



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



Stefan Pfenninger

Climate Policy Group, Department of Environmental Systems Science, ETH Zürich, Switzerland

HIGHLIGHTS

- Comparison of resampling, clustering, and heuristics to reduce model time resolution.
- Suitable approach depends on input data and model constraint setup.
- Heuristic approaches appear more stable than statistical clustering.
- Results with high renewable shares but few years of input data are unreliable.
- Better modeling and planning methods needed to deal with inter-year variability.

ARTICLE INFO

Article history:

Received 5 October 2016
Received in revised form 3 February 2017
Accepted 10 March 2017

Keywords:

High-resolution energy modeling
Time series data
Variable renewable generation
Modeling methods

ABSTRACT

Using a high-resolution planning model of the Great Britain power system and 25 years of simulated wind and PV generation data, this study compares different methods to reduce time resolution of energy models to increase their computational tractability: downsampling, clustering, and heuristics. By comparing model results in terms of costs and installed capacities across different methods, this study shows that the best method depends heavily on input data and the setup of model constraints. This implies that there is no one-size-fits-all approach to the problem of time step reduction, but heuristic approaches appear promising. In addition, the 25 years of time series demonstrate considerable inter-year variability in wind and PV power output. This further complicates the problem of time detail in energy models as it suggests long time series are necessary. Model results with high shares of PV and wind generation using a single or few years of data are likely unreliable. Better modeling and planning methods are required to determine robust scenarios with high shares of variable renewables. The methods are implemented in the freely available open-source modeling framework Calliope.

Copyright © 2017 The Author. Published by Elsevier Ltd. Open Access Article licensed CC-BY 4.0 (<https://creativecommons.org/licenses/by/4.0/>)

1. Introduction

Energy system models were first developed in the 1970s by the International Energy Agency (IEA) and the International Institute for Applied Systems Analysis (IIASA) in the aftermath of the international oil crisis. Using optimization methods, in particular linear programming, they allowed analysts to structure their assumptions and data, forming them into internally coherent scenarios of how energy is extracted, converted, transported, and used, and how these processes might change in the future. Today, with the increasing deployment of variable renewable generation, the global energy system is again undergoing a fundamental transition. Global installed wind power capacity reached about 417 GW in

2015, up from 17 GW in 2000, while solar photovoltaics (PV) has experienced an even higher growth rate, with capacity rising from below 1 GW in 2000 to 222 GW in 2015 globally [1]. Energy models are important decision-making aids to help navigate the transformation of the current fossil-fuel based energy system to one based on clean and renewable energy [2].

In this context, the rising importance of variable renewable generation has presented two crucial and related problems to energy modelers. The first problem is procuring data on the generation potential for wind and PV power with sufficient resolution in space and time, then integrating this data into power system models such as LIMES-EU [3] or larger energy system models such as TIMES and TIAM [4]. Having data with temporal resolution of one hour or better allows a model to depict the hour-by-hour and day-by-day fluctuations in power output from

E-mail address: stefan.pfenninger@usys.ethz.ch

these technologies, how they correlate with each other and with power demand [5]. The second problem pertains to the inter-annual variability of renewable generation, which requires many years of data to fully address. Recent work has started to address the first problem but in the main, studies are limited to a single or a small number of years [6–9]. The provision of longer time series requires input data ideally spanning multiple decades. This is becoming possible through the use of global reanalysis data for energy modeling [10,11].

Both of these problems overlap in one crucial area: for large models to be computationally tractable, it is often not feasible for them to include full hourly time series for an entire year, let alone for several decades. This is primarily due to the computational requirements of running what are often linear or mixed-integer optimization models. Work has started to emerge investigating ways to reduce the number of time steps in models while retaining relevant detail, primarily by using statistical clustering methods to derive a set of typical days or weeks from a larger input time series, then feeding those days into an energy model [12–15]. However, there are well-known limitations to statistical clustering methods, including the fact that most methods produce clusters even in homogeneous datasets, and that clusters must be validated and their stability assessed [16]. This implies that approaches may vary in performance in different years, and that specific approaches are more or less suitable depending on the structure of the underlying model.

This paper treats this problem systematically, by examining different techniques to reduce the time resolution of energy models and their impact on model performance and results, and by doing so, answers two questions. First, how accurate are different methods to reduce time resolution when compared on the same model and with different model configurations? Second, how can time resolution be reduced in the most efficient way while maintaining scientific accuracy? The rationale when reducing the number of time steps is to balance improved computational performance with model accuracy. This study compares different approaches to achieve this, including downsampling, statistical clustering methods, and heuristic selection of specific days and weeks. Wind and solar generation show significant inter-annual variability, so multi-decade time series data should ideally be used to represent the full range of this variability [11,10]. This further complicates the answers to the two questions posed above: when considering power systems with very high shares of variable renewable generation, it is necessary to also consider whether the differences in accuracy between different time resolution reduction approaches persist across different years or when examining multi-decade time series. The analysis is performed with a model of the UK power system based on the open-source Calliope high-resolution modeling framework [8].

2. Background

Table 1 gives a stylized overview of the ways in which temporal resolution has been included in energy models in order of increasing detail (spatial resolution is included for completeness, but this paper focuses on temporal resolution). With the displacement of traditional power generation by variable renewables expected to increase, energy modelers have been moving downwards in this table from lower to higher resolutions. Approaches such as average availabilities for technologies, which were sufficient when modeling baseload or completely dispatchable generators such as coal or nuclear power, have been replaced with more explicit treatment of time. However, this more explicit treatment comes at a computational cost. Assuming a model with a single year of 8760 hourly time steps, 20 technologies (such as wind generation or electric

heating), 20 locations and 5 time-dependent constraints (such as maximum power generation per location, storage charging, and discharging), more than 17 million total constraints would result. Reducing such a model's size by one or two orders of magnitude by reducing the number of time steps brings with it a concomitant reduction in computational complexity, and thus in required CPU time and memory requirements to solve it.

There are broadly two ways to explicitly include temporal detail in energy models without including full time series data. The first is time slices, that is, a reduced set of time steps chosen to characterize key aspects of temporal variability, for example by covering weekdays and weekends, different times of day, and different seasons. For example, four seasons with four times per day for both weekdays and weekends would result in 32 time slices. Large energy system models such as TIMES generally use time slices e.g., [19,20]. The second way is representative of typical days (or longer time periods) extracted or artificially constructed from full time series. Typical days are an intermediate step on the path towards full time series: selecting a number of specific days from the data covering as much variability as possible, or constructing synthetic days by clustering the data. In both cases the goal is to preserve the relevant statistical properties of the time series and thus minimize impact on model results.

Two problems arise when reducing the resolution of model input time series: concurrency and continuity. It is important that correlation between events is realistic in an energy model, for example, a stormy winter day may induce higher heating demand, reduce PV power output to almost zero, but provide above-average wind power production. A concurrency problem arises when a subset of typical heat demand days are mixed with a subset of typical wind production days and such correlations are lost. This can be circumvented by ensuring that an internally consistent set of input data is used across the entire model. The second problem is that of continuity and appears when there is a state in a model that needs to carry over from one time step to the next. The prime example of this is the state of charge of storage facilities. Ensuring continuity between time steps can be difficult to address when picking representative days, so recent approaches have often used groups of consecutive days [14]. Selecting days should take into consideration a statistical measure of representativeness, as a selection based on a typical (e.g. seasonal basis) has shown to be inferior for high shares of variable renewables [21].

Statistical clustering is a way to group samples into groups called clusters, such that the similarity of samples within a cluster is higher than between clusters. It can be used to select representative time periods, and has been applied in particular when studying demand profiles and to estimate load profiles where data is limited. Rhodes et al. [22] use the k-means clustering technique to group homes with similar hourly electricity demand profiles and use regression to determine which variables influence demand in a cluster. Similar clustering approaches for electricity demand were also used in Räsänen et al. [23] and McLoughlin et al. [24]. Green et al. [12] use k-means clustering for power demand in a mixed-integer Great Britain (GB) power system dispatch model written in GAMS, for one year at a time over the 12 years in the period 1994–2005, reporting no more than around 1% model error in system-wide power costs but a model speedup of a factor of about 60 when using their clustered demand data. Clustering has also been used for more than just demand data. Heuberger et al. [15] use the k-means clustering approach reported in Green et al. [12] for demand and for wind and solar production in a cost-optimizing mixed-integer GB power system planning model, reporting about a 0.6% error in system-wide cost results and a 4% error in technology-specific costs when comparing to unclustered data. Baringo and Conejo [25] compare the use of load duration curves and k-means clustering for an efficient representation of

Table 1

Main approaches to include temporal detail in energy models with an indication of the order of magnitude of resolution in time. LDCs = load duration curves. Time step resolution is typically hourly or higher for time slices, typical days, or full time series.

	Time steps	Time resolution	Spatial correlation	Temporal correlation	Example
Average availabilities/ LDCs	1–10	Monthly-yearly	No	No	LEAP [17]
Time slices	10–100	Hourly-daily	No	No	TIMES [4]
Typical days or weeks	100–1000	Hourly	Method-dependent	Method-dependent	LIMES-EU [3]
Full time series	>1000	Hourly and better	Yes	Yes	RREOM [18]

the spatial correlation between wind generation and power demand for investment planning purposes, finding that k-means clustering is more accurate than load duration curves, as it can depict the correlation between different locations. Mena et al. [26] use hierarchical clustering to speed up a differential evolution algorithm for the optimal design of a distributed renewable generation system. Similarly, Nahmacher et al. [14] use hierarchical clustering of time steps using input data from the ERA-Interim reanalysis for 1979–2011 for the long-term European power system planning model LIMES-EU, a cost-minimizing linear optimization model. They find that 6 representative days covered by 48 3-hourly time slices are sufficient when comparing to 100 representative days, with total system costs just 4% lower.

Clustering is not the only possible method to reduce time resolution and other work has taken different approaches. For example, Poncelet et al. [13] present an optimization approach to selecting time steps. Samsatli and Samsatli [27] hierarchically decompose the complete time series into slices of different granularity, in a refinement of the time slice approach used in models like TIMES. Hsu [28] summarizes recent work on clustering to identify groups of electricity consumers, and is the only recent study to assess two different methods with respect to prediction accuracy and cluster stability: k-means and clusterwise regression. He finds that k-means clustering results in relatively stable clusterings, but poorer predictive accuracy compared to clusterwise regression, concluding that given this varying accuracy, appropriate clustering methods must be identified for particular use cases. Pappas [29] argues that for energy systems, reducing a problem to typical states, such as an average winter evening demand, is an oversimplification, and that by using typical states, a model ends up underestimating the energy system's true costs. On the other hand, Ludwig et al. [30] showed that with higher time step resolution, a long-term energy model for part of the German power grid computes higher overall system costs, but the effect of this on climate mitigation costs was negligible. One can conclude that the necessity to use hourly rather than lower-resolved (6, 12, and 24-hourly) time series data when considering high shares of renewables depends on the research question being answered.

While there is a recent wealth of work using various methods to reduce time resolution in energy models, this work leaves open the question of appropriateness of different methods for different model configurations and whether some approaches could achieve accuracy with less detail. Most past work does not motivate the choice of a specific method of reducing time resolution versus the many available options. There are therefore two open questions. First, how accurate are different methods to reduce time resolution when compared on the same model and with different model configurations? Second, how can time resolution be reduced in the most efficient way, balancing accuracy and data? Also, most past work has not assessed systems with very high shares of variable renewable generation, but such systems are vulnerable to the significant inter-annual variability of the renewable resource [10]. Thus it is also necessary to consider whether differences in accuracy between time resolution reduction approaches persist across different years or when examining multi-decade time series. This final point leads from modeling problems (dealing with data

volume and model size) to planning problems (planning power systems with high shares of renewables given the inter-year weather variability and resulting wind and PV power variability).

3. Methods

3.1. Model and data

This analysis uses a cost-minimizing linear optimization model that can determine installed capacity, power plant dispatch, transmission line capacity and use between model zones using net transfer capacities, line losses, storage capacity and charging/discharging, based on the Calliope energy modeling framework [8]. The model determines installed capacities while simultaneously establishing the hour-by-hour operation of all plants. In all model runs, plant capacity is decided simultaneously for all units without consideration of different planning horizons or deployment over time. The basic structure of Calliope was derived from the power nodes modeling framework [31] and uses a model formulation based on nodes defined by a set of technologies and locations, with nodes able to supply, store, transmit, or demand energy depending on the constraints specified by the defined technologies. The model balances electricity supply and demand across the modeled system. The structure is described in more detail in Pfenninger and Keirstead [8] and in the Calliope documentation.¹ Calliope reports leveled costs of electricity (LCOE) for each technology by accounting for both construction and operational costs. System-wide LCOE are computed by a production-weighted average of all technology-specific LCOEs, including those of transmission and storage technologies. The objective function is to minimize total system cost, so all decisions, including power plant dispatch or storage operation, are taken from the perspective of a central planner with perfect information.

The model used for this study is based on a modified version of the 20-zone GB power system model from Pfenninger and Keirstead [8] (see Fig. 1). The zones are based on the National Grid transmission system [32], with transmission constraints between the zones derived from data on power flows across their boundaries [33]. The technologies considered are onshore wind, offshore wind, PV, hydro, pumped hydro and battery storage. To simplify the model used, non-renewable generation is represented by two technologies: baseload (with higher capital costs and lower operating costs) and dispatch (with lower capital but higher operating costs). No capacity additions for hydro or pumped hydro are allowed. These simplifications allow focusing on the key characteristics of interest for this study: high shares of wind and PV power and the effect of different methods to reduce time step resolution on model results.

The Calliope framework was designed for high resolution data in space and time. This analysis uses high resolution in time, with 25 years of hourly time series data for wind and PV generation. These data are obtained using the extensively validated Renewables.ninja PV and wind simulation methods [10,11]. Wind power

¹ calliope.readthedocs.io.

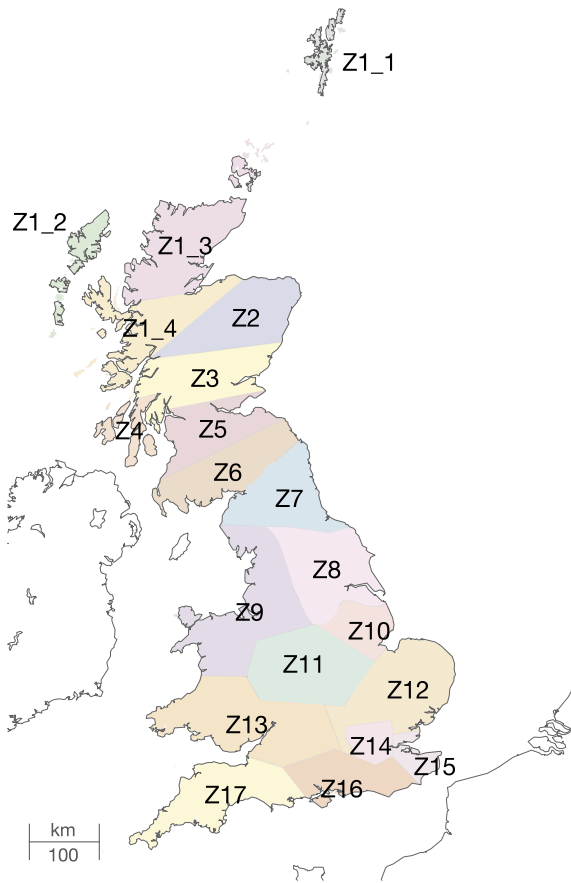


Fig. 1. Model zones; from Pfenninger and Keirstead [8].

Table 2
Methods to reduce time resolution.

Method	Key parameters	Normalized	Weighted
Downsampling	Resolution	Yes	No
Heuristic selection	Selection method; resample or drop	Yes	Yes
K-means clustering	Defining observations; No. of clusters; centroids or closest days	Yes	Yes
Hierarchical clustering	Defining observations; No. of clusters; centroids or closest days	Yes	Yes
Combination of methods	Chosen methods and their parameters	Yes	Yes

gives an overview of the different approaches used to reduce time resolution. The most simple approach is downsampling, where the entire time series is simply downsampled to a lower resolution (e.g. from 1-hourly to 3-hourly). Heuristic selection is the selection of days or full calendar weeks based on criteria such as the week containing the maximum or minimum daily average for the given time series, or the relative difference between time series being maximal or minimal. Two clustering methods are compared: hierarchical and k-means clustering, described in more detail below.

These methods are combined in various ways to yield a total of 42 time resolution reduction approaches, some of which are single methods, others a combination of multiple methods. A full list of the approaches and their parameters is given in Table 3. These approaches are chosen to test a wide range of methods on the same model. They are assessed for their ability to deliver accurate results with as few time steps as possible. The process of applying any of the 42 approaches is as follows. Before applying clustering or any of the heuristic methods, time series are normalized by the maximum value across all time steps and model zones (for downsampling, no normalization is performed). If a heuristic selection is applied, the remaining (unchosen by the heuristic) data can be either removed, downsampled or clustered, depending on the configuration of the chosen approach. The clustering methods are described below. After applying downsampling, heuristic selection, clustering, or a combination of these, a weight is determined for each remaining time step according to the number of original time steps it represents (i.e. a downsampled 6-hourly time step gets a weight of 6). The final time series is then re-scaled to match the mean of the original time series. This overall approach is similar to that presented in Nahmmacher et al. [14].

Two clustering methods are compared. The first is hierarchical clustering using Ward's method as in Nahmmacher et al. [14]. The clustering algorithm determines clusters by recursively starting to group observations together based on Euclidian distance, stopping once a pre-determined maximum distance or a pre-determined number of clusters is reached. For this study, it is always the number of clusters that was fixed, as outlined below, and to match the same number of clusters as used for the second clustering method, k-means. Clustering with k-means also uses Euclidian distance, but with a heuristic algorithm. To assess the stability of clusters given the initial random selection of cluster centers in the k-means algorithm, the algorithm was run 1000 times for 5 clusters using hourly 2014 data. The resulting cluster labeling was compared with the chance-adjusted Rand Index [39] and Adjusted Mutual Information [40], both of which had a mean of 0.90 and a standard deviation of 0.041 and 0.039 respectively, suggesting that clustering is sufficiently stable.

Examining the sum of squared error (SSE) between the cluster centroid and all cluster members as a function of the number of clusters (the elbow method) to test 1–30 clusters on the 2014 data, between 10 and 15 clusters was found to be the point where SSE

is obtained with the Virtual Wind Farm model [11] by extrapolating wind speeds from the NASA MERRA reanalysis [34] to a hub height of 80 m, then simulating power production using aggregated power curves of the UK's five most common turbine models [35]. 1 MW offshore and onshore farms are simulated at each offshore/onshore MERRA grid point [8] and the resulting time series are averaged to the 20 model zones. PV production is calculated with the Global Solar Energy Estimator model [10], which uses an indirect irradiance estimation based on Ridley et al. [36] and a PV performance model based on Huld et al. [37] to simulate PV production from southwards facing panels tilted at 35 degrees in each onshore MERRA grid point. Demand data from National Grid for the years 2002–2013 are included, based on 30-min National Grid data [38], and disaggregated spatially into the model zones by using total annual demand data per local authority and assigning each local authority to one of the 20 zones. This assumes that demand shape is the same in each zone. For years before 2002, where no demand data is available, 2002 data is used. The complete model code and data are made available online (see [supplementary material](#)).

3.2. Methods to reduce time resolution

For the work presented here, the Calliope framework was extended with implementations of all the time resolution reduction methods compared. Time series data is stored at 1-hourly resolution, and any adjustments are performed in-memory based on specifications in Calliope's model configuration before the model is solved. This makes it possible to reproduce the changes made to input data and to adjust those changes for future model runs without losing the higher detail contained in the input data. Table 2

Table 3
Time resolution reduction methods used for single-year models.

Name	Approach	Num.	Selection
Timesteps	Downsampling	4	12, 6, 3, 1 h
k-means	Clustering	8	Closest/mean; 5, 10, 15, 20 days
k-means - wind_offshore	Clustering	6	Closest/mean; 5, 10, 15 days
Hierarchical	Clustering	8	Closest/mean; 5, 10, 15, 20 days
Min/max solar and wind days	Heuristic	1	Drop
Min/max solar and wind days – downsample	Combination	3	Downsample 6 h, 12 h, 24 h
Min/max solar and wind days – kmeans	Combination	4	Closest/mean; 5, 10 days
Min/max solar and wind days – Hierarchical	Combination	4	Closest/mean; 5, 10 days
Min/max wind weeks	Heuristic	1	Drop
Min/max wind and pv weeks	Heuristic	1	Drop
Min/max wind-demand weeks	Heuristic	1	Drop
Min/max wind-demand weeks – downsample	Combination	1	12 h
Min/max wind-demand weeks – kmeans	Combination	1	Mean; 10 days

flattens off. 5, 10, 15, and 20 clusters were therefore selected for the full analysis. In addition, the impact of the normalization used was assessed by comparing load duration curves (LDCs) for 5, 10, 15, and 20 k-means-derived clusters for the 2014 data. An LDC is a cumulative frequency diagram with an inverted x -axis, showing the distribution over all time steps of capacity factors. Capacity factor is the ratio of power generation to hypothetical power generation if running at full nameplate capacity over a period of time, so a wind capacity factor of 0.6 means that wind is generating at 60% of its nameplate capacity. LDCs are often used in power system analysis to visually summarize the availability of a generation technology over an entire year. The R-squared of the original, 1-hourly LDC compared to non-normalized data was 0.64 (standard deviation 0.084), and 0.91 compared to normalized data (standard deviation 0.019), suggesting that normalization delivers significantly more accurate clustering. To more fully judge the impact different approaches to reduce time resolution have on model results, additional metrics are used below.

For both clustering approaches, the observations on which clustering is performed are all time steps for a given day and across all model regions, for four variables: onshore wind, offshore wind and PV generation, and demand. That is, the Euclidean distance is calculated by reshaping the data from a 3-dimensional time steps \times locations \times variable matrix to a 2-dimensional days \times n matrix, where n contains time steps per day, locations, and variables. Furthermore, different numbers of clusters and two choices for determining typical days are compared: cluster centroids (means) or the real day closest to the cluster centroid, from all days contained in a cluster. The closest real day is found by the minimal Euclidean distance between each real day and the cluster centroid, computed with the Frobenius matrix norm. Both clustering methods use their implementation in the SciPy Python package,² and the implementation of all time resolution reduction approaches is available in the Calliope source code.³ To give an example, the method labeled “min/max solar and wind days - kmeans - 10 days - closest” would first pick the days with maximum and minimum PV and wind generation (up to 4 days), then remove them from consideration and apply k-means clustering to the remaining data, grouping it into 10 clusters, then selecting the closest real days to each cluster centroid, resulting in up to 10 additional days for a total selection of up to 14 days.

3.3. Model runs and analyses performed

Table 3 lists all the compared time resolution reduction approaches and the parameters used. In the case of the clustering methods, the derivation of actual days is one parameter (either

by selecting cluster centroids/ mean values, labeled “mean”, or selecting the day from within a given cluster that is closest to the cluster centroid, labeled “closest”), and the number of clusters to look for is the second parameter. The heuristic methods with “drop” in the selection column simply drop all non-selected time steps, i.e. in the case of “min/max wind weeks”, all but the weeks with maximum and minimum wind output are dropped.

First, to test the 42 approaches, each one is applied to a year (2014) of hourly data, therefore simulated renewable generation and reported demand data for that year are used. The resulting models with reduced time resolution are run for three different scenarios for each time resolution reduction approach. The three scenarios are defined by varying the constraint that controls the amount of variable renewable generation: (1) averaged over the entire modeled time period, a minimum of 50% of power must be supplied from PV and wind (called the “50% renewable generation” scenario), (2) a minimum of 90% (“90% renewable generation”), and (3) a minimum of 90% with the possibility of deploying a large amount of battery storage (“90% with storage”). Unless in this last case, the primary means of balancing variable renewables available to the model are deploying the generic dispatchable and baseload technologies. The deliberate lack of options to balance variable renewable output means that these scenarios are difficult to achieve and result in high levelized system costs, so they are not representative of real power system configurations, but are useful test cases to compare the effect of time resolution on models with both a 50% and 90% share of renewable power generation.

Second, a smaller set of time resolution reduction methods is applied to the full 25 years of data. Table 4 lists these eight approaches. The number of clusters (days) selected is higher here than in the methods applied to a single year only, in order to attempt to capture inter-annual variability. In addition, a different heuristic is used to select a single or two extreme days for each year of data, such that a representation of all annual extremes is available. “Extreme days” here means for each of wind generation, PV generation and power demand, the day per year with the maximum and the minimum mean value.

The primary goal when reducing time resolution is to improve computational tractability of the model, while retaining as much detail as possible. The metric for computational tractability is the number of time steps. The CPU time required to solve a model is also compared. As it is machine and solver dependent, it is normalized by the maximum CPU time over all runs. The metrics for examining the performance of a method are the percent deviation from the reference (1-hourly) model of system-wide LCOE and of installed technology capacities. In addition, the efficiency of an approach is determined by $1/(n_t * abs(C))$, where n_t is the number of time steps and C is the relative deviation of LCOE from the 1-hourly reference case. The efficiency is used to rank all time

² www.scipy.org.

³ www.github.com/calliope-project/calliope.

Table 4
Time resolution reduction methods used for the full 25 years of data.

Approach	Days	Selection
k-means	100	Closest
k-means	100	Mean
Hierarchical	100	Closest
Hierarchical	100	Mean
k-means	50	Closest
k-means – single extreme days	50	Closest
k-means – annual extreme days	50	Closest
k-means – 2 annual extreme days	10	Closest

resolution methods compared, and to choose the five most efficient ones for further inspection in the sections below. Other measures to assess the stability and suitability of clustering are possible, for example information-theoretic ones e.g.: [41]. However, here we want to look primarily at the optimization model outcomes, since those are the quantities of interest.

4. Results

4.1. Time resolution impact on model results

4.1.1. Downsampling

We first examine uniform downsampling of the input time series to 3-hourly, 6-hourly, 12-hourly and 24-hourly time steps, for the year 2014. Fig. 2 compares the effect of this downsampling on CPU time required to solve a given model, and the resulting optimal wind generation capacity. CPU time is normalized by the maximum reported CPU time since the absolute time is computer-dependent. The figure shows a significant reduction in normalized CPU time with each reduction of time resolution. The reduction in CPU time is most noticeable in the “90% with storage” scenario, since that model has more complex constraints with respect to time by including storage charging and discharging. Time resolution influences the system configuration found by the model considerably, as seen in the figure by the amount of installed offshore wind capacity. Since the optimization finds minimal cost configuration and is forced to match supply with demand, averaging renewable generation and demand time series over ever longer time periods results in lower installed capacities. This is because as demand peaks and production troughs are smoothed, it becomes easier for the model to fulfill the balancing constraint. The availability of storage in the system in the “90% with storage” scenario reduces the difference in installed wind capacity between time resolutions, since the smoothing of output and demand fluctuations now take place through storage. These results are a first indication that the appropriate method to reduce time resolution depends heavily on the model setup and constraints used.

One issue with downsampling of input data is the loss of intra-time step variability, which, as Fig. 2 shows, has a particularly strong effect on results from models with high shares of variable renewable generation but little storage. Fig. 3 compares model results from a 6-hourly and a 1-hourly run, for the first seven days in June 2014. Since the model used optimizes both plant deployment and dispatch simultaneously, the loss of detail through smoothing of peaks and troughs (for example on June 5th) can strongly affect model results. The loss of detail seen in this figure explains why downsampling the input data leads to a reduction of installed capacity in the “90% renewable generation” scenario. It removes the need for peaking capacity during particularly non-windy or non-sunny hours, since they are simply smoothed out in the time series. This effect is not as strong in the “50% renewable generation” case given that 50% of demand is covered by non-variable generation.

4.1.2. Clustering and heuristic selection

We now examine more complex methods to reduce time resolution and whether they can offset some of the undesirable side effects of downsampling. We first examine the effect these methods have on the structure of the demand and wind/PV generation time series for 2014 data, by using LDCs and the correlation between wind and PV. Table 5 shows the configurations we analyze more closely: the two methods with the highest efficiency (labeled 1 and 2, both combinations of individual methods), a k-means and a hierarchical clustering method for comparison (labeled 3, 4), and 6-hourly downsampling (labeled 5). Table S1 in the supplementary material contains the full list of all tested approaches. The efficiency measure used to select methods 1 and 2 is as defined in the methods section: the inverse of number of time steps multiplied by the relative deviation of model-reported LCOE from the 1-hourly reference case, in other words, a measure of how well a time resolution reduction method can approximate the 1-hourly reference case with as few time steps as possible. The selection of the most efficient methods will be discussed in more detail in the next section.

Fig. 4 shows load duration curves for offshore wind and PV for 1-hourly 2014 data, for the five methods listed in Table 5. The actual time series always contain less than 8760 time steps, but to draw the figure, each clustered time step gets replicated along the x-axis according to how many actual time steps it represents. Uniform downsampling of the time series to 6 h does not substantially affect the shape of the load duration curve for wind, whereas for PV, a 6-hourly resolution is insufficient to capture the diurnal solar cycle. The shape of the PV load duration curve is therefore skewed towards more hours with lower capacity factors.

Similarly, different time resolution reduction methods affect the shape of the load duration curve by emphasizing certain parts of it over others depending on the criteria used to choose the time steps. The method labeled number 2 in the figure (“min/max wind-demand weeks - drop”) picks the weeks with the minimum and maximum difference between wind power output and demand, but it does not replicate the LDC of PV well. The extreme ends of the load duration curve are of potential interest as they represent periods of very low or very high production. For the wind time series, both the k-means and hierarchical clustering approaches (labeled 3 and 4) flatten the LDCs substantially, that is, they remove much of the extremes and skew the LDC towards the average. In the case of PV, the LDCs from k-means and hierarchical clustering generally lead to an overprediction of PV capacity factors throughout the curve. The figure also shows that some of the methods considered here skew the correlation between wind and PV generation quite significantly compared to the original 1-hourly data (shown in black).

4.1.3. Model results and efficiency of different approaches

Above (Fig. 4), the method with the highest efficiency is a combination of selecting extreme days heuristically and clustering the remaining time periods (labeled 1 in the figure). Because this method represents LDCs and the PV-wind correlation worse than other methods, it is necessary to compare model results to judge the performance of different approaches.

Two model results are of particular interest: the deployed capacity of key technologies and the levelized cost of electricity (LCOE). Fig. 5a shows the installed capacity across the three different scenarios, for all 42 examined time reduction methods. As above, these results are from the 2014 model. Three key technologies are shown: offshore wind, PV, and the dispatchable fossil technology. We see that the range of installed capacities for offshore wind ranges from just above 40–170 GW. It should be noted that a small number of time resolution reduction approaches result in more wind capacity than the 1-hourly reference case (shown with

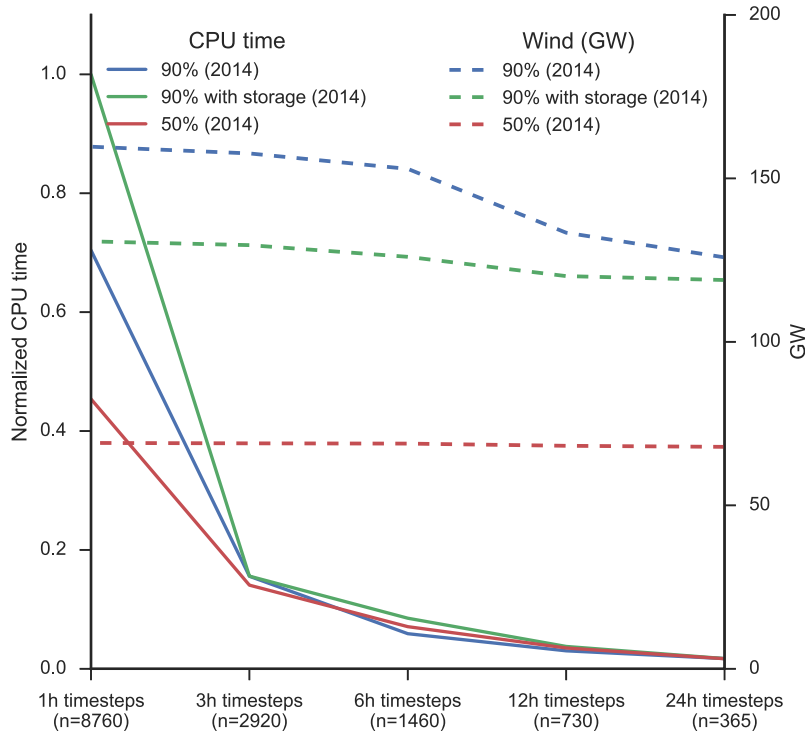


Fig. 2. Normalized CPU time and installed offshore wind capacity from 2014 runs with uniformly downsampled time resolution. On the machine used, the 1-hourly run (maximum run time) took about 7.2 h. All CPU times are normalized by dividing by that value.

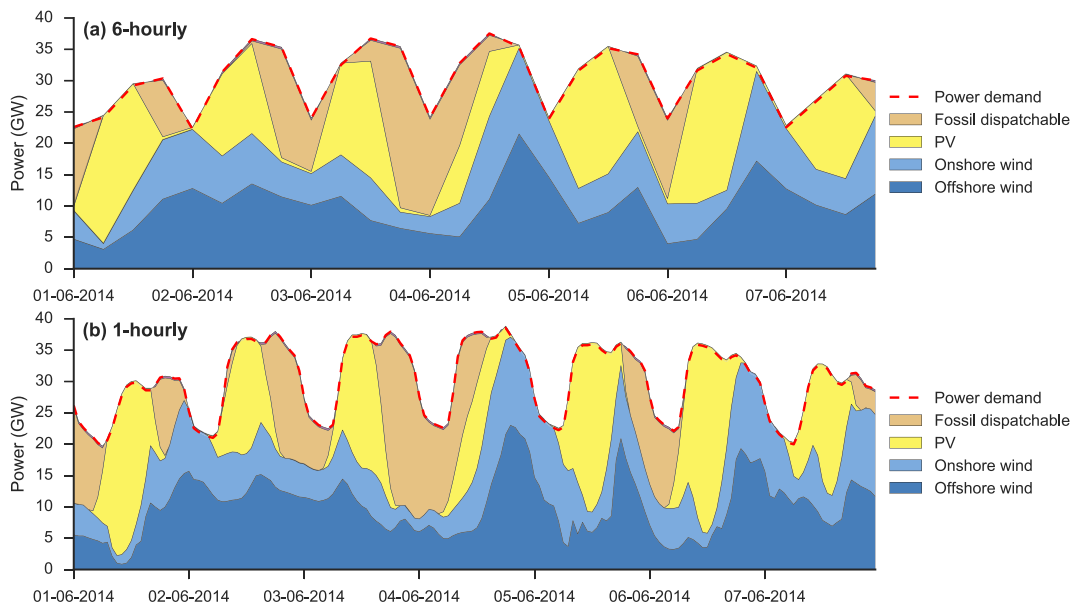


Fig. 3. Comparison of plant dispatch decisions from model runs with different time resolutions. (a) 6-hourly, (b) 1-hourly.

Table 5
 Compared time resolution reduction methods. Normalized CPU time is given in % of the CPU time for the full 1-hourly reference model.

Method	Time steps	Norm. CPU Time
1 Min/max solar and wind days – kmeans – 10 days – closest	336	1.97
2 Min/max wind-demand weeks – drop	168	1.15
3 k-means – 10 days – mean	240	1.64
4 Hierarchical – 10 days – closest	216	1.27
5 6 h timesteps	1460	8.40

the thick line). In the case of PV capacity, this effect is even more noticeable. These are methods that choose extreme days only and disregard the rest of the time series data. If we run a capacity deployment model for only days with extreme wind conditions, the model may well deploy substantially more wind generation than necessary. Thus, these methods are simply not useful in practice when modeling very high shares of variable renewable generation.

Fig. 5b shows the system-wide LCOE (generation-weighted average LCOE across technologies) computed by the 2014 model.

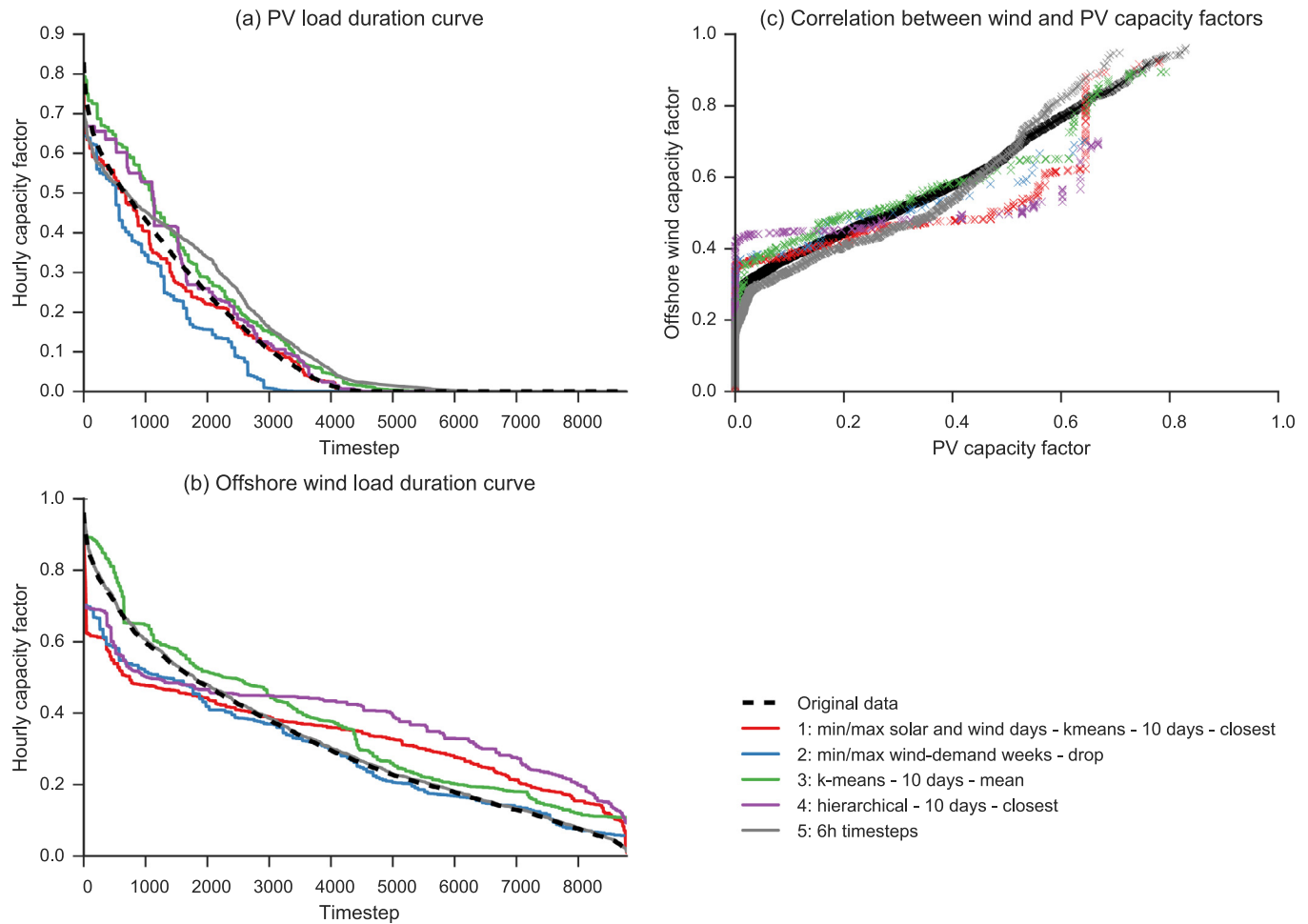


Fig. 4. Accuracy of load duration curves (LDC) after applying different time resolution reduction methods. (a) LDC for PV generation. (b) LDC for offshore wind generation. (c) Correlation between PV and wind capacity factors.

The LCOE range is significantly smaller for the “50% renewable generation” scenario. This is because in this scenario, the comparatively low share of variable renewable generation and conversely the higher share of dispatchable generation means that weather events, particularly extreme events for either wind or PV generation, have a lesser effect on the optimal model solution. The choice of time step resolution reduction method therefore plays a lesser role, and the range of results is narrower. Allowing for additional storage capacity in the “90% with storage” scenario also reduces the range of costs, by providing an alternate means for balancing variable generation and thus also reducing the importance of weather variability for the model solution. Overall, it appears that the method to reduce time resolution is of less importance unless very high shares (on the order of 90%) of variable renewables are modeled with little backup or balancing possibilities.

From amongst the complete 42 methods, we now select those five time resolution adjustment methods with the highest efficiency (as defined in the methods section) when applied to the 2014 model. Table 6 shows these methods. The two topmost have already been used in the above section when examining LDCs and correlation between wind and PV. A similar table with results for all examined methods and for the “50% renewable generation” and “90% with storage” scenarios can be found in the [supplementary material](#) in Tables S2–S4. While Fig. 5 showed that some of the 42 methods perform inadequately on the accuracy of both installed capacities and levelized costs, Table 6 shows that certain approaches can achieve values very close to the 1-hourly reference

case even with an order of magnitude fewer time steps. These results suggest a high time resolution is important for models that analyze systems with high shares of variable renewable generation. However while there is a trade-off between accuracy and the number of time steps in the model, the relationship is not linear. Significant model complexity gains can be made by careful selection of time steps while still retaining high accuracy in the results.

4.2. Inter-annual variability

4.2.1. Variability of wind and PV generation

Examining the results from just a single year leaves open the question of whether a method proven as most viable in that single year is equally suitable for other years. Fig. 6 shows the inter-annual variability from 25 years of PV and wind simulations for the UK, from the simulated data used for the model runs presented here. There are seasonal trends for both wind and PV generation, but also considerable inter-annual variability. A sunny day in winter can be almost as productive for PV output across the UK as a cloudy day in summer. The average UK wind capacity factor can range from close to zero to almost one on the same calendar day across the 25 years of simulations shown. This implies an inadequacy in using merely a single year of data. When examining multiple years of data and time series longer than a single year, two distinct problems emerge. The first is a modeling problem and requires the application of the time resolution reduction

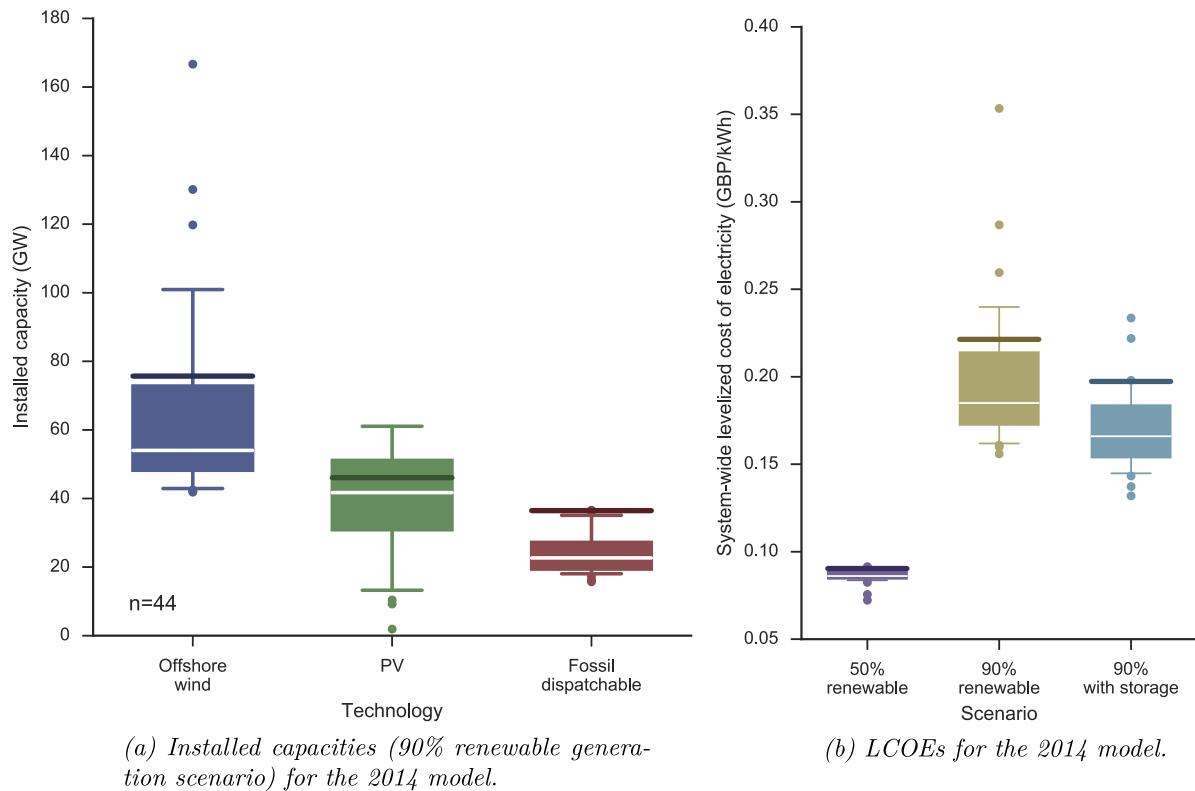


Fig. 5. Installed capacities and system-wide LCOEs for 2014 model runs. The thick line shows the reference model run (full 1-hourly data). The box spans the .25–.75 percentile and the white line in the box is the median.

Table 6

Relative difference in percentage points to the reference case (1-h, or full dispatch model), for system-wide LCOE and for offshore wind capacity deployment, for the year 2014 model and the “90% renewable generation” scenario. Results in 1-hourly reference case are: LCOE = 0.18 GBP/kW h, offshore wind capacity = 41.8 GW.

	LCOE (% diff.)	Wind cap. (% diff.)	Time steps	Efficiency
Min/max solar and wind days – kmeans – 10 days – closest	–0.2	–10.6	336	1.32
Min/max wind-demand weeks – drop	0.6	19.2	168	0.98
Min/max solar and wind days – kmeans – 5 days – mean	2.3	–1.0	264	0.16
Min/max solar and wind days – hierarchical – 5 days – mean	3.5	–0.1	264	0.11
Min/max solar and wind days – drop	–8.1	–14.6	144	0.09
6 h timesteps	–3.4	–4.4	1460	0.02

approaches described above to more than a single year of data. The second is a planning and policy problem. Despite the inter-year variability in the renewable resource, power systems with high shares of variable renewables will likely become a reality, so energy modelers must propose solutions for their stable operation.

4.2.2. Variability impact on time resolution reduction methods

To examine the modeling problem, Fig. 7 shows the percent deviation from the one-hourly reference model runs for the system-wide LCOE, for three different time resolution reduction methods, over five years of model runs. A negative deviation means that the method is underestimating the LCOE relative to the 1-hourly model runs. The three methods were chosen to illustrate the wide range of behavior as follows. Method A (min/max solar and wind days - downsample 6 h) shows the lowest mean absolute LCOE error across the five years, method B (min/max wind-demand weeks - drop) has the highest mean efficiency across years, and method C (k-means - 5 days - closest) is chosen as an illustrative example of an approach with clustering only. This last approach does well in some years, but leads to substantial errors in other years. Thus, by validating a clustering method based on just a single year of data, one may erroneously conclude that it is more

robust than it really is. Method A (min/max solar and wind days - downsample 6 h) performs relatively consistently across all five years, most likely because it heuristically selects the extreme events that drive model results. However, this method is quite far from being efficient: it requires 1580 time steps. Method B, with only 168 time steps, is substantially more efficient computationally, but does not show the same consistent behavior over the five years compared.

If we repeat the analysis shown in Table 6 above for the year 2012, as shown in Table 7, the resulting list of approaches is different. However, a consistent pattern is that methods with heuristic selection of the maximum and minimum solar and wind days (“min/max solar and wind days”) are amongst those that work best.

4.2.3. Model planning results over multiple decades

The variability of input data across multiple years translates into significant differences in model results. Fig. 8 shows the range of installed generation capacities when running single-year models for each of the 25 years, for the “90% renewable generation” and the “90% with storage” scenarios. The considerable range of installed capacities casts doubt on model results using only a single

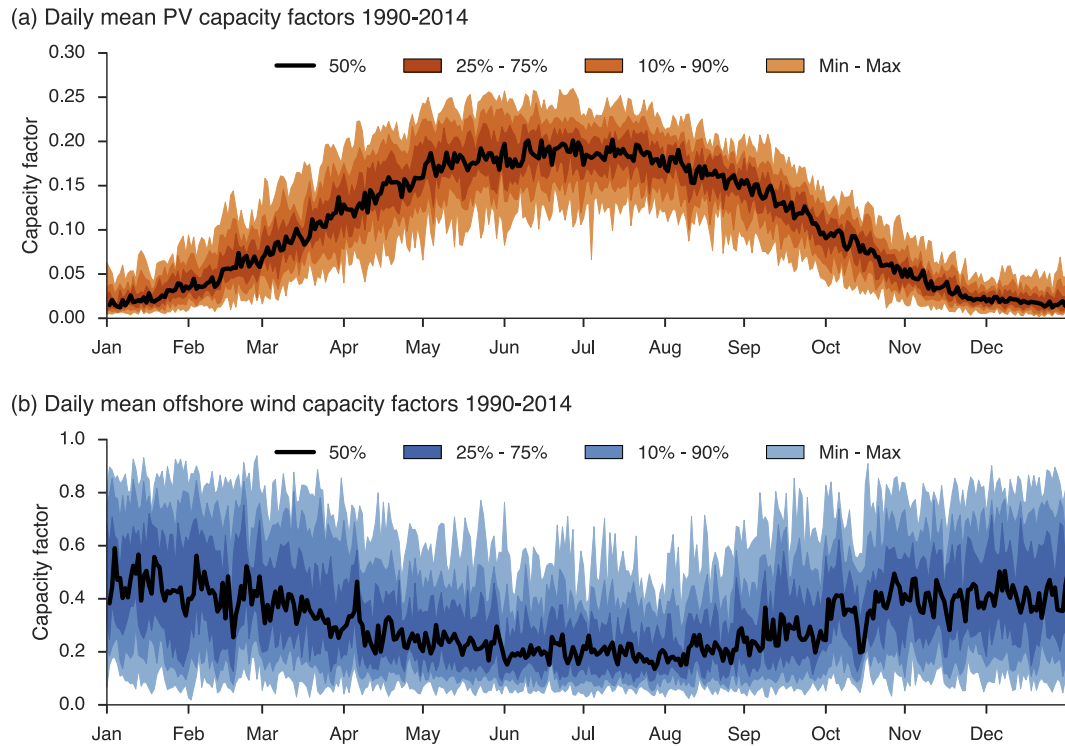


Fig. 6. Inter-annual variability of daily capacity factors from 1990 to 2014. (a) PV. (b) Offshore wind. The black line shows median daily capacity factors for each day of year from the 25 years of hourly data used in this study, while the shaded areas indicate their possible range over those 25 years.

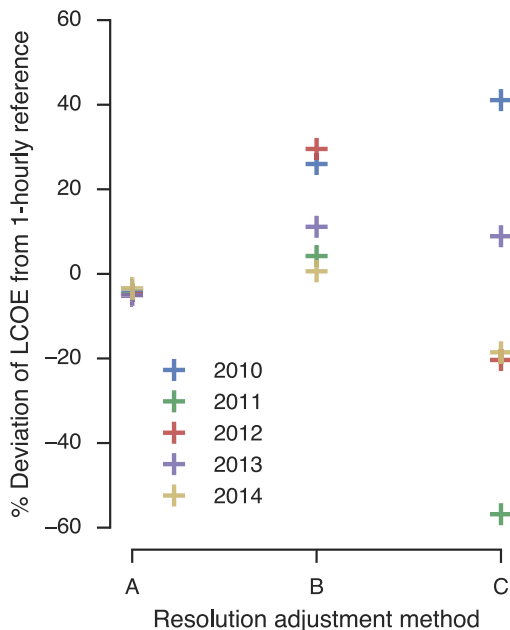


Fig. 7. LCOE variability across five years when applying three time resolution adjustment methods. A: min/max solar and wind days – downsample 6 h, B: min/max wind-demand weeks – drop, C: k-means – 5 days – closest.

or a small number of years of data. While allowing additional storage (Fig. 8b) reduces the absolute capacity as well as the range somewhat by providing an option to balance weather variability, the model results still show a considerable range of capacity.

Fig. 9 shows installed offshore wind and PV capacities from each of the 25 years separately, for the “90% renewable generation” scenario, sorted by wind capacity. Wind capacity is relatively equally distributed throughout the range spanned by the 25 years.

The year 1991 could be considered an outlier, and discarded as an exceptional extreme condition, or it could be considered a condition that energy planning needs to address. Even when discarding this outlier, the remaining years span a 30 GW range of difference in installed capacity however. By breaking out installed offshore wind capacity into the different model zones, across the 25 years, as seen in Fig. 10, we see that this variability stems from only a few model locations. For robust capacity planning, therefore, these locations in particular would be of interest.

Finally, we examine the results from applying different time resolution reduction methods to the full 25 year time series (see Table 4 for the methods used for this). Fig. 11 shows the results in a similar manner as the boxplots of deployed capacity above, but additionally showing as dots the individual results contained within the boxplots. The range of installed generation capacities is even larger than the range from running individual years, which would be expected given that some of the time resolution reduction methods specifically select (and thus focus on) extreme events from the data. For the purposes of robust capacity planning in the face of inter-annual variability, this could imply that certain approaches to selecting time steps are suitable to determine a “mean” system design across all years. However, there is also the question of weighting the rarer extreme events. By having too high a weight, they will skew model results towards overcapacity. Which of the resulting capacities is “true” is not a straightforward determination, and these results also do not give a clear answer to the question of how an energy model can determine a robust planning approach in the face of inter-annual variability whilst also working with a reduced set of time steps drawn from a longer time series.

5. Discussion and conclusion

This paper has compared different methods to reduce the time resolution of energy model input time series, and assessed their

Table 7

Relative difference in percentage points to the reference case (1-h, or full dispatch model), for system-wide LCOE and for offshore wind capacity deployment, for the year 2012 model and the “90% renewable generation” scenario. Results in 1-hourly reference case are: LCOE = 0.19 GBP/kW h, offshore wind capacity = 55.9 GW.

	LCOE (% diff.)	Wind cap. (% diff.)	Timesteps	Efficiency
Min/max solar and wind days – hierarchical – 5 days – closest	0.1	–4.2	264	3.53
Min/max solar and wind days – kmeans – 5 days – closest	1.3	–1.8	264	0.29
k-means – 10 days – closest	–3.5	–0.4	144	0.20
Hierarchical – 15 days – closest	4.8	6.4	240	0.09
Min/max solar and wind days - hierarchical – 5 days – mean	–5.4	–8.3	264	0.07
6 h timesteps	–5.0	–6.2	1464	0.01

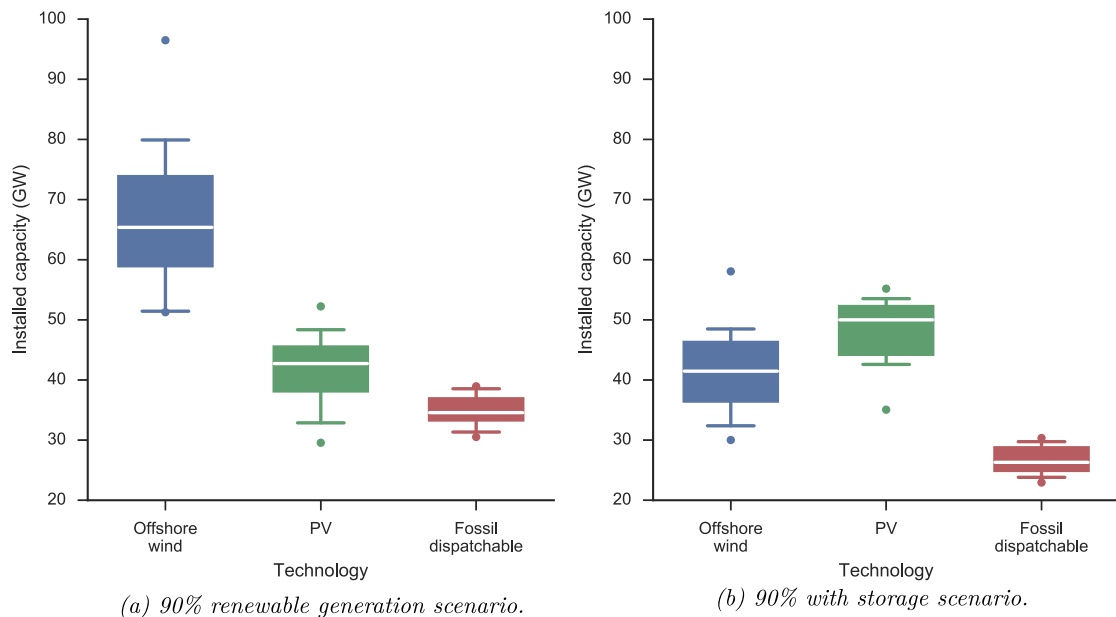


Fig. 8. Variability in model results for installed capacity from the 25 yearly model runs. This variability in results is driven by wind and PV variability in the input data. The box spans the .25–.75 percentile and the white line in the box is the median.

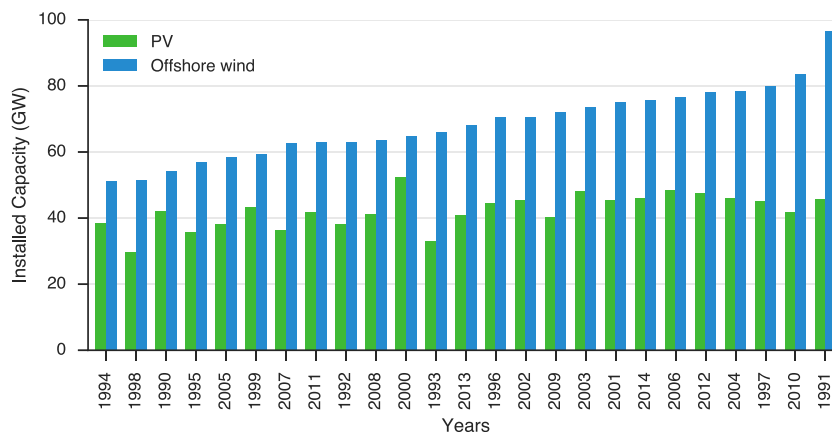


Fig. 9. Offshore wind and PV capacity model results for each of the 25 annual runs, sorted by offshore wind capacity.

effect on the outputs from a power system planning model for Great Britain. The results show that different methods, including downsampling, heuristic selection of time steps, and clustering, lead to substantially different model results, in particular when modeling high shares of variable renewable generation. Approaches including heuristic methods appear to be more stable when applied across different models of individual years, and may therefore be preferable for those with high shares of variable renewables. In the “50% renewable generation” scenario (with a

constraint that at least 50% of total generation over the entire modeled period must come from wind and PV), model results still showed acceptable accuracy at resolutions of 6-hourly and lower, when comparing to the reference 1-hourly model.

The “90% renewable generation” scenario substantially increased the difference in results between different amounts of time steps and time resolution reduction methods. This underscores that modeling renewables adequately requires high resolution input data, but also that documenting the processing steps

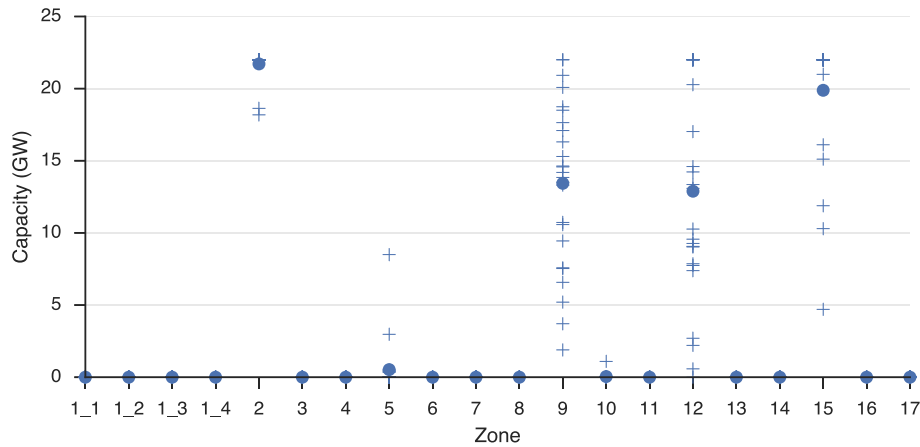


Fig. 10. Offshore wind capacity by zone, for each of the 25 annual model runs. Crosses are yearly model decisions, dots are the mean across all 25 years.

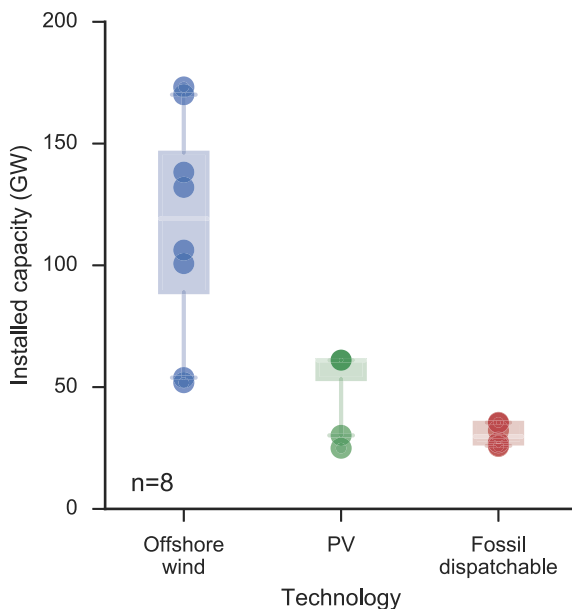


Fig. 11. Variability in model results for installed capacity over the different time resolution adjustment methods applied to the full 25 years of time series data. Dots are the individual model results aggregated into the boxplots.

applied to this input data in the course of any given analysis is important to understand and evaluate model results. Results pertaining to installed capacity requirements, levelized costs and the overall feasibility of certain power or energy system configurations, should be used with caution when high shares of renewables are modeled with time resolutions much below 1-hourly or with only a small number of typical days. Giving a model additional inter-time step balancing opportunities like storage further decreases the optimization's computational tractability, but it also reduces the importance of high time resolution. Overall, the results suggest that the appropriate method with which to reduce the number of time steps in a model depends on the model setup. Rather than falling back on the same method for all model runs, studies should therefore include justification for why a specific approach was chosen and validate that method for their specific model configuration.

The situation is further complicated by the substantial inter-annual variability in wind and PV power output. This variability results in the need to use not just a single year (for example, a typical meteorological year) but as long a time series as possible,

ideally several decades. The resulting modeling problem is how to deal with the increased amount of data while maintaining computational tractability. It is also a planning problem, since it involves the question of how much variable generation capacity to build, and where, to deal with this variability. Furthermore, it entails setting criteria to apply for acceptable levels of supply security and system stability. These security and stability requirements are influenced by the variability of renewable generation and the availability and cost of backup and balancing options for variable renewables. Assessing them therefore requires that model results not be distorted by inaccurate time series data. For all these reasons, there is the need for further research into how to integrate multi-decade time series of renewable generation into energy planning models, so that they in turn can help plan for secure and affordable energy systems with high shares of variable generation.

The work presented here has several limitations. The optimization model used is a simplified one and does not therefore lead directly to real-world planning insights. The resulting levelized costs, for example, for the “90% renewable generation” and “90% with storage” scenarios, should not be interpreted as representative of the costs of real power systems, but are used to compare the relative differences between a model run with different time step configurations. Furthermore, the relevance of spatial detail is not further examined in this work. Temporal resolution is important for the balancing of supply and demand and for depicting the variability of wind and PV. Spatial resolution allows a model to capture dynamics such as a weather system moving across a country and the impact this has on variable renewable generation in different model locations. Depending on a model's purpose – for example, examining decentralized generation – higher spatial resolution may be of crucial importance. The analyses performed here cannot quantify the effect of spatial resolution or its relative importance compared to temporal resolution, so this is another avenue for future work.

This study has assessed methods to address time resolution systematically, extending knowledge on what methods are appropriate and where, and underscoring the fact that published results based on models should always carefully specify what adjustments were made to the time series data. The methods and data are freely available to build upon. This makes it possible for other modelers to easily adapt or extend the approaches presented here, and apply them to their own models and data. The results suggest that while heuristic approaches appear promising, there is no one-size-fits-all approach to reduce time resolution while also covering long-term variability. However given the rising importance of variable renewable generation, there is both a need and ample room for more research on these problems.

Acknowledgements

The author thanks the participants of the WholeSEM workshop on integrated energy system models incorporating spatial and temporal detail (London, May 2016) for useful comments. Funding for this work was provided by the European Research Council via grant StG 2012–313553.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.apenergy.2017.03.051>.

References

- [1] IRENA. Installed renewable energy power capacity; 2016. <https://public.tableau.com/shared/TZPQPDW32?toolbar=no&:display_count=no&:showVizHome=no> [Accessed: 2016-07-06].
- [2] Krey V. Global energy-climate scenarios and models: a review. *Global energy-climate scenarios and models*. Wiley Interdiscip Rev: Energy Environ 2014;3(4):363–83. <http://dx.doi.org/10.1002/wene.98>.
- [3] Haller M, Ludig S, Bauer N. Decarbonization scenarios for the EU and MENA power system: considering spatial distribution and short term dynamics of renewable generation. *Energy Policy* 2012;47:282–90. <http://dx.doi.org/10.1016/j.enpol.2012.04.069>.
- [4] Loulou R, Labriet M. ETSAP-TIAM: the TIMES integrated assessment model Part I: Model structure. *Comput Manage Sci* 2008;5(1):7–40.
- [5] Haydt G, Leal V, Pina A, Silva CA. The relevance of the energy resource dynamics in the mid/long-term energy planning models. *Renew Energy* 2011;36(11):3068–74. <http://dx.doi.org/10.1016/j.renene.2011.03.028>.
- [6] Widen J. Correlations between large-scale solar and wind power in a future scenario for Sweden. *IEEE Trans Sustain Energy* 2011;2(2):177–84. <http://dx.doi.org/10.1109/TSTE.2010.2101620>.
- [7] Schroeder A, Oei P-Y, Sander A, Hankel L, Laurisch LC. The integration of renewable energies into the German transmission grid—A scenario comparison. *Energy Policy* 2013;61:140–50. <http://dx.doi.org/10.1016/j.enpol.2013.06.006>.
- [8] Pfenninger S, Keirstead J. Renewables, nuclear, or fossil fuels? Scenarios for Great Britain's power system considering costs, emissions and import dependency. *Appl Energy* 2015;152:83–93. <http://dx.doi.org/10.1016/j.apenergy.2015.04.102>.
- [9] Rodriguez RA, Becker S, Greiner M. Cost-optimal design of a simplified, highly renewable pan-European electricity system. *Energy* 2015;83:658–68. <http://dx.doi.org/10.1016/j.energy.2015.02.066>.
- [10] Pfenninger S, Staffell I. Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data. *Energy* 2016;114:1251–65. <http://dx.doi.org/10.1016/j.energy.2016.08.060>.
- [11] Staffell I, Pfenninger S. Using bias-corrected reanalysis to simulate current and future wind power output. *Energy* 2016;114:1224–39. <http://dx.doi.org/10.1016/j.energy.2016.08.068>.
- [12] 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 Manage, Early Access Online* 2014. <http://dx.doi.org/10.1109/TEM.2013.2284386>.
- [13] Poncelet K, Hoschle 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;PP(99). <http://dx.doi.org/10.1109/TPWRS.2016.2596803>. p. 1–1.
- [14] 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;112:430–42. <http://dx.doi.org/10.1016/j.energy.2016.06.081>.
- [15] Heuberger C, Staffell I, Shah N, Mac Dowell N. A systems approach to quantifying the value of power generation and storage technologies in future electrical networks. *Comput Chem Eng* 2016. in preparation.
- [16] Hennig C. Cluster-wise assessment of cluster stability. *Comput Stat Data Anal* 2007;52(1):258–71. <http://dx.doi.org/10.1016/j.csda.2006.11.025>.
- [17] Heaps C. Introduction to LEAP. Technical report, SEI; 2008. <<https://www.energycommunity.org/documents/LEAPIntro.pdf>> [Accessed: 2016-09-27].
- [18] Budischak C, Sewell D, Thomson H, Mach L, Veron DE, Kempton W. Cost-minimized combinations of wind power, solar power and electrochemical storage, powering the grid up to 99.9% of the time. *J Power Sources* 2013;225:60–74. <http://dx.doi.org/10.1016/j.jpowsour.2012.09.054>.
- [19] Pina A, Silva C, Ferrão P. Modeling hourly electricity dynamics for policy making in long-term scenarios. *Energy Policy* 2011;39(9):4692–702. <http://dx.doi.org/10.1016/j.enpol.2011.06.062>.
- [20] Kannan R, Turton H. A long-term electricity dispatch model with the TIMES framework. *Environ Model Assess* 2012;18(3):325–43. <http://dx.doi.org/10.1007/s10666-012-9346-y>.
- [21] Poncelet K, Delarue E, Duerinckx J, Six D, D'haeseleer W. Impact of temporal and operational detail in energy-system planning models. TME working paper – energy and environment, ?KU Leuven Energy Institute; 2015. <https://www.mech.kuleuven.be/en/tme/research/energy_environment/Pdf/wp-en201420-2.pdf> [Accessed: 2015-04-23].
- [22] Rhodes JD, Cole WJ, Upshaw CR, Edgar TF, Webber ME. Clustering analysis of residential electricity demand profiles. *Appl Energy* 2014;135:461–71. <http://dx.doi.org/10.1016/j.apenergy.2014.08.111>.
- [23] Räsänen T, Voukantsis D, Niska H, Karatzas K, Kolehmainen M. Data-based method for creating electricity use load profiles using large amount of customer-specific hourly measured electricity use data. *Appl Energy* 2010;87(11):3538–45. <http://dx.doi.org/10.1016/j.apenergy.2010.05.015>.
- [24] McLoughlin F, Duffy A, Conlon M. A clustering approach to domestic electricity load profile characterisation using smart metering data. *Appl Energy* 2015;141:190–9. <http://dx.doi.org/10.1016/j.apenergy.2014.12.039>.
- [25] Baringo L, Conejo AJ. Correlated wind-power production and electric load scenarios for investment decisions. *Appl Energy* 2013;101:475–82. <http://dx.doi.org/10.1016/j.apenergy.2012.06.002>.
- [26] Mena R, Hennebel M, Li Y-F, Zio E. Self-adaptable hierarchical clustering analysis and differential evolution for optimal integration of renewable distributed generation. *Appl Energy* 2014;133:388–402. <http://dx.doi.org/10.1016/j.apenergy.2014.07.086>.
- [27] Samsatli S, Samsatli NJ. A general spatio-temporal model of energy systems with a detailed account of transport and storage. *Comput Chem Eng* 2015;80:155–76. <http://dx.doi.org/10.1016/j.compchemeng.2015.05.019>.
- [28] Hsu D. Comparison of integrated clustering methods for accurate and stable prediction of building energy consumption data. *Appl Energy* 2015;160:153–63. <http://dx.doi.org/10.1016/j.apenergy.2015.08.126>.
- [29] Parpas P. Integrated multiscale models for the optimal integration of renewable and distributed electricity generation. Technical report; 2010. <<http://www.eprg.group.cam.ac.uk/wp-content/uploads/2010/03/Panos-Parpas.pdf>> [Accessed: 2013-02-13].
- [30] Ludig S, Haller M, Schmid E, Bauer N. Fluctuating renewables in a long-term climate change mitigation strategy. *Energy* 2011;36(11):6674–85. <http://dx.doi.org/10.1016/j.energy.2011.08.021>.
- [31] Heussen K, Koch S, Ulbig A, Andersson G. Energy storage in power system operation: the power nodes modeling framework. In: Innovative smart grid technologies conference Europe (ISGT Europe), 2010 IEEE PES. p. 1–8. <http://dx.doi.org/10.1109/ISGT-EUROPE.2010.5638865>.
- [32] Miranda M, Dunn R. Spatially correlated wind speed modelling for generation adequacy studies in the UK. In: IEEE power engineering society general meeting, 2007. p. 1–6. <http://dx.doi.org/10.1109/PES.2007.386268>.
- [33] National Grid. 2013 Electricity ten year statement. Technical report, National Grid plc, Warwick, UK; 2013. <<http://www2.nationalgrid.com/UK/Industry-information/Future-of-Energy/Electricity-ten-year-statement/Current-statement/>> [Accessed: 2014-06-06].
- [34] Rienecker MM, Suarez MJ, Gelaro R, Todling R, Bacmeister J, Liu E, et al. MERRA: NASA's modern-era retrospective analysis for research and applications. *J Climate* 2011;24(14):3624–48. <http://dx.doi.org/10.1175/JCLI-D-11-00015.1>.
- [35] Staffell I, Green R. How does wind farm performance decline with age? *Renew Energy* 2014;66:775–86. <http://dx.doi.org/10.1016/j.renene.2013.10.041>.
- [36] Ridley B, Boland J, Lauret P. Modelling of diffuse solar fraction with multiple predictors. *Renew Energy* 2010;35(2):478–83. <http://dx.doi.org/10.1016/j.renene.2009.07.018>.
- [37] Huld T, Gottschalg R, Beyer HG, Topič M. Mapping the performance of PV modules, effects of module type and data averaging. *Sol Energy* 2010;84(2):324–38. <http://dx.doi.org/10.1016/j.solener.2009.12.002>.
- [38] National Grid. National grid data explorer > historical demand data; 2015. <<http://www2.nationalgrid.com/UK/Industry-information/Electricity-transmission-operational-data/Data-Explorer/>> [Accessed: 2015-09-21].
- [39] Hubert L, Arabie P. Comparing partitions. *J Classif* 1985;2(1):193–218. <http://dx.doi.org/10.1007/BF01908075>.
- [40] Vinh NX, Epps J, Bailey J. Information theoretic measures for clusterings comparison: variants, properties, normalization and correction for chance. *J Machine Learn Res* 2010;11(Oct):2837–54.
- [41] Pascual D, Pla F, Sánchez JS. Cluster validation using information stability measures. *Pattern Recogn Lett* 2010;31(6):454–61. <http://dx.doi.org/10.1016/j.patrec.2009.07.009>.