric, clustering method, , because the number of clustering methods used for ESMs is small. Figure 6 shows the mandatory steps for time series clustering used for ESMs, which are presented in the following. Figure 7 shows the time series of photovoltaic capacity factors represented by 12 typical days (TDs) using k-means clustering and the python package tsam [62].
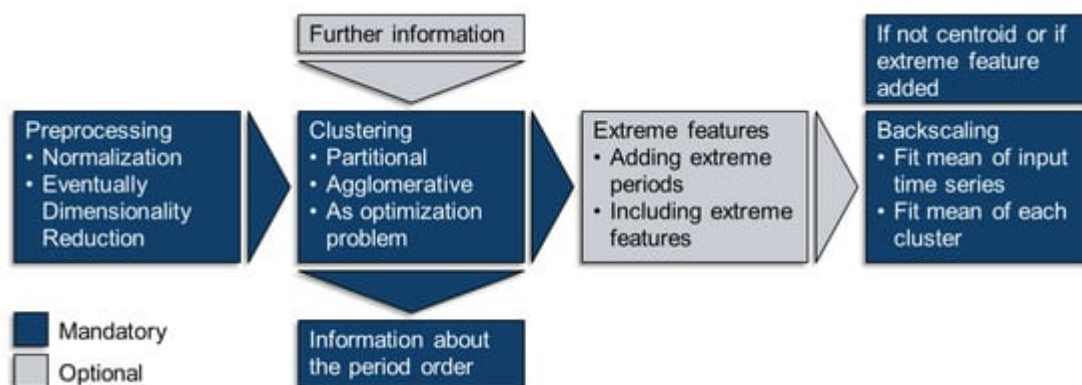

Figure 6. Steps for clustering time series for energy system models (ESMs).
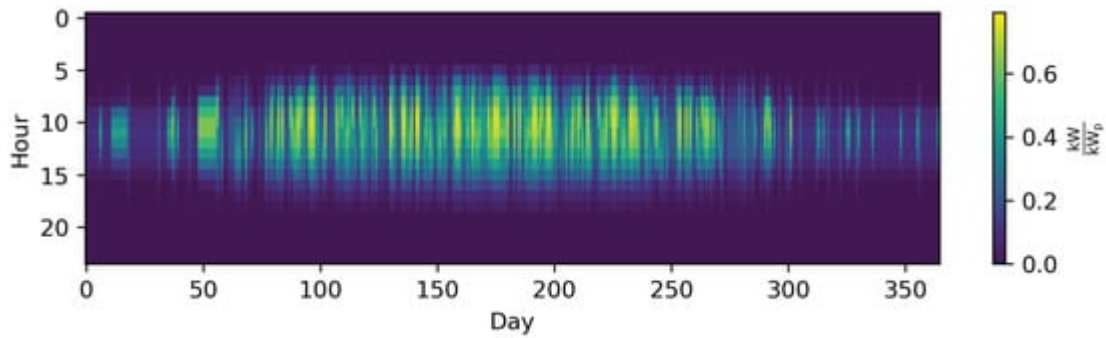
Figure 7. The time series of photovoltaic capacity factors represented by twelve typical days (TDs) using k-means clustering and the python package tsam [62] (i.e., 288 different time steps).

2.2.2.1. Preprocessing and Normalization

Clustering normally starts with preprocessing the time series, which includes a normalization step, an optional dimensionality reduction and an alignment step. Because of the diversity of scales and units amongst different attributes, they must be normalized before applying clustering algorithms to them. Otherwise, distance measures used in the clustering algorithm would focus on large-scaled attributes and other attributes would not be properly represented by the cluster centers. For example, capacity factors are defined as having values of between zero and one, whereas electricity demands can easily reach multiple gigawatts. Although a vast number of clustering algorithms exist, the min-max normalization is used in the majority of publications [14][18][104][62][72][42][66][67][119][122][123][124][125]. For the time series of an attribute $a \in A=\{1,\ldots,Na\}$ consisting of $s \in S=\{1,\ldots,Ns\}$ time steps, the normalization to the values assigned to a in time step s is calculated as follows:

$$x_{a,s} = \frac{x'_{a,s} - \min(x'_a)}{\max(x'_a) - \min(x'_a)} \tag{1}$$

In cases in which the natural lower limit is zero, such as time series for electricity demands, this is sometimes [34][59][61][68][126][127][128][129] reduced to:

$$x_{a,s} = \frac{x'_{a,s}}{\max(x'_a)} \tag{2}$$

Another normalization that can be found in the literature [130][71][131][132][133] is the z-normalization that directly accounts for the standard deviation, rather than for the maximum and minimum outliers, which implies a normal distribution with different spreads amongst different attributes:

$$x_{a,s} = \frac{x'_{a,s} - \bar{x}'_a}{\sigma(x'_a)} \tag{3}$$

In the following, the issue of dimensionality reduction will not be considered due to the fact that it is only used in a small number of publications [119][120][121] and transforms the data into eigenvectors to tackle the non-trivial behavior of distance measures used for clustering in hyper-dimensional spaces [117].

A time series can further be divided into a set of periods P and a set of time steps within each period T, i.e., S=P×T. The periods are clustered into non-overlapping subsets PC, which are then represented by a representative period, respectively. A representative period consists of at least one discrete time step and, depending on the number and duration of time steps, it is often referred to as a typical hour, snapshot or system state, typical or representative day, or typical week. The data D=A×P×T can thus be rearranged so that each period is represented by a row vector in which all inter-period time steps of all attributes are concatenated, i.e.,

$$\mathbb{D}_{arr} = \begin{pmatrix} x_{1,1,1} & \cdots & x_{1,1,N_t} & x_{1,2,1} & \cdots & x_{1,N_a,N_t} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ x_{N_p,1,1} & \cdots & x_{N_p,1,N_t} & x_{N_p,2,1} & \cdots & x_{N_p,N_a.N_t} \end{pmatrix} \quad \text{with} \quad \begin{array}{l} a \in \mathbb{A} = \{1, \ldots, N_a\} \\ p \in \mathbb{P} = \{1, \ldots, N_p\} \\ t \in \mathbb{T} = \{1, \ldots, N_t\} \end{array} \tag{4}$$

The row vectors of Darr are now grouped with respect to their similarity. Finally, yet importantly, it must be highlighted that the inner-period time step values can also be sorted in descending order, which means that, in this case, the duration curves of the periods are clustered as done in other studies [18][124][134][135]. This can reduce the averaging effect of clustering time series without periodic patterns such as wind time series.

2.2.2.2. Algorithms, Distance Metrics, Representation

Although a vast number of different clustering algorithms exist [70][136] and have been used for time series clustering in general [114], only a relatively small number of regular clustering algorithms have been used for clustering input data for energy system optimization problems, which will be presented in the following. Apart from that, a number of modified clustering methods have been implemented in order to account for certain properties of the time series. The goal of all clustering methods is to meaningfully group data based on their similarity, which means minimizing the intra-cluster difference (homogeneity) or maximizing the inter-cluster difference (separability) or a combination of the two [137]. However, this depends on the question of how the differences are defined. To begin with, the clustering algorithms can be separated into partitional and deterministic hierarchical algorithms.

Partitional Clustering

One of the most common partitional clustering algorithms used in energy system optimization is the k-means algorithm, which has been used in a variety of studies [14][15][24][34][62][72][33][42][47][51][56][57][58][59][60][71][121][122][123][125][126][129][131][132][133][138][139][140][141][142][143][144][145][146]. The objective of the k-means algorithm is to minimize the sum of the squared distances between all cluster members of all clusters and the corresponding cluster centers, i.e.,

$$\min \sum_{k=1}^{N_k} \sum_{p \in \mathbb{C}_k} \text{dist}\left(x_p, c_k\right)^2 \tag{5}$$

The distance metric in this case is the Euclidean distance between the hyperdimensional period vectors with the dimension dim(vec(T×A))

and their cluster centers ck, i.e.,

$$\text{dist}\left(x_p, c_k\right) = \sqrt{\sum_{a=1}^{N_a} \sum_{t=1}^{N_t} \left(x_{p,a,t} - c_k\right)^2} \tag{6}$$

where the cluster centers are defined as the centroid of each cluster, i.e.:

$$c_k = \frac{1}{|\mathbb{C}_k|} \sum_{p \in \mathbb{C}_k} x_{p,a,t} \tag{7}$$

This NP-hard problem is generally solved by an adopted version [49] of Lloyd's algorithm [48], a greedy algorithm that heuristically converges to a local minimum. As multiple runs are performed in order to improve the local optimum, improved versions (such as k-means++) for setting initial cluster centers have also been proposed in the literature [147].

The only difference regarding the k-medoids algorithm is that the cluster centers are defined as samples from the dataset that minimize the sum of the intra-cluster distances, i.e., that are closest to the clusters' centroids.

$$c_k = \underset{x_l \in \mathbb{C}_k}{\arg \min} \frac{1}{N_k} \sum_{p \in \mathbb{C}_k} \text{dist}\left(x_p, x_l\right)^2 \tag{8}$$

This clustering algorithm was used by numerous authors [14][19][130][113][62][72][47][51][59][123][125][134][135][144][148][149][150][151][152], either by using the partitioning around medoids (PAM) introduced by Kaufman et al. [153] or by using an

MILP formulation introduced by Vinod et al. [154] and used in several studies [14][130][113][62][123][144][149]. The MILP can be formulated as follows:

$$\min \sum_{i=1}^{N_p} \sum_{j=1}^{N_p} \text{dist}\,(x_i, x_j) \times z_{i,j} \tag{9}$$

Subject to:

$$\sum_{j=1}^{N_p} z_{i,j} = 1 \ \forall \ j \in 1, \ldots, N_i \tag{10}$$

$$z_{i,j} \leq y_i \ \forall \ i, j \in 1, \ldots, N_i \tag{11}$$

$$\sum_{i=1}^{N_i} y_i = N_k \tag{12}$$

In a number of publications [130][62][72][59][123][125][144], k-medoids clustering was directly compared to k-means clustering. The general observation is that k-medoids clustering is more capable of preserving the intra-period variance, while k-means clustering underestimates extreme events more gravely. Nevertheless, the medoids lead to higher root mean squared errors compared to the original time series. This leads to the phenomenon that k-medoids outperforms k-means in the cases of energy systems sensitive to high variance, as in self-sufficient buildings, e.g., as shown by Kotzur et al. [62] and Schütz et al. [123]. In contrast to that, k-means outperforms k-medoids clustering in the case of smooth demand time series and non-rescaled medoids that do not match the overall annual demand in the case of k-medoids clustering, as shown by Zatti et al. [125] for the energy system of a university campus.

Agglomerative Clustering

In contrast to partitional clustering algorithms that iteratively determine a set consisting of k clusters in each iteration step, agglomerative clustering algorithms such as Ward's hierarchical algorithm [46] stepwise merge clusters aimed at minimizing the increase in intra-cluster variance

$$SSE = \sum_{p \in C_k} \text{dist}\,(x_p, c_k)^2 \tag{13}$$

in each merging step until the data is agglomerated to k clusters. The algorithm is thus deterministic and does not require multiple random starting point initializations. Analogously to k-means and k-medoids, the cluster centers

can either be represented by their centroids [130][144] or by their medoids [18][34][130][62][45][59][119][124][127][128][133][144]. The general property that centroids underestimate the intra-period variance more severely due to the averaging effect is equivalent to the findings when using k-means instead of k-medoids.

Rarely Used Clustering Algorithms

Apart from the frequently used clustering algorithms in the literature, two more clustering algorithms were used in the context of determining typical periods based on unsorted time intervals of consistent lengths.

K-medians clustering is another partitional clustering algorithm that is closely related to the k-means algorithm and has been used in other studies [72][123]. Taking into account that the Euclidean distance is only the special case for of the Minkowski distance [155]

$$\text{dist}\,(x_p, c_k) = \left(\sum_{a=1}^{N_a} \sum_{t=1}^{N_t} |x_{p,a,t} - c_k|^{\gamma} \right)^{\frac{1}{\gamma}} \tag{14}$$

K-medians generally tries to minimize the sum of the distances of all data points to their cluster center in the Manhattan norm, i.e., for $\gamma=1$

and the objective function [156][157]:

$$\min \sum_{k=1}^{N_k} \sum_{p \in C_k} \text{dist}\,(x_p, c_k) \quad \text{with} \quad \text{dist}\,(x_p, c_k) = \sum_{a=1}^{N_a} \sum_{t=1}^{N_t} |x_{p,a,t} - c_k| \tag{15}$$

For this, the L1 distance is usually used in the assignment step [156] and the median is calculated in each direction to minimize the L1 distance within each cluster [157]. However, Schütz et al. [72][123] used the Euclidean distance (like for k-means) in the assignment step to isolate the influence of using dimension-wise medians instead of dimension-wise means (i.e., centroids). Thus, all values come from the original dataset, but not necessarily from the same candidates [72].

Moreover, Schütz et al. [72][123] used k-centers clustering, which minimizes the maximum distance of all candidates to its cluster center, i.e., according to Har-Peled [158],

$$\min_{C,|C|=k} \left(\max_{p \in C_k} (\text{dist}\,(x_p, c_k)) \right) \quad \text{with} \quad \text{dist}\,(x_p, c_k) = \sqrt{\sum_{a=1}^{N_a} \sum_{t=1}^{N_t} (x_{p,a,t} - c_k)^2} \tag{16}$$

Time Shift-Tolerant Clustering Algorithms

The last group of clustering algorithms applied for TSA in ESMs is time shift-tolerant clustering algorithms. These algorithms not only compare to the values of different time series at single time steps (pointwise), but also compare values along the time axis with those of other time series (pairwise). In the literature [130][144], dynamic time warping (DTW) and the k-shape algorithm are used, both of which are based on distance measures that are not sensitive to phase shifts within a typical period, which is the case for the Euclidean distance. The dynamic time-warping distance is defined as:

$$\text{dist}\,(x_p, c_k) = \min_{w} \sqrt{\sum_{l=1}^{L} w_l} \qquad (17)$$

where w describes the so-called warping path, which is the path of minimal deviations across the matrix of cross-deviations between any entry of x_p and any entry of c_k [130][159]. The cluster centers c_k are determined using DTW Barycenter averaging, which is the centroid of each time series value (within an allowed warping window) assigned to the time step [160]. Moreover, a warping window [130][144] can be determined that limits the assignment of entries across the time steps. Shape-based clustering uses a similar algorithm and tries to maximize the cross-correlation amongst the periods. Here, the distance measure to be minimized is the cross-correlation and the period vectors are uniformly shifted against each other to maximize it [130][144][159][161]. It must be highlighted that both dynamic time warping and shape-based distance, have only been applied on the clustering of electricity prices, i.e., only one attribute [130][144]. Moreover, Liu et al. [133] also applied dynamic time warping to demand, solar, and wind capacity factors simultaneously. However, it is unclear how it was guaranteed that different attributes were not compared to each other within the warping window which remains a field of future research. Furthermore, a band distance, which is also a pairwise rather than a pointwise distance measure, was used in a k-medoids algorithm by Tupper et al. [152], leading to significantly less loss of load when deriving operational decisions for the next day using a stochastic optimization model.

Due to the fact that not all of the methods rely on the representation of each cluster by its centroid (i.e., the mean in each dimension), these typical periods do not meet the overall average value when weighted by their number of appearances and must be rescaled. This also holds true for the consideration of extreme periods, which will be explained in the following chapters. Accordingly, the following section will be referred to if rescaling is considered in the implementation of extreme periods. To the best of our knowledge, the first work that used clustering not based on centroids was that of Domínguez-Muñoz et al. [113], in which the exact k-medoids approach was chosen as per Vinod et al. [154]. Here, each attribute (time series) of each TD was rescaled to the respective cluster's mean, i.e.,

$$c_{k,a,t}^* = c_{k,a,t} \frac{\sum_{p \in \mathbb{C}_k} \sum_{t=1}^{N_t} x_{p,a,t}}{|\mathbb{C}_k| \sum_{t=1}^{N_t} c_{k,a,t}} \; \forall \, k, a, t \qquad (18)$$

Furthermore, Domínguez-Muñoz et al. [113] discarded the extreme values that were manually added from the rescaling procedure. A similar procedure, which was applied for each time series, but not for each TD, was introduced by Nahmmacher et al. [127], who used hierarchical clustering based on Ward's algorithm [46] and chose medoids as representatives, which was later used in a number of other studies [14][18][130][62][124][144]. Here, all representative days were rescaled to fit the overall yearly average when multiplied by their cardinality and summed up, but not the average of their respective clusters, i.e.,

$$c_{k,a,t}^* = c_{k,a,t} \frac{\sum_{p=1}^{N_p} \sum_{t=1}^{N_t} x_{p,a,t}}{\sum_{k=1}^{N_k} \left(|\mathbb{C}_k| \sum_{t=1}^{N_t} c_{k,a,t} \right)} \; \forall \, k, a, t \qquad (19)$$

Schütz et al. [72][123], Bahl et al. [47], and Marquant et al. [134][135] refer to the method of Domínguez-Muñoz et al. [113], but some used it time series-wise and not cluster- and time series-wise. Schütz et al. [72][123] were the first to highlight that both approaches are possible. It also needs to be highlighted that these methods are not the only methods, as Zatti et al. [125], for instance, presented a method to choose medoids within the optimization problem without violating a predefined maximum deviation from the original data, but for the sake of simplicity, it focused on the most frequently used post-processing approaches. Additionally, other early publications, such as by Schiefelbein et al. [148], did not use rescaling at all. Finally, yet importantly, the rescaling combined with the min-max normalization could lead to values over one. Accordingly, these values were reset to one so as to not overestimate the maximum values and the rescaling process was re-run in several studies [14][18][62][124][127]. In contrast, Teichgräber et al. [130][144] used the z-normalization with rescaling in accordance with Nahmmacher et al. [127], but did not assure that the original extreme values were not overestimated by rescaling.

2.2.3. Modified Feature-Based Merging

Apart from the methods that are based on the direct clustering of the time series' values or periods, a number of methods exist that group time series in a consecutive manner [162], by means of other features, such as sorted time series (i.e., duration curves) [18][20][66][67][68][69][124][128][134][135][163] or other statistical features such as the average, variance, minimal and maximal values [33], or predefine the clusters based on additional information [61]. These methods will be presented in the following.

With respect to grouping consecutive typical periods, an early publication by Balachandra et al. [162] started by grouping daily residual load profiles by month, and then applied multiple discriminant analysis to these groups and reclassified the days at the beginning or end of a group (month) to the preceding or subsequent group if they were

more similar to it. This resulted in nine consecutive groups represented by their centroids. However, this aggregation was not applied to an energy system optimization.

Furthermore, a number of publications [20][68][69][128] rely on the principle introduced by Poncelet et al. [163]. For this, the normalized duration curves were placed into bins, i.e., how many hours of the year surpass a certain level between zero and the maximum level of the specific attribute. The same was performed for each candidate day. Then, the sum of absolute differences between the hours at which the reference curve surpassed a bin border and the hours at which the curve derived from a linear combination of a given number of candidates surpassed the same bin borders was minimized in an MILP.

Another approach aimed at reproducing a yearly duration curve was introduced by de Sisternes et al. [66][67]. Here, the duration curve of power feed-in by wind and solar at a certain penetration level was calculated and approximated by an exhaustive search for a combination from a subset of typical weeks. As this was a combinatorial problem, the computation time rapidly increased for higher numbers of weeks. In a later publication [66], the variability of the selected weeks was used as an additional metric.

Instead of clustering the original time series, the yearly duration curve was approximated in a number of publications [18][124][134][135]. For this, the candidate days were simply sorted prior to being clustered. This decreased the averaging effect of statistical events, such as wind, as the largest value and second largest, etc. always lay in the first dimension and second dimension, etc.

With respect to the clustering of other statistical features apart from the distribution curve (duration curve), Agapoff et al. [33] applied k-means clustering to snapshots (i.e., TTSs) and used different features for the clustering: either absolute values or the average, minimum, maximum, and standard deviation of all considered regions for either price differences, non-controllable demand and generation, or both. This is an extension with promising results to all thus far used clustering algorithms only applied to normalized absolute values.

Finally, yet significantly, Lythcke-Jørgensen et al. [61] introduced a so-called CHOP-method that was based on splitting the range of each attribute, in this case the power price and relative heat demand on a five-year basis, into different intervals based on important values (e.g., zero-price) and even divisions between them. Then, all values (i.e., hours) were transferred in a 2d space in which the intervals for both attributes formed a grid. From each cell, the centroid was subsequently calculated if it contained any candidate hours. As information about the chronology of these TTSs was lost, the design of storage technologies resulted in large deviations from the reference case.

These cases highlight that methods based on well-known approaches are constantly customized for specific ESMs and improved where possible, which illustrates that the development of TSA methods is a dynamic process.

2.2.4. Linking Typical Periods

As mentioned above, some components, such as storage components of ESMs, link consecutive time steps by means of intertemporal constraints. The representation of time series by a few TDs or weeks does not generally

take their order across the entire time horizon into account. This means that the modeling of filling levels is normally only possible within these typical periods with a periodic boundary condition for the state of charge. In this case, the order of typical periods no longer plays a role. On the other hand, seasonal storage cannot be sufficiently modeled by this method. Yet, this is especially important for energy systems based on a high share of RES. As per Bauer et al. [54], central solar heating plants with introduced short-term heat storages can typically supply 15–20% of the total residential heating demand. With seasonal heat storages, this fraction can be increased to about 50%. For a long period of time, the only approach to model seasonal storages was to drastically reduce TR, as by Tveit et al. [164], making it impossible to model short-term storages. To overcome this issue, different methods have been developed that take the linking of TDs into account.

As far as we know, the TIMES framework was the first to deal with linking TSs not only consecutively, but also inter-period storages that work on a larger time scale [165][166][84][85]. However, since the inter-period storages are meant to work between different years, e.g., as waste disposal sites [165], they are not linked to the intra-period storages, which only link consecutive TSs (segments) within one typical period, such as weekdays in spring.

Welsch et al. [82] and Samsatli et al. [16] independently developed a non-uniform hierarchical time discretization that is based on the selection of TSs. In two publications [16][17], Samsatli et al. chose two TDs with hourly data for both the week and weekend which was done for each season consisting of 13 weeks. This resulted in 192 time steps. For the modeling of the seasonal storage, the energy surplus across each time scale was determined and added up. As the chosen days appeared in a regular order within each season, the capacity constraints were not postulated for each time step. Instead, they were only defined for the first and last instance of each day type, the first and last week of each season, and the first and last season of each year, if a multiple year approach was chosen. Welsch et al. [82] chose a similar approach that consisted of three TSs for a workday and a weekend day in each season. However, the case study was only run with one TD with an hourly resolution.

Both approaches did not consider a self-discharge rate. The approach of Welsch et al. [82] was later developed by Timmerman et al. [83] to handle self-discharge and re-used by van der Heijde et al. [69]. Since the typical days in these publications [16][17][69][82][83] are aligned in a regular manner, the critical storage levels can only be reached at certain time steps which significantly reduces the number of side constraints. Taking the configuration used by Timmerman et al. [83] as an example, five identical workdays alternate 13 times per season with two identical weekend days. As each week consists of only two day types, of which the first is repeated five times and the second is repeated twice, the intermediate weekdays representing Tuesday, Wednesday, and Thursday cannot include critical states of charge-neither for a rising state of charge across the weekdays (the critical day would be Friday), nor for a decreasing state of charge across the weekdays (the critical day would be Monday). The same holds true for the intermediate weeks in each season. As they are repeated 13 times, either the first or last week of each season is critical with respect to the state of charge of seasonal storage and their capacity.

Similarly, but again independently, Spiecker et al. [94] developed a comparable approach that linked workdays and weekend days for every second month in an inter-day manner for pumped storage plants and an inter-month

manner for large-scale storage systems in the E2M2s model. Moreover, the TDs were based on a recombining decision tree of 2 h segments and were thus capable of modeling the storage size stochastically.

Gabrielli et al. [15] developed a method to couple TDs using a function that assigns each day of the original time series to the TD it is represented by. This function is used to couple the state of charge of consecutive (typical) days in an additional equation and means that the operation of the components is modeled for a number of TDs, while the state of charge of the storages is modeled for the entire time horizon represented by a sequence of TDs. The approach was tested for a different number of TDs, as well as in a later publication [15][145].

Wogrin et al. [58] earlier proposed the same approach as Gabrielli et al. [15] for TTSs and took the information of the clustering indices, i.e., which original time step was represented by which TTS, to link TTS in order to consider start-up and shut-down costs, which was later re-used by Tejada-Arango et al. [19] for the calculation of storage levels using typical periods (days and weeks). However, in contrast to Gabrielli et al. [15], the storage levels were not constrained for each time step by Tejada-Arango et al. [19], but only at intervals of one week. Additionally, a similar method was applied to avoid unnecessary unit transitions at the border between two consecutive TDs.

Like the idea of Gabrielli et al. [15], Kotzur et al. [14] introduced a similar method of linking TDs in a chronologically correct order. Instead of directly linking each state of charge to the preceding one, the superposition principle was used to distinguish intraday and interday states of charge. Here, the interday state of charge describes the state of charge at the beginning of each day, while the intraday state of charge is defined to be zero at the beginning of each day but is defined for each hour of each TD. The sum of both values, i.e., the intraday state of charge for a given number of TDs, along with the interday state of charge, which was determined by the sum of storage level differences of each TD in the corresponding sequence, was then used to determine the storage levels at each time step. This approach was also used in later publications dealing with seasonal storage [18][124].

Another slight deviation of this method was applied by van der Heijde et al. [20], who also used the superposition principle discussed by Kotzur et al. [14] to couple TDs. However, they did not use clustering algorithms to group similar days and represented these by one TD for each cluster, but instead searched for a linear combination of days that minimized the deviation from the yearly duration curve; a procedure introduced by Poncelet et al. [128]. In contrast to clustering algorithms, this procedure did not directly lead to an assignment of original days to groups represented by single TDs. This meant that this had to be performed in a separate step. For this, a mixed integer quadratic programm (MIQP) problem was formulated that sought to minimize the sum of squared errors of each day of the original time series to the TDs. The outcome of this was a sequence of TDs that represented the original time series, which was crucial for linking the TDs in accordance with the aforementioned approach of Kotzur et al. [14]. Recently, Baumgärtner et al. [50] included the storage formulation of Kotzur et al. [14] in their rigorous synthesis of energy systems using aggregation approaches to define upper and lower bounds for the objective function with full time resolution.

The fact that a number of methods for linking typical periods were independently developed [14][15][16][58][82] shows the value of giving an overview over all of the methods employed so far.

2.3. Random Sampling

Another minor group of publications uses TSA based on random sampling. This means that the time steps or periods are randomly chosen from the original time series and considered to be representative for the entire time series. Most of the methods in the following deal with single time steps instead of periods, which is an acceptable simplification when the impact of storage capacity or other intertemporal constraints on the system design can be neglected [151].

2.3.1. Unsupervised

As with supervised random sampling methods, unsupervised random sampling methods can be applied to typical periods or single time steps. However, they appeared earlier than the supervised methods (2011 and 2012).

Ortiga et al. [167] introduced a graphical method for which a number of days from the dataset had to be defined. In a second step, the algorithm minimized the deviation between the duration curve of the original dataset and a duration curve of the chosen periods multiplied by a set of variable factors for the number of appearances of each TD.

With respect to the random sampling of time steps, Van der Weijde et al. [168] sampled 500 out of 8760 h to capture major correlations of the input data for seven regions.

However, in the years since 2012, these methods were substituted by supervised random sampling methods.

2.3.2. Supervised

Munoz et al. [169] applied supervised random sampling for 1 up to 300 daily samples out of a dataset of seven years, which were then benchmarked against the k-means clustering of typical hours. A similar method was used by Frew et al. [170], who took two extreme days and eight random days from the dataset and weighted each day so that the sum of squared errors to the original wind, solar and load distribution was minimized. This procedure was then repeated for ten different sets of different days, with the average of each optimization outcome calculated at the end. With respect to time steps, Härtel et al. [59] either systematically determined samples taking every n_{th} element from the time series or randomly chose 10,000 random samples from the original dataset and selected the one that minimized the deviation to the original dataset with respect to moments (e.g., correlation, mean and standard variation). Another complex algorithm for representing seasonal or monthly wind time series was proposed by Neniškis et al. [79] and tested in the MESSAGE model. This approach took into account both the output distribution (duration curve) for a TD and the inter-daily variance, not to be exceeded by more than a predefined tolerance, while using a random sampling process. However, only the typical days for wind were calculated in this way, whereas the other time series (electricity and heat) were chosen using TS. Recently, Hilbers et al. [151] used the sampling method twice with different numbers of random initial samples drawn from 36 years. From a first run, the 60 most expensive random samples were taken and included in a second run with the same number of samples.

These methods are fairly comparable to the method of clustering TTSs. However, the initial selection of samples is based on random choice.

2.4. Miscellaneous Methods

Apart from the random sampling methods that cannot be systematically categorized with the scheme in Figure 1, an even smaller number of publications cannot be grouped in any way with respect to their TSA methods. For the sake of completeness, however, they are presented in the following.

Lee et al. [171] used an improved particle swarm optimization to optimize the UC of a power system with respect to fuel and outage costs. This method was based on an evolutionary algorithm that iteratively determined the "fittest" solutions and thus was quite comparable to supervised random sampling methods. However, the use of an own class of optimization algorithm is a unique feature. A similar approach to solve the UC problem of a grid-connected building with renewable energy sources and a battery was presented by Quang et al. [172]. In their work, a genetic algorithm and a particle swarm algorithm were used for different charge and discharge rates of the battery based on half-hourly time steps. It is worth mentioning that apart from these publications, a number of other works exist which use, among other methods, genetic algorithms or particle swarm algorithms to UC models. A comprehensive review on the methods to address the UC problem was given by Saravanan et al. [173]. However, these approaches are based on a survival of the fittest principle and not on a classic optimization problem so that an aggregation can only be applied by downsampling the time steps used for simulation. Moreover, these approaches are not directly applicable to combined UC and GEP models. Therefore, these methods are not further analyzed within the scope of this paper.

Xiao et al. [174] optimized the capacity of a battery and a diesel generator for an island system by searching for the optimal cut-off frequency at which the running of a diesel generator was more convenient without causing overly high fuel costs, whereas the battery capacity would be too large if it was run on a high frequency band. For this, an analysis based on discrete Fourier transform (DFT) was used, highlighting the different specific cost-dependent time scales on which different technologies operate.

More recently, Pöstges et al. [175] introduced an analytical approach to aggregate the time steps of a demand duration curve for a simple ESM without storage units and with only one energy type. Interestingly, this method led to a simplified problem formulation based on a minimum number of time steps without causing an error in the objective function. In this case, the supply technology costs are based on capacity- and operation-specific costs and the approach was inspired by an earlier work of Sherali et al. [6]. Sherali et al. proved in 1982 that the cost optimal operation of these simple systems can be interpreted as an optimization problem which is closely related to the peak load pricing theory introduced by Boiteux in 1949 [3] (English translation in 1960 [4]) and Steiner in 1957 [5].

To summarize, special methods that cannot be categorized in any way appear in an irregular manner, but can have special implications for the improvement of preexisting methods.

References

1. Robinius, M.; Otto, A.; Heuser, P.; Welder, L.; Syranidis, K.; Ryberg, D.S.; Grube, T.; Markewitz, P.; Peters, R.; Stolten, D. Linking the Power and Transport Sectors—Part 1: The Principle of Sector Coupling. Energies 2017, 10, 956.

2. Barnett, H.J. Energy Uses and Supplies 1950, 1947, 1965, Bureau of Mines: Washington, DC, USA, 1950.

3. Boiteux, M. La Tarification des Demandes en Pointe. Rev. Gen. De L'electricite 1949, 58, 157–179.

4. Boiteux, M. Peak-Load Pricing. J. Bus. 1960, 33, 157–179.

5. Steiner, P.O. Peak loads and efficient pricing. Q. J. Econ. 1957, 71, 585–610.

6. Sherali, H.D.; Soyster, A.L.; Murphy, F.H.; Sen, S. Linear programming based analysis of marginal cost pricing in electric utility capacity expansion. Eur. J. Oper. Res. 1982, 11, 349–360.

7. Helm, D. Energy policy: Security of supply, sustainability and competition. Energy Policy 2002, 30, 173–184.

8. Hoffman, K.C.; Wood, D.O. Energy System Modeling and Forecasting. Annu. Rev. Energy 1976, 1, 423–453.

9. Lopion, P.; Markewitz, P.; Robinius, M.; Stolten, D. A review of current challenges and trends in energy systems modeling. Renew. Sustain. Energy Rev. 2018, 96, 156–166.

10. Caramanis, M.C.; Tabors, R.D.; Nochur, K.S.; Schweppe, F.C. The Introduction of Non-Dispatchable Technologies as Decision Variables in Long-Term Generation Expansion Models. Ieee Power Eng. Rev. 1982, PER-2, 40–41.

11. Bhattacharyya, S.C.; Timilsina, G.R. A review of energy system models. Int. J. Energy Sect. Manag. 2010, 4, 494–518.

12. Pfenninger, S.; Hawkes, A.; Keirstead, J. Energy systems modeling for twenty-first century energy challenges. Renew. Sustain. Energy Rev. 2014, 33, 74–86.

13. Ringkjøb, H.-K.; Haugan, P.M.; Solbrekke, I.M. A review of modelling tools for energy and electricity systems with large shares of variable renewables. Renew. Sustain. Energy Rev. 2018, 96, 440–459.

14. Kotzur, L.; Markewitz, P.; Robinius, M.; Stolten, D. Time series aggregation for energy system design: Modeling seasonal storage. Appl. Energy 2018, 213, 123–135.

15. Gabrielli, P.; Gazzani, M.; Martelli, E.; Mazzotti, M. Optimal design of multi-energy systems with seasonal storage. Appl. Energy 2018, 219, 408–424.

16. Samsatli, S.; Samsatli, N.J. A general spatio-temporal model of energy systems with a detailed account of transport and storage. Comput. Chem. Eng. 2015, 80, 155–176.

17. Samsatli, S.; Staffell, I.; Samsatli, N.J. Optimal design and operation of integrated wind-hydrogen-electricity networks for decarbonising the domestic transport sector in Great Britain. Int. J. Hydrog. Energy 2016, 41, 447–475.

18. Welder, L.; Ryberg, D.; Kotzur, L.; Grube, T.; Robinius, M.; Stolten, D. Spatio-Temporal Optimization of a Future Energy System for Power-to-Hydrogen Applications in Germany. Energy 2018, 158, 1130–1149.

19. Tejada-Arango, D.A.; Domeshek, M.; Wogrin, S.; Centeno, E. Enhanced Representative Days and System States Modeling for Energy Storage Investment Analysis. IEEE Trans. Power Syst. 2018, 33, 6534–6544.

20. van der Heijde, B.; Vandermeulen, A.; Salenbien, R.; Helsen, L. Representative days selection for district energy system optimisation: A solar district heating system with seasonal storage. Appl. Energy 2019, 248, 79–94.

21. Ören, T.I. Computer-Aided Systems technology: Its role in advanced computerization. In Computer Aided Systems Theory; Pichler, F., Moreno Díaz, R., Eds.; Springer Berlin Heidelberg: Berlin/Heidelberg, Germany, 1994; pp. 11–20.

22. Sass, S.; Mitsos, A. Optimal Operation of Dynamic (Energy) Systems: When are Quasi-Steady Models Adequate? Comput. Chem. Eng. 2019.

23. Morales-España, G.; Tejada-Arango, D. Modelling the Hidden Flexibility of Clustered Unit Commitment. IEEE Trans. Power Syst. 2018.

24. Lara, C.L.; Mallapragada, D.S.; Papageorgiou, D.J.; Venkatesh, A.; Grossmann, I.E. Deterministic electric power infrastructure planning: Mixed-integer programming model and nested decomposition algorithm. Eur. J. Oper. Res. 2018, 271, 1037–1054.

25. Lopion, P.; Markewitz, P.; Stolten, D.; Robinius, M. Cost Uncertainties in Energy System Optimisation Models: A Quadratic Programming Approach for Avoiding Penny Switching Effects. Energies 2019, 12, 4006.

26. Klinge Jacobsen, H. Integrating the bottom-up and top-down approach to energy–economy modelling: The case of Denmark. Energy Econ. 1998, 20, 443–461.

27. Subramanian, A.; Gundersen, T.; Adams, T. Modeling and simulation of energy systems: A review. Processes 2018, 6, 238.

28. Böhringer, C.; Rutherford, T.F. Integrating bottom-up into top-down: A mixed complementarity approach. Zew-Cent. Eur. Econ. Res. Discuss. Pap. 2005, 05-028.

29. Herbst, M.; Toro, F.; Reitze, F.; Eberhard, J. Bridging Macroeconomic and Bottom up Energy Models-the Case of Efficiency in Industry. EceeNeth. 2012.

30. Helgesen, P.I. Top-down and Bottom-up: Combining energy system models and macroeconomic general equilibrium models. Censes: TrondheimNor. 2013.

31. Renaldi, R.; Friedrich, D. Multiple time grids in operational optimisation of energy systems with short- and long-term thermal energy storage. Energy 2017, 133, 784–795.

32. Nanopoulos, A.; Alcock, R.; Manolopoulos, Y. Feature-based classification of time-series data. In Information Processing and Technology; Nikos, M., Stavros, D.N., Eds.; Nova Science Publishers, Inc.: Hauppauge, NY, USA, 2001; pp. 49–61.

33. Agapoff, S.; Pache, C.; Panciatici, P.; Warland, L.; Lumbreras, S. Snapshot selection based on statistical clustering for Transmission Expansion Planning. In Proceedings of the 2015 IEEE Eindhoven PowerTech, Eindhoven, The Netherlands, 29 June–2 July 2015; pp. 1–6.

34. Pfenninger, S. Dealing with multiple decades of hourly wind and PV time series in energy models: A comparison of methods to reduce time resolution and the planning implications of inter-annual variability. Appl. Energy 2017, 197, 1–13.

35. Stenzel, P.; Linssen, J.; Fleer, J.; Busch, F. Impact of temporal resolution of supply and demand profiles on the design of photovoltaic battery systems for increased self-consumption. In Proceedings of the 2016 IEEE International Energy Conference (ENERGYCON), Leuven, Belgium, 4–8 April 2016; pp. 1–6.

36. Deane, J.P.; Drayton, G.; Gallachóir, B.Ó. The impact of sub-hourly modelling in power systems with significant levels of renewable generation. Appl. Energy 2014, 113, 152–158.

37. Beck, T.; Kondziella, H.; Huard, G.; Bruckner, T. Assessing the influence of the temporal resolution of electrical load and PV generation profiles on self-consumption and sizing of PV-battery systems. Appl. Energy 2016, 173, 331–342.

38. Yokoyama, R.; Hasegawa, Y.; Ito, K. A MILP decomposition approach to large scale optimization in structural design of energy supply systems. Energy Convers. Manag. 2002, 43, 771–790.

39. Yokoyama, R.; Shinano, Y.; Taniguchi, S.; Ohkura, M.; Wakui, T. Optimization of energy supply systems by MILP branch and bound method in consideration of hierarchical relationship between design and operation. Energy Convers. Manag. 2015, 92, 92–104.

40. Yokoyama, R.; Shinano, Y.; Wakayama, Y.; Wakui, T. Model reduction by time aggregation for optimal design of energy supply systems by an MILP hierarchical branch and bound method. Energy 2019, 181, 782–792.

41. Mavrotas, G.; Diakoulaki, D.; Florios, K.; Georgiou, P. A mathematical programming framework for energy planning in services' sector buildings under uncertainty in load demand: The case of a

hospital in Athens. Energy Policy 2008, 36, 2415–2429.

42. Fazlollahi, S.; Bungener, S.L.; Mandel, P.; Becker, G.; Maréchal, F. Multi-objectives, multi-period optimization of district energy systems: I. Selection of typical operating periods. Comput. Chem. Eng. 2014, 65, 54–66.

43. Bungener, S.; Hackl, R.; Van Eetvelde, G.; Harvey, S.; Marechal, F. Multi-period analysis of heat integration measures in industrial clusters. Energy 2015, 93, 220–234.

44. Deml, S.; Ulbig, A.; Borsche, T.; Andersson, G. The role of aggregation in power system simulation. In Proceedings of the 2015 IEEE Eindhoven PowerTech, Eindhoven, The Netherlands, 29 June–2 July 2015; pp. 1–6.

45. Pineda, S.; Morales, J.M. Chronological Time-Period Clustering for Optimal Capacity Expansion Planning With Storage. IEEE Trans. Power Syst. 2018, 33, 7162–7170.

46. Ward, J.H. Hierarchical Grouping to Optimize an Objective Function AU-Ward, Joe H. J. Am. Stat. Assoc. 1963, 58, 236–244.

47. Bahl, B.; Söhler, T.; Hennen, M.; Bardow, A. Typical Periods for Two-Stage Synthesis by Time-Series Aggregation with Bounded Error in Objective Function. Front. Energy Res. 2018, 5.

48. Lloyd, S. Least squares quantization in PCM. IEEE Trans. Inf. Theory 1982, 28, 129–137.

49. MacQueen, J. Some methods for classification and analysis of multivariate observations. In Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability; University of California Press: Berkeley, CA, USA, 1967; Volume 1, pp. 281–297.

50. Baumgärtner, N.; Temme, F.; Bahl, B.; Hennen, M.; Hollermann, D.; Bardow, A. RiSES4 Rigorous Synthesis of Energy Supply Systems with Seasonal Storage by Relaxation and Time—Series Aggregation to Typical Periods. In Proceedings of the ECOS 2019, Wroclaw, Poland, 23–28 June 2019.

51. Baumgärtner, N.; Bahl, B.; Hennen, M.; Bardow, A. RiSES3: Rigorous Synthesis of Energy Supply and Storage Systems via time-series relaxation and aggregation. Comput. Chem. Eng. 2019, 127, 127–139.

52. Stein, D.V.; Bracht, N.V.; Maaz, A.; Moser, A. Development of adaptive time patterns for multi-dimensional power system simulations. In Proceedings of the 2017 14th International Conference on the European Energy Market (EEM), Dresden, Germany, 6–9 June 2017; pp. 1–5.

53. Georgios Savvidis, K.H. How well do we understand our power system models? In Proceedings of the 42nd International Association for Energy Economics (IAEE) Annual Conference, Montréal, QC, Canada, 29 May–1 June 2019.

54. Bauer, D.; Marx, R.; Nußbicker-Lux, J.; Ochs, F.; Heidemann, W.; Müller-Steinhagen, H. German central solar heating plants with seasonal heat storage. Sol. Energy 2010, 84, 612–623.

55. Sorknæs, P. Simulation method for a pit seasonal thermal energy storage system with a heat pump in a district heating system. Energy 2018, 152, 533–538.

56. Wogrin, S.; Galbally, D.; Reneses, J. Optimizing Storage Operations in Medium-and Long-Term Power System Models. IEEE Trans. Power Syst. 2016, 31, 3129–3138.

57. Tejada-Arango, D.A.; Wogrin, S.; Centeno, E. Representation of Storage Operations in Network-Constrained Optimization Models for Medium- and Long-Term Operation. IEEE Trans. Power Syst. 2018, 33, 386–396.

58. Wogrin, S.; Dueñas, P.; Delgadillo, A.; Reneses, J. A New Approach to Model Load Levels in Electric Power Systems With High Renewable Penetration. IEEE Trans. Power Syst. 2014, 29, 2210–2218.

59. Härtel, P.; Kristiansen, M.; Korpås, M. Assessing the impact of sampling and clustering techniques on offshore grid expansion planning. Energy Procedia 2017, 137, 152–161.

60. Ploussard, Q.; Olmos, L.; Ramos, A. An operational state aggregation technique for transmission expansion planning based on line benefits. IEEE Trans. Power Syst. 2016, 32, 2744–2755.

61. Lythcke-Jørgensen, C.E.; Münster, M.; Ensinas, A.V.; Haglind, F. A method for aggregating external operating conditions in multi-generation system optimization models. Appl. Energy 2016, 166, 59–75.

62. Kotzur, L.; Markewitz, P.; Robinius, M.; Stolten, D. Impact of different time series aggregation methods on optimal energy system design. Renew. Energy 2018, 117, 474–487.

63. Buoro, D.; Casisi, M.; Pinamonti, P.; Reini, M. Optimal synthesis and operation of advanced energy supply systems for standard and domotic home. Energy Convers. Manag. 2012, 60, 96–105.

64. Harb, H.; Schwager, C.; Streblow, R.; Mueller, D. Optimal design of energy systems in residential districts WITH interconnected local heating and electrical networks. In Proceedings of the 14th International IBPSA Conference, Hyderabad, India, 7–9 December 2015.

65. Merkel, E.; McKenna, R.; Fichtner, W. Optimisation of the capacity and the dispatch of decentralised micro-CHP systems: A case study for the UK. Appl. Energy 2015, 140, 120–134.

66. de Sisternes, F.J.; Jenkins, J.D.; Botterud, A. The value of energy storage in decarbonizing the electricity sector. Appl. Energy 2016, 175, 368–379.

67. De Sisternes Jimenez, F.; Webster, M.D. Optimal Selection of Sample Weeks for Approximating the Net Load in Generation Planning Problems; Massachusetts Institute of Technology: Cambridge, MA, USA, 2013.

68. Patteeuw, D.; Helsen, L. Combined design and control optimization of residential heating systems in a smart-grid context. Energy Build. 2016, 133, 640–657.

69. van der Heijde, B.; Scapino, L.; Vandermeulen, A.; Patteeuw, D.; Helsen, L.; Salenbien, R. Using Representative Time Slices for Optimization of Thermal Energy Storage Systems in Low-Temperature District Heating Systems. In Proceedings of the ECOS 2018 31st International Conference on Efficiency, Cost, Optimization, SImulation and Environmental Impact of Energy Systems, Guimarães, Portugal, 17–22 June 2018.

70. Murty, M.N.; Jain, A.K.; Flynn, P. Data clustering: A review. ACM Comput Surv. ACM Comput. Surv. 1999, 31, 264–323.

71. Brodrick, P.G.; Brandt, A.R.; Durlofsky, L.J. Optimal design and operation of integrated solar combined cycles under emissions intensity constraints. Appl. Energy 2018, 226, 979–990.

72. Schütz, T.; Schraven, M.; Fuchs, M.; Remmen, P.; Mueller, D. Comparison of clustering algorithms for the selection of typical demand days for energy system synthesis. Renew. energy 2018, 129, 570–582.

73. Marton, C.H.; Elkamel, A.; Duever, T.A. An order-specific clustering algorithm for the determination of representative demand curves. Comput. Chem. Eng. 2008, 32, 1365–1372.

74. Lozano, M.A.; Ramos, J.C.; Serra, L.M. Cost optimization of the design of CHCP (combined heat, cooling and power) systems under legal constraints. Energy 2010, 35, 794–805.

75. Schütz, T.; Schiffer, L.; Harb, H.; Fuchs, M.; Müller, D. Optimal design of energy conversion units and envelopes for residential building retrofits using a comprehensive MILP model. Appl. Energy 2017, 185, 1–15.

76. Harb, H.; Reinhardt, J.; Streblow, R.; Mueller, D. MIP approach for designing heating systems in residential buildings and neighbourhood. J. Build. Perform. Simul. 2015.

77. Kools, L.; Phillipson, F. Data granularity and the optimal planning of distributed generation. Energy 2016, 112, 342–352.

78. Voll, P.; Klaffke, C.; Hennen, M.; Bardow, A. Automated superstructure-based synthesis and optimization of distributed energy supply systems. Energy 2013, 50, 374–388.

79. Neniškis, E.; Galinis, A. Representation of wind power generation in economic models for long-term energy planning. Energetika 2018, 64.

80. Nicolosi, M. The Importance of High Temporal Resolution in Modeling Renewable Energy Penetration Scenarios; Lawrence Berkeley National Lab.(LBNL): Berkeley, CA, USA, 2010.

81. Haydt, G.; Leal, V.; Pina, A.; Silva, C.A. The relevance of the energy resource dynamics in the mid/long-term energy planning models. Renew. Energy 2011, 36, 3068–3074.

82. Welsch, M.; Howells, M.; Bazilian, M.; DeCarolis, J.F.; Hermann, S.; Rogner, H.H. Modelling elements of Smart Grids—Enhancing the OSeMOSYS (Open Source Energy Modelling System) code. Energy 2012, 46, 337–350.

83. Timmerman, J.; Hennen, M.; Bardow, A.; Lodewijks, P.; Vandevelde, L.; Van Eetvelde, G. Towards low carbon business park energy systems: A holistic techno-economic optimisation model. Energy 2017, 125, 747–770.

84. Loulou, R.; Remne, U.; Kanudia, A.; Lehtila, A.; Goldstein, G. Documentation for the TIMES Model PART I; TIMES: London, UK, 2005.

85. Loulou, R.; Lehtilä, A.; Kanudia, A.; Remne, U.; Goldstein, G. Documentation for the TIMES Model PART II; TIMES: London, UK, 2005.

86. Wakui, T.; Kawayoshi, H.; Yokoyama, R. Optimal structural design of residential power and heat supply devices in consideration of operational and capital recovery constraints. Appl. Energy 2016, 163, 118–133.

87. Wakui, T.; Yokoyama, R. Optimal structural design of residential cogeneration systems with battery based on improved solution method for mixed-integer linear programming. Energy 2015, 84, 106–120.

88. Wakui, T.; Yokoyama, R. Optimal structural design of residential cogeneration systems in consideration of their operating restrictions. Energy 2014, 64, 719–733.

89. Lozano, M.A.; Ramos, J.C.; Carvalho, M.; Serra, L.M. Structure optimization of energy supply systems in tertiary sector buildings. Energy Build. 2009, 41, 1063–1075.

90. Weber, C.; Shah, N. Optimisation based design of a district energy system for an eco-town in the United Kingdom. Energy 2011, 36, 1292–1308.

91. Stadler, M.; Groissböck, M.; Cardoso, G.; Marnay, C. Optimizing Distributed Energy Resources and building retrofits with the strategic DER-CAModel. Appl. Energy 2014, 132, 557–567.

92. Casisi, M.; Pinamonti, P.; Reini, M. Optimal lay-out and operation of combined heat & power (CHP) distributed generation systems. Energy 2009, 34, 2175–2183.

93. Pina, A.; Silva, C.A.; Ferrão, P. High-resolution modeling framework for planning electricity systems with high penetration of renewables. Appl. Energy 2013, 112, 215–223.

94. Spiecker, S.; Vogel, P.; Weber, C. Evaluating interconnector investments in the north European electricity system considering fluctuating wind power penetration. Energy Econ. 2013, 37, 114–127.

95. Wouters, C.; Fraga, E.S.; James, A.M. An energy integrated, multi-microgrid, MILP (mixed-integer linear programming) approach for residential distributed energy system planning—A South Australian case-study. Energy 2015, 85, 30–44.

96. Yang, Y.; Zhang, S.; Xiao, Y. Optimal design of distributed energy resource systems coupled with energy distribution networks. Energy 2015, 85, 433–448.

97. Ameri, M.; Besharati, Z. Optimal design and operation of district heating and cooling networks with CCHP systems in a residential complex. Energy Build. 2016, 110, 135–148.

98. Bracco, S.; Dentici, G.; Siri, S. DESOD: A mathematical programming tool to optimally design a distributed energy system. Energy 2016, 100, 298–309.

99. Moradi, S.; Ghaffarpour, R.; Ranjbar, A.M.; Mozaffari, B. Optimal integrated sizing and planning of hubs with midsize/large CHP units considering reliability of supply. Energy Convers. Manag. 2017, 148, 974–992.

100. Swider, D.J.; Weber, C. The costs of wind's intermittency in Germany: Application of a stochastic electricity market model. Eur. Trans. Electr. Power 2007, 17, 151–172.

101. Rosen, J. The Future Role of Renewable Energy Sources in European Electricity Supply: A Model-Based Analysis for the EU-15; KIT Scientific Publishing: Karlsruhe, Germany, 2008.

102. Oluleye, G.; Vasquez, L.; Smith, R.; Jobson, M. A multi-period Mixed Integer Linear Program for design of residential distributed energy centres with thermal demand data discretisation. Sustain. Prod. Consum. 2016, 5, 16–28.

103. Poncelet, K.; Delarue, E.; Duerinck, J.; Six, D.; D'haeseleer, W. The Importance of Integrating the Variability of Renewables in Long-term Energy Planning Models; TME: Rome, Italy, 2014.

104. Mallapragada, D.S.; Papageorgiou, D.J.; Venkatesh, A.; Lara, C.L.; Grossmann, I.E. Impact of model resolution on scenario outcomes for electricity sector system expansion. Energy 2018, 163, 1231–1244.

105. Pina, A.; Silva, C.; Ferrão, P. Modeling hourly electricity dynamics for policy making in long-term scenarios. Energy Policy 2011, 39, 4692–4702.

106. Devogelaer, D. Towards 100% Renewable Energy in Belgium by 2050; FPB: Brussels, Belgium, 2012.

107. Kannan, R.; Turton, H. A Long-Term Electricity Dispatch Model with the TIMES Framework. Environ. Model. Assess. 2013, 18, 325–343.

108. Simões, S.; Nijs, W.; Ruiz, P.; Sgobbi, A.; Radu, D.; Yilmaz Bolat, P.; Thiel, C.; Peteves, E. The JRC-EU-TIMES model—Assessing the long-term role of the SET Plan Energy technologies. JRC's Inst. Energy Transport Tech. Rep. 2013.

109. Poncelet, K.; Delarue, E.; Six, D.; Duerinck, J.; D'haeseleer, W. Impact of the level of temporal and operational detail in energy-system planning models. Appl. Energy 2016, 162, 631–643.

110. Merrick, J.H. On representation of temporal variability in electricity capacity planning models. Energy Econ. 2016, 59, 261–274.

111. Mehleri, E.D.; Sarimveis, H.; Markatos, N.C.; Papageorgiou, L.G. A mathematical programming approach for optimal design of distributed energy systems at the neighbourhood level. Energy 2012, 44, 96–104.

112. Mehleri, E.D.; Sarimveis, H.; Markatos, N.C.; Papageorgiou, L.G. Optimal design and operation of distributed energy systems: Application to Greek residential sector. Renew. Energy 2013, 51, 331–342.

113. Domínguez-Muñoz, F.; Cejudo-López, J.M.; Carrillo-Andrés, A.; Gallardo-Salazar, M. Selection of typical demand days for CHP optimization. Energy Build. 2011, 43, 3036–3043.

114. Aghabozorgi, S.; Seyed Shirkhorshidi, A.; Ying Wah, T. Time-series clustering—A decade review. Inf. Syst. 2015, 53, 16–38.

115. Beyer, K.; Goldstein, J.; Ramakrishnan, R.; Shaft, U. When Is "Nearest Neighbor" Meaningful? In Proceedings of the Database Theory—ICDT'99; Beeri, C., Buneman, P., Eds.; Springer: Berlin/Heidelberg, Germany, 1999; pp. 217–235.

116. Keogh, E.; Mueen, A. Curse of Dimensionality. In Encyclopedia of Machine Learning; Sammut, C., Webb, G.I., Eds.; Springer US: Boston, MA, USA, 2010; pp. 257–258.

117. Aggarwal, C.C.; Hinneburg, A.; Keim, D.A. On the Surprising Behavior of Distance Metrics in High Dimensional Space. In Proceedings of the Database Theory—ICDT 2001; Van den Bussche, J., Vianu, V., Eds.; Springer Berlin Heidelberg: Berlin/Heidelberg, Germany, 2001; pp. 420–434.

118. Guo, X.; Gao, L.; Liu, X.; Yin, J. Improved Deep Embedded Clustering with Local Structure Preservation. In Proceedings of the 26th International Joint Conference on Artificial Intelligence; AAAI Press: Melbourne, Australia, 2017; pp. 1753–1759.

119. Sun, M.; Teng, F.; Zhang, X.; Strbac, G.; Pudjianto, D. Data-Driven Representative Day Selection for Investment Decisions: A Cost-Oriented Approach. IEEE Trans. Power Syst. 2019, 34, 1.

120. Sun, M.; Konstantelos, I.; Strbac, G. C-Vine Copula Mixture Model for Clustering of Residential Electrical Load Pattern Data. IEEE Trans. Power Syst. 2017, 32, 2382–2393.

121. Almaimouni, A.; Ademola-Idowu, A.; Nathan Kutz, J.; Negash, A.; Kirschen, D. Selecting and Evaluating Representative Days for Generation Expansion Planning. In 2018 Power Systems Computation Conference; IEEE: Piscataway Township, NJ, USA, 2018; pp. 1–7.

122. Fazlollahi, S.; Girardin, L.; Maréchal, F. Clustering Urban Areas for Optimizing the Design and the Operation of District Energy Systems. In Computer Aided Chemical Engineering; Klemeš, J.J., Varbanov, P.S., Liew, P.Y., Eds.; Elsevier: Amsterdam, The Netherlands, 2014; Volume 33, pp. 1291–1296.

123. Schütz, T.; Schraven, M.; Harb, H.; Fuchs, M.; Mueller, D. Clustering Algorithms for the Selection of Typical Demand Days for the Optimal Design of Building Energy Systems. In Proceedings of

the ECOS 2016: 29th International Conference on Efficiency, Cost, Optimization, Simulation, and Environmental Impact of Energy Systems, Portoroz, Slovenia, 16–23 June 2016.

124. Kannengießer, T.; Hoffmann, M.; Kotzur, L.; Stenzel, P.; Schuetz, F.; Peters, K.; Nykamp, S.; Stolten, D.; Robinius, M. Reducing Computational Load for Mixed Integer Linear Programming: An Example for a District and an Island Energy System. Energies 2019, 12, 2825.

125. Zatti, M.; Gabba, M.; Freschini, M.; Rossi, M.; Gambarotta, A.; Morini, M.; Martelli, E. k-MILP: A novel clustering approach to select typical and extreme days for multi-energy systems design optimization. Energy 2019, 181, 1051–1063.

126. Fitiwi, D.Z.; de Cuadra, F.; Olmos, L.; Rivier, M. A new approach of clustering operational states for power network expansion planning problems dealing with RES (renewable energy source) generation operational variability and uncertainty. Energy 2015, 90, 1360–1376.

127. Nahmmacher, P.; Schmid, E.; Hirth, L.; Knopf, B. Carpe diem: A novel approach to select representative days for long-term power system modeling. Energy 2016, 112, 430–442.

128. Poncelet, K.; Höschle, H.; Delarue, E.; Virag, A.; D'haeseleer, W. Selecting Representative Days for Capturing the Implications of Integrating Intermittent Renewables in Generation Expansion Planning Problems. IEEE Trans. Power Syst. 2016, 32, 1936–1948.

129. Voulis, N.; Warnier, M.; Brazier, F.M.T. Understanding spatio-temporal electricity demand at different urban scales: A data-driven approach. Appl. Energy 2018, 230, 1157–1171.

130. Teichgraeber, H.; Brandt, A.R. Clustering methods to find representative periods for the optimization of energy systems: An initial framework and comparison. Appl. Energy 2019, 239, 1283–1293.

131. Brodrick, P.G.; Brandt, A.R.; Durlofsky, L.J. Operational optimization of an integrated solar combined cycle under practical time-dependent constraints. Energy 2017, 141, 1569–1584.

132. Teichgraeber, H.; Brodrick, P.G.; Brandt, A.R. Optimal design and operations of a flexible oxyfuel natural gas plant. Energy 2017, 141, 506–518.

133. Liu, Y.; Sioshansi, R.; Conejo, A.J. Hierarchical Clustering to Find Representative Operating Periods for Capacity-Expansion Modeling. IEEE Trans. Power Syst. 2018, 33, 3029–3039.

134. Marquant, J.; Omu, A.; Evins, R.; Carmeliet, J. Application of Spatial-Temporal Clustering to Facilitate Energy System Modelling. In 14th International Confrence of IBPSA Building Simulation 2015; Khare, V.R., Gaurav, C., Eds.; IIIT Hyderabad: Hyderabad, India, 2015; pp. 551–558.

135. Marquant, J.F.; Mavromatidis, G.; Evins, R.; Carmeliet, J. Comparing different temporal dimension representations in distributed energy system design models. Energy Procedia 2017, 122, 907–912.

136. Jain, A.K. Data clustering: 50 years beyond K-means. Pattern Recognit. Lett. 2010, 31, 651–666.

137. Saxena, A.; Prasad, M.; Gupta, A.; Bharill, N.; Patel, O.P.; Tiwari, A.; Er, M.J.; Ding, W.; Lin, C.-T. A review of clustering techniques and developments. Neurocomputing 2017, 267, 664–681.

138. Adhau, S.P.; Moharil, R.M.; Adhau, P.G. K-Means clustering technique applied to availability of micro hydro power. Sustain. Energy Technol. Assess. 2014, 8, 191–201.

139. Green, R.; Staffell, I.; Vasilakos, N. Divide and Conquer? k-Means Clustering of Demand Data Allows Rapid and Accurate Simulations of the British Electricity System. IEEE Trans. Eng. Manag. 2014, 61, 251–260.

140. Brodrick, P.G.; Kang, C.A.; Brandt, A.R.; Durlofsky, L.J. Optimization of carbon-capture-enabled coal-gas-solar power generation. Energy 2015, 79, 149–162.

141. Lin, F.; Leyffer, S.; Munson, T. A two-level approach to large mixed-integer programs with application to cogeneration in energy-efficient buildings. Comput. Optim. Appl. 2016, 65, 1–46.

142. Bahl, B.; Kümpel, A.; Seele, H.; Lampe, M.; Bardow, A. Time-series aggregation for synthesis problems by bounding error in the objective function. Energy 2017, 135, 900–912.

143. Heuberger, C.F.; Staffell, I.; Shah, N.; Dowell, N.M. A systems approach to quantifying the value of power generation and energy storage technologies in future electricity networks. Comput. Chem. Eng. 2017, 107, 247–256.

144. Teichgraeber, H.; Brandt, A.R. Systematic Comparison of Aggregation Methods for Input Data Time Series Aggregation of Energy Systems Optimization Problems. In Computer Aided Chemical Engineering; Eden, M.R., Ierapetritou, M.G., Towler, G.P., Eds.; Elsevier: Amsterdam, The Netherlands, 2018; Volume 44, pp. 955–960.

145. Gabrielli, P.; Fürer, F.; Mavromatidis, G.; Mazzotti, M. Robust and optimal design of multi-energy systems with seasonal storage through uncertainty analysis. Appl. Energy 2019, 238, 1192–1210.

146. Zhang, H.; Lu, Z.; Hu, W.; Wang, Y.; Dong, L.; Zhang, J. Coordinated optimal operation of hydro–wind–solar integrated systems. Appl. Energy 2019, 242, 883–896.

147. Arthur, D.; Vassilvitskii, S. k-means++: The advantages of careful seeding. In Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms; Society for Industrial and Applied Mathematics: Philadelphia, PA, USA, 2007; pp. 1027–1035.

148. Schiefelbein, J.; Tesfaegzi, J.; Streblow, R.; Müller, D. Design of an optimization algorithm for the distribution of thermal energy systems and local heating networks within a city district. Proc. Ecos 2015.

149. Zhu, Q.; Luo, X.; Zhang, B.; Chen, Y. Mathematical modelling and optimization of a large-scale combined cooling, heat, and power system that incorporates unit changeover and time-of-use electricity price. Energy Convers. Manag. 2017, 133, 385–398.

150. Stadler, P.; Girardin, L.; Ashouri, A.; Maréchal, F. Contribution of Model Predictive Control in the Integration of Renewable Energy Sources within the Built Environment. Front. Energy Res. 2018, 6.

151. Hilbers, A.P.; Brayshaw, D.J.; Gandy, A. Importance subsampling: Improving power system planning under climate-based uncertainty. Appl. Energy 2019, 251, 113114.

152. Tupper, L.L.; Matteson, D.S.; Anderson, C.L.; Zephyr, L. Band Depth Clustering for Nonstationary Time Series and Wind Speed Behavior. Technometrics 2018, 60, 245–254.

153. Kaufman, L.; Rousseeuw, P.J. Clustering by means of medoids. Statistical Data Analysis based on the L1 Norm. Y. DodgeEd 1987, 405–416.

154. Vinod, H. Integer Programming and the Theory of Grouping. J. Am. Stat. Assoc. 1969, 64.

155. Singh, A.; Yadav, A.; Rana, A. K-means with Three different Distance Metrics. Int. J. Comput. Appl. 2013, 67.

156. Bradley, P.S.; Mangasarian, O.L.; Street, W.N. Clustering via concave minimization. In Advances in Neural Information Processing Systems; Mit Press: Cambridge, MA, USA, 1997; pp. 368–374.

157. Whelan, C.; Harrell, G.; Wang, J. Understanding the K-Medians Problem. In Proceedings of the International Conference on Scientific Computing (CSC), The Steering Committee of The World Congress in Computer Science, Computer, San Diego, CA, USA, 27–30 July 2015; p. 219.

158. Har-Peled, S. Geometric Approximation Algorithms; American Mathematical Soc.: Providence, RI USA, 2006.

159. Paparrizos, J.; Gravano, L. k-Shape: Efficient and Accurate Clustering of Time Series. Sigmod Rec. 2016, 45, 69–76.

160. Petitjean, F.; Ketterlin, A.; Gançarski, P. A global averaging method for dynamic time warping, with applications to clustering. Pattern Recognit. 2011, 44, 678–693.

161. Niennattrakul, V.; Srisai, D.; Ratanamahatana, C. Shape-based template matching for time series data. Knowl.-Based Syst. 2012, 26.

162. Balachandra, P.; Chandru, V. Modelling electricity demand with representative load curves. Energy 1999, 24, 219–230.

163. Poncelet, K.; Höschle, H.; Delarue, E.; D'haeseleer, W. Selecting Representative Days for Investment Planning Models; kU Leuven: Leuven, Belgium, 2015.

164. Tveit, T.-M.; Savola, T.; Gebremedhin, A.; Fogelholm, C.-J. Multi-period MINLP model for optimising operation and structural changes to CHP plants in district heating networks with long-term thermal storage. Energy Convers. Manag. 2009, 50, 639–647.

165. Loulou, R.; Kanudia, A.; Goldstein, G. Documentation for the times model part ii. Energy Technol. Syst. Anal. Programme. 2016.

166. Loulou, R.; Goldstein, G.; Kanudia, A.; Lettila, A.; Remne, U. Documentation for the TIMES Model PART I; TIMES: London, UK, 2016.

167. Ortiga, J.; Bruno, J.C.; Coronas, A. Selection of typical days for the characterisation of energy demand in cogeneration and trigeneration optimisation models for buildings. Energy Convers. Manag. 2011, 52, 1934–1942.

168. van der Weijde, A.H.; Hobbs, B.F. The economics of planning electricity transmission to accommodate renewables: Using two-stage optimisation to evaluate flexibility and the cost of disregarding uncertainty. Energy Econ. 2012, 34, 2089–2101.

169. Munoz, F.D.; Mills, A.D. Endogenous Assessment of the Capacity Value of Solar PV in Generation Investment Planning Studies. IEEE Trans. Sustain. Energy 2015, 6, 1574–1585.

170. Frew, B.A.; Jacobson, M.Z. Temporal and spatial tradeoffs in power system modeling with assumptions about storage: An application of the POWER model. Energy 2016, 117, 198–213.

171. Lee, T.-Y.; Chen, C.-L. Unit commitment with probabilistic reserve: An IPSO approach. Energy Convers. Manag. 2007, 48, 486–493.

172. Phan, Q.A.; Scully, T.; Breen, M.; Murphy, M.D. Determination of optimal battery utilization to minimize operating costs for a grid-connected building with renewable energy sources. Energy Convers. Manag. 2018, 174, 157–174.

173. Saravanan, B.; Das, S.; Sikri, S.; Kothari, D.P. A solution to the unit commitment problem—A review. Front. Energy 2013, 7, 223–236.

174. Xiao, J.; Bai, L.; Li, F.; Liang, H.; Wang, C. Sizing of Energy Storage and Diesel Generators in an Isolated Microgrid Using Discrete Fourier Transform (DFT). IEEE Trans. Sustain. Energy 2014, 5, 907–916.

175. Pöstges, A.; Weber, C. Time series aggregation—A new methodological approach using the "peak-load-pricing" model. Util. Policy 2019, 59, 100917.