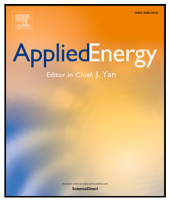
Reducing climate risk in energy system planning: a posteriori time series aggregation for models with storage

Applied Energy 334 (2023) 120624

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy



Reducing climate risk in energy system planning: A posteriori time series aggregation for models with storage

Adriaan P. Hilbers [a,*], David J. Brayshaw [b], Axel Gandy [a]

[a] Department of Mathematics, Imperial College London, United Kingdom
[b] Department of Meteorology, University of Reading, United Kingdom

ARTICLE INFO

Keywords:
Energy system modelling
Energy system optimisation model
Capacity expansion planning
Time series aggregation
Storage
Climate

ABSTRACT

The growth in variable renewables such as solar and wind is increasing the impact of climate uncertainty in energy system planning. Addressing this ideally requires high-resolution time series spanning at least a few decades. However, solving capacity expansion planning models across such datasets often requires too much computing time or memory.

To reduce computational cost, users often employ *time series aggregation* to compress demand and weather time series into a smaller number of time steps. Methods are usually *a priori*, employing information about the input time series only. Recent studies highlight the limitations of this approach, since reducing statistical error metrics on input time series does not in general lead to more accurate model outputs. Furthermore, many aggregation schemes are unsuitable for models with storage since they distort chronology.

In this paper, we introduce *a posteriori* time series aggregation schemes that preserve chronology and hence allow modelling of storage technologies. Our methods adapt to the underlying energy system model; aggregation may differ in systems with different technologies or topologies even with the same time series inputs. They do this by using operational variables (generation, transmission and storage patterns) in addition to time series inputs when aggregating.

We investigate a number of approaches. We find that *a posteriori* methods can perform better than *a priori* ones, primarily through a systematic identification and preservation of relevant extreme events. We hope that these tools render long demand and weather time series more manageable in capacity expansion planning studies. We make our models, data, and code publicly available.

1. Introduction

1.1. Capacity expansion planning models

The growth of variable renewables such as solar and wind has created new computational challenges in optimisation-based energy system planning. This is because accurate representation of such technologies' variability requires both a high spatiotemporal resolution [1–3] and long simulation lengths [4–12]. This leads to long solution times, since algorithms to solve the associated optimisation problems scale quickly (often exponentially) in the number of time steps [13,14]. It also hampers the study of climate impacts, since the use of climate model data – typically various multi-year samples from an ensemble of simulations – requires too much computing time or memory [15,16].

In this paper, we consider *capacity expansion planning* models, used to inform investments into energy infrastructure [17]. They determine the system design[1] that minimises the sum of install and subsequent operation costs given a sample of demand and weather data [18]. We view them as functions Φ_{plan} from a demand and weather time series $(\xi_t)_{t\in\mathcal{T}}$ to the associated optimal system design (installed capacities of generation, transmission and storage technologies) $\mathbf{D}_{\mathcal{T}}$:

$$\mathbf{D}_{\mathcal{T}} = \Phi_{\text{plan}}((\xi_t)_{t\in\mathcal{T}}). \tag{1}$$

The vector ξ_t contains time series values in period t. For example, for daily periods with hourly demand levels and wind speeds,

$$\xi_t = [d_{t,1}, \cdots, d_{t,24}, w_{t,1}, \cdots, w_{t,24}] \tag{2}$$

where $d_{t,i}$ and $w_{t,i}$ are the demand and wind speed respectively in the ith hour of day t.

Each planning problem has an associated *operation* problem (sometimes called the *production cost model*), in which we fix the system

* Corresponding author.
E-mail address: a.hilbers17@imperial.ac.uk (A.P. Hilbers).

[1] In this paper, we consider *build-from-scratch* models that determine the full system design. Ideas generalise naturally to investments into existing systems.

https://doi.org/10.1016/j.apenergy.2022.120624
Received 15 October 2022; Received in revised form 13 December 2022; Accepted 29 December 2022
Available online 13 January 2023

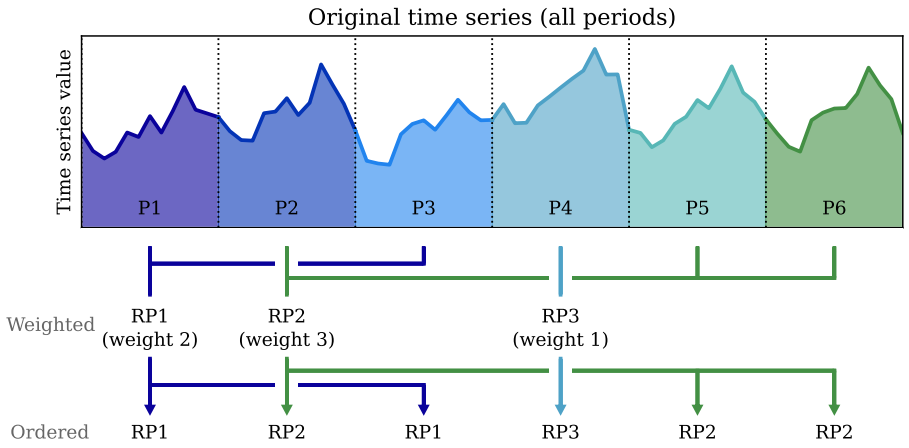


Fig. 1. Time series aggregation from six periods (P1-6) to three *representative periods* (RP1-3), either *weighted* (appearing once, weighted by number of occurrences in full time series) or *ordered* (in same order as full time series).

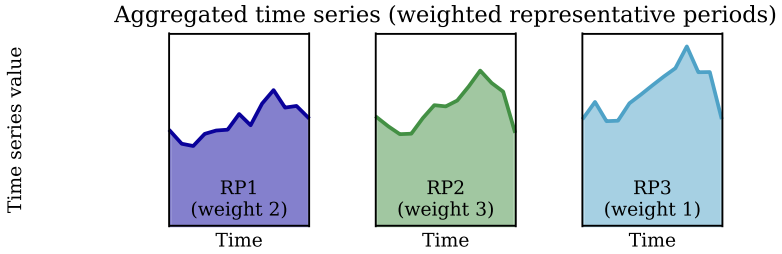


Fig. 2. Time series after aggregation into weighted representative days.

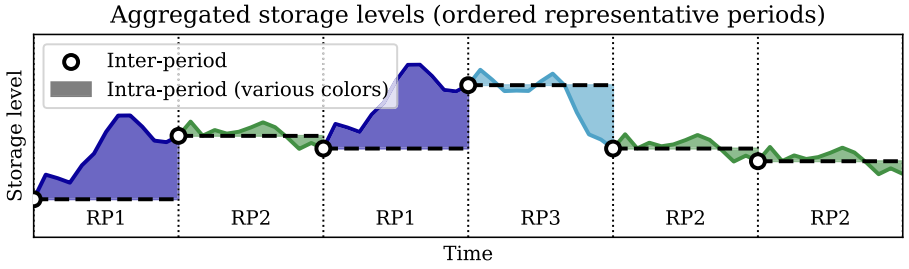


Fig. 3. Storage levels after aggregation into ordered representative days and decomposition into *inter-period* (level at start of period) and *intra-period* (change compared to start of period) levels. Intra-period contributions are equal in each replication of the same representative period.

design $\mathbf{D}$ and optimise the system operation:

$$(\mathbf{O}_t)_{t\in\mathcal{T}} = \Phi_{\text{operate}}((\xi_t)_{t\in\mathcal{T}} \mid \mathbf{D}) \tag{3}$$

where $\mathbf{O}_t$ contains the *operational variables* in period t: generation, transmission and storage (dis)charge decisions.

1.2. Time series aggregation

Time series aggregation, as reviewed by Hoffmann et al. [19] and Teichgraeber and Brandt [20], creates compressed time series that planning models are subsequently solved across. Many approaches create a smaller set of *representative periods* as shown in Fig. 1. An example mapping $t \mapsto \mathcal{A}(t)$, from six periods to three, is:

$$\begin{aligned}(\xi_t)_{t\in\mathcal{T}} &= (\xi_1 \quad \xi_2 \quad \xi_3 \quad \xi_4 \quad \xi_5 \quad \xi_6)\\ &\qquad\qquad\downarrow\\ (\hat{\xi}_{\mathcal{A}(t)})_{t\in\mathcal{T}} &= (\hat{\xi}_1 \quad \hat{\xi}_2 \quad \hat{\xi}_1 \quad \hat{\xi}_3 \quad \hat{\xi}_2 \quad \hat{\xi}_2).\end{aligned} \tag{4}$$

The number of unique representative periods $\{\mathcal{A}(t) \mid t \in \mathcal{T}\}$ is usually significantly smaller than the number of original periods $\{t \mid t \in \mathcal{T}\}$. The mapping $\mathcal{A}$ can be determined in various ways, such as choosing days from each season [21], minimising the deviation of load duration curves [22,23] or clustering vectors of each period's time series values [23–28].

Without constraints linking periods, time series aggregation reduces computational cost since operational decision variables appear once per *representative* period, weighted in the objective function by its number of occurrences [29]. These are *weighted* representative periods in Figs. 1 and 2. We discuss linked periods in Section 1.3.

Relevant extreme events disproportionately drive accurate estimates of optimal design or cost [30]. By *extreme* events, we mean those which require maximum generation, transmission or storage capacities to meet demand. In older energy systems, these are mostly peak demand events. However, with renewable technologies or storage, they may be those with peak demand minus available renewable (e.g. solar or wind) supply, or the end of a prolonged period of low renewable output, when storage levels are depleted. Exactly what these events are depends on the underlying energy system; identifying them is one of the key innovations of *a posteriori* methods (Section 1.4).

Heuristics such as including the maximum demand or minimum renewable potential day [26,27] are sometimes employed to identify extreme events. However, such approaches may fail to identify the

extremes relevant to the particular model. For example, peak demand may not require peak capacity if there is ample renewable generation or if stored or imported energy is available.

1.3. Inter-period links and storage

Constraints linking periods, such as storage, complicate time series aggregation since they require chronology of representative periods to be preserved. A number of solutions have been proposed; they include merging only periods that are adjacent chronologically [31–33], aggregating periods from different parts of the year separately [21,34,35], and *linking* storage levels between representative periods [36–40].

Another complication is that aggregating time series inputs no longer automatically reduces the number of decision variables, since absolute storage levels may differ in replications of the same representative period (Fig. 3). For this reason, modellers often assume storage (dis)charge decisions are identical in each replication [36,37,39]. Kotzur et al. [38] exploit this assumption by decomposing storage levels into *inter-period* (level at start of period, one per original period) and *intra-period* (deviation from start of period, one per representative period) contributions. Fig. 3 shows this decomposition and how *ordered* representative days with linked storage levels preserve chronology. For a detailed discussion, see [41].

1.4. A posteriori methods

Most time series aggregation schemes are what [19] call *a priori*: they use information about the input time series only, creating identical aggregation for any model with the same time series inputs irrespective of technologies or topology. Such an approach would generate, for example, the same representative days for a model with various renewable generation, transmission and storage technologies as for a model with primarily fossil fuels. This has been criticised by Wogrin [42], since reducing error metrics on time series inputs alone does not necessarily improve estimates of model outputs, i.e. optimal system design or cost [20,41].

A posteriori (also known as *adaptive*) methods use information about the underlying energy system model to tailor aggregation. For example, Sun et al. [43] and Zhang et al. [44] cluster vectors of planning model outputs (run on each individual day) instead of the time series itself. Bahl et al. [45] and Teichgraeber et al. [46] alternate between a planning model on aggregated data and an operation model on the full time series to iteratively identify and include days with unmet demand; this ensures design estimates have adequate generation capacity for such events. Hilbers et al. [47] identify and include system-relevant extreme events using their generation cost, also calculated using an operation model. Li et al. [48] combines elements of both such approaches.

2. This paper's contribution

In this paper, we introduce *a posteriori* time series aggregation schemes for capacity expansion planning models with storage. These schemes (1) tailor aggregation to the underlying energy system model and (2) preserve chronology, allowing the representation of long-term storage patterns. To our best knowledge, we are the first to combine these approaches. We make our models, time series data and code available at doi:10.5281/zenodo.7178301.

We introduce a framework that uses a model's operational variables – generation, transmission and storage patterns – to improve aggregation. We generalise work by Hilbers et al. [47] and Teichgraeber et al. [46] in a number of ways. Firstly, we allow the modelling of storage technologies by preserving chronology. This requires significant modifications to the problem's constraints – in particular the linking of storage levels between representative periods in order, and their decomposition into inter- and intra-period levels as in the method introduced by Kotzur et al. [38]. We also allow the use of (and explain how to use) arbitrary functions of operational variables (e.g. energy price), instead of only the generation cost or unserved energy. This makes our approach applicable to more general models, such as those without a notion of unserved energy (when energy demand is elastic). Finally, we use operational variables (in particular, storage (dis)charge patterns) when clustering, a previously unseen innovation that enhances optimal storage capacity estimation. The methods of Hilbers et al. [47] and Teichgraeber et al. [46] are specific applications of our method in a simpler setting that (1) conduct stratified sampling using generation cost or unmet demand respectively (instead of a more general function), (2) do not use any operational variables when clustering into representative days and (3) do not allow storage technologies. We compare a number of possible implementations of our framework, and discuss their relative (dis)advantages.

These methods fill a research gap. From a modelling perspective, recent studies indicate that *a priori* aggregation using time series inputs alone may lead to significant errors (Section 1.4), while accurately representing storage technologies is increasingly important. From an applied perspective, our methods allow the consideration of multi-year samples in planning models at significantly lower error than current approaches; this is important for robust decision-making under climate uncertainty as discussed in Section 1.1. They also allow the use of climate model data – typically long time series produced by an ensemble of climate models – in planning studies.

This paper is structured as follows. Section 3 introduces the method, including the intuition behind its machinery. Section 4 provides a case study application. In Section 5, we discussion conclusions, implications and possible extensions.

3. Methods

3.1. Overview and intuition

Consider two hypothetical planning models A and B. They take the same time series inputs, but A allows only fossil-fuel technologies, while B contains primarily variable renewables. *A priori* aggregation leads to the same representative periods for both models, even though weather variables are unimportant for model A but very important for model B.

Our framework uses a model's *operational variables* – generation, transmission and storage patterns – to customise time series aggregation. We can do this in two ways. The first is to model a selection of relevant extreme events – as identified by an *importance* function such as generation cost, electricity price or unmet demand – at higher resolution (see Section 3.3 for discussion). This reduces the truncation of extreme events as illustrated in Fig. 4. The second is by using operational variables when clustering. For example, concatenating storage (dis)charge decisions to the vectors we cluster encourages periods with similar storage patterns to be mapped to the same representative.

Unlike time series inputs, operational variables are not available *a priori*. For example, we do not know a model's generation levels before simulations. We hence propose a two-stage approach. We determine a first-stage optimal design estimate $\mathbf{D}_{\mathcal{A}_0}$ using *a priori* aggregation. We then calculate operational variables via an operational model (Eq. (3)) across the full time series given $\mathbf{D}_{\mathcal{A}_0}$. These variables are used in a second planning model with *a posteriori* aggregation.

3.2. Framework: storage importance subsampling

Suppose we have a planning model Φ_{plan} and want to estimate the optimal design $\mathbf{D}_{\mathcal{T}} = \Phi_{\text{plan}}((\xi_t)_{t\in\mathcal{T}})$ across a long sample $(\xi_t)_{t\in\mathcal{T}}$ of

demand and weather data. We estimate it by $\mathbf{D}_{\mathcal{A}_1}$, determined from the following algorithm.

Inputs:

- $(\xi_t)_{t\in\mathcal{T}}$: demand and weather time series, length $n_{\mathcal{T}}$ periods
- $n_{\mathcal{A}}$: number of unique representative periods to aggregate into
- $p_e \in [0,1]$: proportion of periods in $\mathcal{T}$ considered "extreme"
- IMP: real-valued *importance* function of operational variables $\mathbf{O}_t$

Steps:

1. Get preliminary optimal design estimate $\mathbf{D}_{\mathcal{A}_0}$:
   (a) Aggregate $\mathcal{T}$ into $(\hat{\xi}_{\mathcal{A}_0(t)})_{t\in\mathcal{T}}$, $n_{\mathcal{A}}$ unique representative periods, using *a priori* scheme.
   (b) Solve planning problem:
   $$\mathbf{D}_{\mathcal{A}_0} = \Phi_{\text{plan}}((\hat{\xi}_{\mathcal{A}_0(t)})_{t\in\mathcal{T}}). \quad (5)$$
2. Create *importance subsample*:
   (a) Determine system operation across full time series:
   $$(\mathbf{O}_t)_{t\in\mathcal{T}} = \Phi_{\text{operate}}((\xi_t)_{t\in\mathcal{T}} \mid \mathbf{D}_{\mathcal{A}_0}). \quad (6)$$
   (b) Calculate *importance* of each period:
   $$\text{imp}_t = \text{IMP}(\mathbf{O}_t) \quad \text{for all } t \in \mathcal{T}. \quad (7)$$
   (c) Partition time series $\mathcal{T}$ into:
   - $\mathcal{T}_e$: $p_e n_{\mathcal{T}}$ "extreme" periods (with highest *importance*)
   - $\mathcal{T}_r$: $(1-p_e)n_{\mathcal{T}}$ "regular" periods (those remaining).

   (d) Aggregate into $(\hat{\xi}_{\mathcal{A}_1(t)})_{t\in\mathcal{T}}$, $n_{\mathcal{A}}$ unique representative periods, with:
   - $\frac{n_{\mathcal{A}}}{2}$ "extreme" representative periods aggregated from $\mathcal{T}_e$
   - $\frac{n_{\mathcal{A}}}{2}$ "regular" representative periods aggregated from $\mathcal{T}_r$.

   Aggregate using input time series and/or operational variables.
3. Get final optimal design estimate $\mathbf{D}_{\mathcal{A}_1}$:
   $$\mathbf{D}_{\mathcal{A}_1} = \Phi_{\text{plan}}((\hat{\xi}_{\mathcal{A}_1(t)})_{t\in\mathcal{T}}). \quad (8)$$

Output: $\mathbf{D}_{\mathcal{A}_1}$: optimal design estimate

3.3. Remarks

The method above is an *a posteriori* scheme that tailors aggregation to the underlying planning model using (estimated) operational variables. An *a priori* scheme would finish after step 1 and return $\mathbf{D}_{\mathcal{A}_0}$. We instead use $\mathbf{D}_{\mathcal{A}_0}$ to construct a new aggregation $\mathcal{A}_1$ for a second, hopefully more accurate, optimal design estimate $\mathbf{D}_{\mathcal{A}_1}$. We preserve chronology for storage technologies using ordered and linked representative days as discussed in Section 1.3 and Figs. 1–3.

The *importance* function IMP in step 2(b) identifies relevant extreme events. It is one-dimensional to allow stratification into "extreme" and "regular" periods in step 2(c). It should identify events that require peak generation, transmission and storage capacities. In our case study (Section 4), we examine two candidates. The first is the generation cost, which naturally identifies extremes since expensive measures (e.g. peaking plants or load curtailment) are used only in settings where there are otherwise supply shortages. The second is unserved energy, which occurs only at times of insufficient supply. These are roughly those used by Hilbers et al. [47] and Teichgraeber et al. [46] respectively for models without storage. Expert knowledge can motivate others, e.g. electricity price.

Table 1
Simulations for (a) *validation* (Section 4.3) and (b) *example* (Section 4.4) experiments. For simplicity, we detail one *a priori* (B) and one *a posteriori* method (F). For Method F, we disaggregate solution times into first planning, operation and second planning model runs (steps 1(b), 2(a) and 3 in Section 3.2).

Aggregation	Number of (repr.) days	Solution time across 40 runs [minutes] Mean	(2.5–97.5% range)
(a) **Validation**: 3-year base time series			
None (benchmark)	1095[a]	1005	(322–3432)
B (*a priori*)	30	2	(2–4)
F (*a posteriori*)	30	2+18+2 = 22	(16–34)
B (*a priori*)	120	299	(46–658)
F (*a posteriori*)	120	299+19+165 = 483	(133–887)
(b) **Example**: 30-year base time series			
None (benchmark)	10950[a]	*very long, estimated* > 1 *year*	
B (*a priori*)	120	968	(88–1660)
F (*a posteriori*)	120	968+301+717 = 1986	(847–3770)

[a]May include additional leap days

Using this framework requires a number of choices. One is the *importance* function discussed above. Another is p_e, the proportion of periods in the full time series $\mathcal{T}$ considered extreme. This is a trade-off; a larger p_e value (with the number of representative days $n_{\mathcal{A}}$ fixed) means more periods are considered extreme, but are modelled at a lower resolution each (see Fig. 4). For simplicity, we use equal numbers of representative periods ($\frac{n_{\mathcal{A}}}{2}$) for both "extreme" and "regular" regions. We must also specify the aggregation used in steps 1(a) and 2(d); options include e.g. k-means/medoids or hierarchical clustering.

4. Simulation studies

4.1. Overview

In this section we examine the performance of a number of schemes, both *a priori* and *a posteriori* (in the framework of Section 3) on an example energy system planning model. We conduct two experiments as detailed in Table 1. The first is a validation exercise on a relatively short time series. Here, we calculate the "true" (non-aggregated) optimal design $\mathbf{D}_{\mathcal{T}}$ and compare it with aggregated estimates $\mathbf{D}_{\mathcal{A}}$. The second uses a longer time series, for which calculating $\mathbf{D}_{\mathcal{T}}$ directly requires too much computing time; in this case we examine computational costs under aggregation. In both experiments, we calculate *unserved energy* – demand unable to be met by a system with design $\mathbf{D}_{\mathcal{A}}$ – by running an operation problem with this design across the full time series.

We run experiments across the six-region planning model illustrated in Fig. 5. It determines the generation (baseload, peaking and wind), transmission and storage capacities that minimise the sum of install and operation costs given hourly demand levels and wind *generation potentials* (generation as a fraction of rated capacity, also called *capacity factors*). We use this model since it is a benchmark test system, used in other studies; it is based on a renewable version of the *IEEE six-bus* system introduced by Kamalinia et al. [49] and Hilbers et al. [47] and publicly available at doi:10.5281/zenodo.7025098. For details, see Appendix B.

This section is structured as follows. Section 4.2 describes our time series aggregation schemes. Sections 4.3 and 4.4 present results from the validation study (on short base time series) and the example exercise (on longer ones) respectively. Section 4.5 discusses results and their implications.

4.2. Setup

We examine six time series aggregation schemes as detailed in Table 2, all with daily periods. Methods A-C are *a priori*. A and B use the cluster mean and medoid (closest real day to mean) respectively

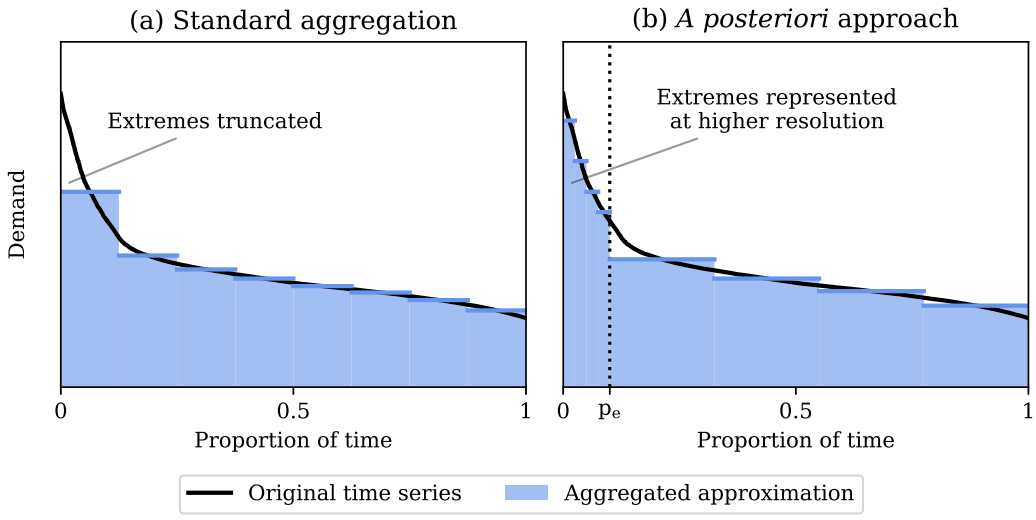
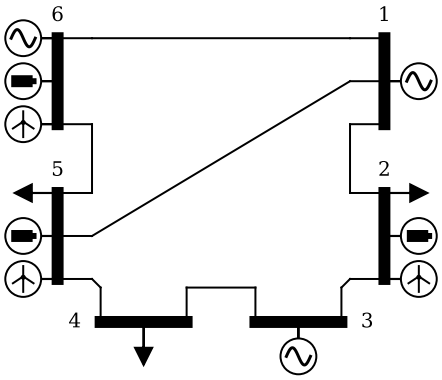


Fig. 4. Impacts of time series aggregation. Black line: *load duration curve* (demand values plotted from highest to lowest). (a) Standard aggregation truncates extremes, removing e.g. high demand events. (b) Our *a posteriori* methods preserve extremes at higher resolution and hence truncate to a smaller degree.



region	technologies	time series
1	baseload, peaking	—
2	wind, storage	demand, wind (DE)
3	baseload, peaking	—
4	—	demand (FR)
5	wind, storage	demand, wind (UK)
6	baseload, peaking, wind, storage	wind (ES)

Fig. 5. Planning model topology. Demand and generation/storage technologies are distributed across six regions, linked by seven transmission lines. Regions 2, 4, 5 and 6 use time series data from Germany (DE), France (FR), the United Kingdom (UK) and Spain (ES) respectively.

Table 2
Time series aggregation schemes: three *a priori* (A–C) and three *a posteriori* (D–E).

Method A (a priori):	Use cluster mean as representative day.
Method B (a priori):	Use cluster medoid (closest real day to mean) as representative day.
Method C (a priori):	Medoid representative, include maximum demand and minimum wind days.
Method D (a posteriori):	Medoid representative, model days with unserved energy at higher resolution.
Method E (a posteriori):	Medoid representative, model days with high generation cost at higher resolution.
Method F (a posteriori):	Medoid representative, model days with high generation cost at higher resolution, cluster on time series inputs and storage patterns.

as representative day. Method C includes the maximum demand and minimum wind days in each region, a common *a priori* way to preserve extremes (Section 1.2). Methods D-F are *a posteriori*, using operational variables as described in Section 3.2. D and E include days with high unserved energy and generation cost, which serve as the *importance* functions and correspond roughly to those used by Teichgraeber et al. [46] and Hilbers et al. [47] respectively in models without storage. Method F is the same as E, but uses storage (dis)charge decisions in the second aggregation (step 2(d)). This is implemented by concatenating storage (dis)charge decisions to the vectors that clustering algorithms are applied to.

We use the following implementation. When we aggregate in steps 1(a) and 2(d), we scale and shift each time series to mean zero and variance one, reshape them to daily vectors and group them using Wald's hierarchical clustering. Representative days are either cluster means or medoids as specified in Table 2. For *a posteriori* methods, we represent $p_e = 0.05$ (5%) of the original time series at higher resolution.

We solve operational problems in step 2(a) sequentially using a horizon of one year and a window of six months. For example, we solve months 1–12 and store months 1–6, then solve 7–18 and store 7–12, etc. To calculate the generation cost in Methods E and F, we assign a value of lost load of £6000/MWh [50] to unserved energy. This value implies ≈1% additional system cost for every 0.01% of energy not met. We also run simulations with different choices than those presented here; these schemes showed similar or worse performance and are discussed in Appendix A.

We repeat experiments 40 times with different base time series $(\xi_t)_{t\in\mathcal{T}}$ created by resampling years with replacement. For example, a three-year sample may be [2011][1992][1992]. For reference, $\mathbf{D}_{\mathcal{T}}$ has mean values (across 40 three-year time series) of 71.5 GW baseload, 62.8 GW peaking, 157.0 GW wind, 153.3 GW transmission and 427.9 GWh storage.

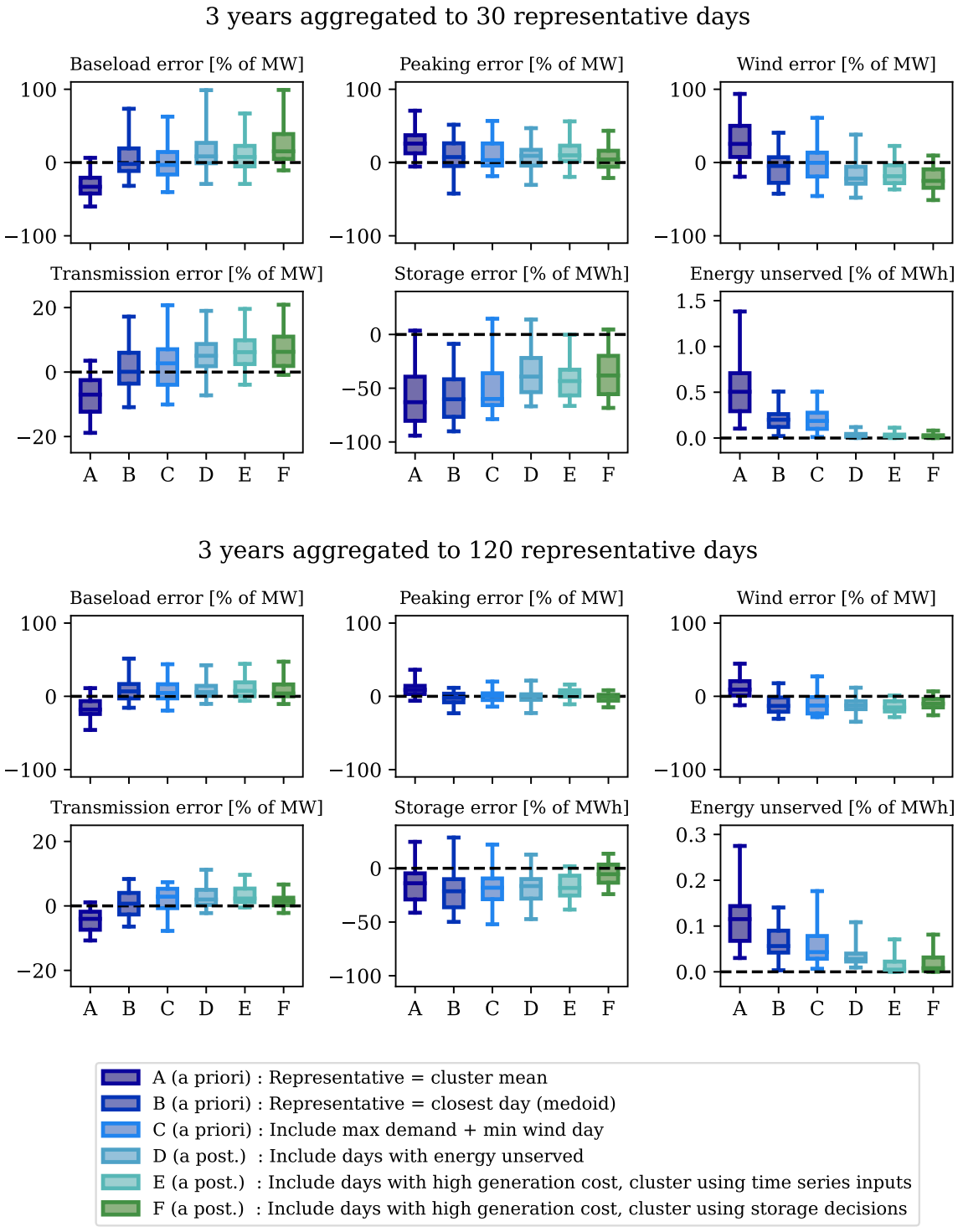


Fig. 6. Distribution of evaluation metrics across 40 simulations for aggregation schemes A-F (Section 4.2, Table 2). The box and whiskers show the 2.5%, 25%, 50%, 75% and 97.5% percentiles. We express values as percentages and denote the original unit, e.g. [% of MW] is a percentage across values with unit MW.

4.3. Validation

Fig. 6 shows results of the validation exercise on six metrics: percentage errors in baseload, peaking, wind, transmission and storage energy capacities (compared to the "true" optimum $\mathbf{D}_{\mathcal{T}}$) as well as levels of unserved energy (MWh) across the full time series. For all methods, increasing the number of representative days from 30 to 120 decreases error metrics, but their relative performances differ. For simplicity, we present only systemwide totals, summed across the model's six regions.

Methods A-C are *a priori*. Method A, with the cluster mean as representative day, overestimates optimal peaking and wind capacities while underestimating optimal baseload, transmission and (by a large margin) storage energy capacities, especially for 30 representative days. We observe unserved energy levels up to 1.5% of demand. Method B, with medoid representative days, performs better; baseload, peaking, wind and transmission capacities are unbiased (median close to true

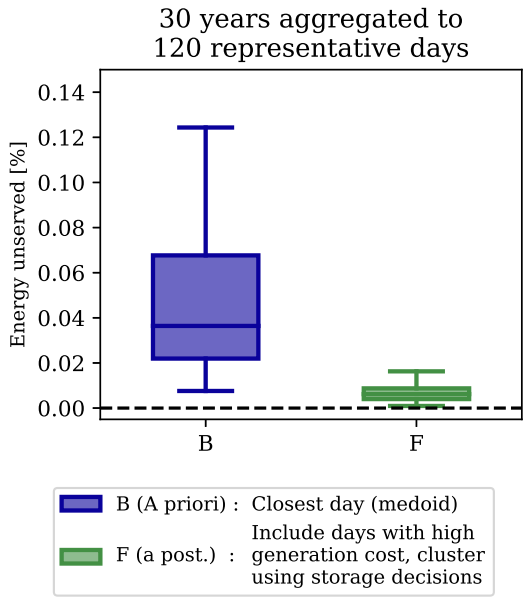


Fig. 7. Distribution of unserved energy (as a percentage of total demand) across 40 simulations for aggregation methods A and F from 30 years to 120 representative days (Section 4.4).

value). Storage remains underestimated, but unserved energy is less than half of Method A. Method C, the *a priori* attempt to include extremes via maximum demand and minimum wind days, does not further enhance performance.

The *a posteriori* Methods D-F have significantly lower levels of unmet demand than A-C. Methods D and E, which identify extremes using unserved energy and generation cost respectively, show similar results as one another. Method F, which uses storage (dis)charge decisions in clustering, more accurately estimates storage capacity.

Table 1(a) shows solution times. For simplicity, we show the best performing *a priori* Method B (A and C have similar solution times) and *a posteriori* Method F (D and E have similar solution times). *A posteriori* solution times consist of two planning runs across representative days and one operational run across the full time series without aggregation. While the operational solution times are constant, the two planning runs take much longer with 120 than 30 representative days.

4.4. Example

Fig. 7 shows the results of aggregating 30 years into 120 representative days. In this case, we are unable to solve the unaggregated benchmark problem; extrapolating from shorter simulation lengths suggests over a year of solution time. We compare Method B with F; these are the *a priori* and *a posteriori* schemes with best performance respectively. While unable to calculate capacity errors – which require the non-aggregated design $\mathbf{D}_{\mathcal{T}}$ – unserved energy levels are around seven times lower on average for Method F. This is a result of system designs that are more robust to extreme events; on average, designs for Method F have 7% more peaking, 5% more transmission and 16% more storage capacity than for Method B.

Table 1(b) shows the distribution of solution times. Method F's solution times are formed mostly of the two planning runs on 120 representative days and not the operational run across the full 30 years.

4.5. Discussion

For the *a priori* methods, using the medoid (closest real day to cluster mean) as representative performs significantly better than using the mean itself. In fact, except in terms of storage energy capacity and unserved energy, it performs similarly to the more complicated *a posteriori* methods. Note that the *a priori* attempt at including extremes via the maximum demand and minimum wind days does not improve performance. As discussed in Section 1.2, this heuristic appears unable to determine those events that truly drive installed capacities.

The main performance enhancements from *a posteriori* methods lie in reduced levels of unserved energy. This is to be expected, since they identify and include relevant extremes, ensuring resultant design estimates are robust to such occurrences. Our simulations suggest that both high generation cost and unserved energy successfully identify relevant extremes and can serve as useful *importance* functions (Section 3.3). The value in using storage (dis)charge decisions when aggregating (Method F) is concentrated primarily in more accurate estimates of optimal storage capacity, which many schemes underestimate significantly.

Different aggregation methods lead to different energy system designs. Designs from *a posteriori* methods have lower wind but higher baseload, transmission and storage capacities. They hence have higher levels of security of supply and are less susceptible to extreme weather events. In fact, Fig. 6 indicates that these systems are (slightly) over-engineered for extreme events; the "true" optimum has slightly more renewables and less baseload, without any additional unserved energy.

Obtaining a design estimate $\mathbf{D}_{\mathcal{A}_1}$ using an *a posteriori* scheme requires two planning model runs (across $n_{\mathcal{A}}$ representative periods each) and one operation model run (across the full time series, length $n_{\mathcal{T}}$ periods). Both the accuracy of design estimates and computational times increase with the number of representative periods. Hence, in practice, the largest number of representative periods for which one can solve the planning problem given available computing resources should be chosen; in this case, the two planning runs usually constitute the majority of solution times, as seen in Section 4.4.

5. Conclusions

5.1. Conclusions

This paper introduces a framework for *a posteriori* time series aggregation schemes for energy system (capacity expansion) planning models with storage. They allow us to estimate optimal system design or investment decisions across long time series at significantly reduced computational cost and with smaller levels of error than established approaches. Our models, data and code are publicly available at doi:10.5281/zenodo.7178301.

The ability to reliably consider long samples in capacity expansion planning problems has two important implications. Firstly, it reduces the impact of sampling uncertainty on model outputs, reducing the risk in incorrect strategical decisions as a result of an unrepresentative demand and weather time series. Secondly, it is a step towards the use of climate model data – typically many multi-decadal samples from an ensemble of simulations – in planning models: without reliable compression techniques this requires too much computing time.

Our framework customises aggregation to the energy system model using its operational variables (generation, transmission and storage patterns). It unifies and generalises methods by Hilbers et al. [47] and Teichgraeber et al. [46] – which use operational variables in models without storage – with that of Kotzur et al. [38] – which allow chronology-preserving aggregation for storage technologies. We also investigate the use of operational variables, such as storage patterns, in clustering (in addition to identifying relevant extreme events).

Users can choose an *a priori* versus an *a posteriori* scheme based on a number of factors. We find that *a posteriori* schemes lead to lower levels of unserved energy than *a priori* methods, even with heuristic adjustments such as including the maximum demand day. Hence, *a posteriori* methods reduce the climate risk in planning; such systems are able to meet demand across a large range of demand and weather scenarios. However, they are also more complicated, requiring two

Table B.3
Nomenclature.

Indices & Sets	
$i \in \mathcal{I}$	Generation technology
$r \in \mathcal{R}$	Region
$t \in \mathcal{T}$	Time step
Parameters	
C_i^{gen}	Annualised generation install cost, technology i [£/MWyr]
$C_{r,r'}^{\text{tr}}$	Annualised transmission install cost, region r to r' [£/MWyr]
C^{sto}	Annualised storage energy install cost [£/MWhyr]
F_i	Generation cost, technology i [£/MWh]
e^{sto}	Storage (dis)charge efficiency [$\in$ [0, 1]]
l^{sto}	Storage self-loss [1/hr]
Time series	
$d_{r,t}$	Demand, region r, time t [MWh]
$\lambda_{i,r,t}$	Generation potential, technology i, region r, time t [$\in$ [0, 1]]
ξ_t	Time series values, time t
Decision variables	
$\text{cap}_{i,r}^{\text{gen}}$	Generation capacity, technology i, region r [MW]
$\text{cap}_{r,r'}^{\text{tr}}$	Transmission capacity, region r to r' [MW]
$\text{cap}_r^{\text{sto}}$	Storage energy capacity, region r [MWh]
$\text{gen}_{i,r,t}$	Generation, technology i, region r, time t [MWh]
$\text{tr}_{r,r',t}$	Transmission, region r to r', time t [MWh]
$\text{ch}_{r,t}$	Storage charging, region r, time t [MWh]
$\text{sto}_{r,t}$	Storage energy level, region r, time t [MWh]
$\mathbf{D}$	Power system design
$\mathbf{O}_t$	Power system operation, time t

planning and one operational model runs and the choice of a number of hyperparameters (see Section 3.3), in particular the *importance* function. We find that medoid-based clustering (which significantly outperforms mean-based methods) provides broadly accurate capacity estimates with a single planning model simulation. A user could, for example, use an *a priori* approach and install spare generation capacity, as opposed to the more accurate, but involved, *a posteriori* methods.

5.2. Extensions

Our framework uses two planning runs to obtain $\mathbf{D}_{\mathcal{A}_0}$ and $\mathbf{D}_{\mathcal{A}_1}$. We can, however, iterate steps 2 and 3, repeatedly using the last design estimate to calculate new operational variables and design estimates $\mathbf{D}_{\mathcal{A}_i}$ for $i > 1$. This is done by Bahl et al. [45] and Teichgraeber et al. [46] with unserved energy as the *importance* function in models without storage. In our experiments, we find further iterations to offer minimal performance gain, but this need not hold in general.

Another extension involves dimensionality reduction. Even in our comparatively simple case study models, our three time series of demand, wind generation potentials and storage (dis)charge decisions give $3 \times 3 \times 24 = 216$ components in each daily vector to cluster in steps 1(a) and 2(d) (Section 3.2). For models with more time series inputs, reducing dimensionality may be necessary.

We may also use more information from operational variables. In our case studies, we use (1) generation cost or unserved energy levels to identify extreme events and (2) storage patterns when clustering. However, the operational variables include more information, such as generation levels of individual technologies and regions, from which we may be able to extract more information. We may also investigate different *importance* functions, such as the electricity price.

Other classes of extensions involve combining different chronology-preserving or *a posteriori* aggregation. For example, we may use the *system states* or *chronological time period clustering* (Section 1.3) to link periods across time, or cluster in solution space as some methods in Section 1.4.

Finally, we can reduce our models' operational foresight. For our planning and operational models, we optimise operation across the full sample and one year ahead respectively. Planning problems without perfect foresight are more realistic but may require different solution methods.

CRediT authorship contribution statement

Adriaan P. Hilbers: Conceptualization, Methodology, Software, Writing, Funding acquisition. **David J. Brayshaw:** Conceptualization, Review, Editing, Supervision. **Axel Gandy:** Conceptualization, Review, Editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The models, data, and full reproducibility code is available at doi: 10.5281/zenodo.7178301.

Acknowledgements

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) Mathematics of Planet Earth Centre for Doctoral Training, United Kingdom, grant number EP/L016613/1.

Appendix A. Simulations not included

A number of experiments showed performance that was either similar or worse than those presented in Section 4. These included using $p_e = 0.1$ instead of 0.05, meaning 10% of the time series was considered "extreme". We also used different vector normalisation, including scaling and shifting time series to lie between zero and one (instead of having mean zero and variance one) and normalising *daily* vectors (so that each hour-of-day in each time series has mean zero and variance one instead of each time series). To determine extreme periods, we included days with peak generation cost or unserved energy instead of maximum integral value. Finally, we calculated additional design estimates $\mathbf{D}_{\mathcal{A}_i}$ for $i > 1$ by repeating steps 2 and 3 in Section 3.2.

Table B.4
Technologies. Install costs are expressed per year of infrastructure lifetime. Carbon emissions are expressed in kg CO_2 equivalent warming potential. Storage efficiency is dimensionless. To avoid solution nonuniqueness, costs are perturbed slightly (< 0.1%) in different regions.

Technology (i)	Monetary cost			Storage	
	Install [£/MWyr]	Install [£/MWhyr]	Generation [£/MWh]	Efficiency [1]	Self-loss [1/hr]
Generation	C_i^{gen}		F_i		
Baseload (b)	300,000	–	5	–	–
Peaking (p)	100,000	–	35	–	–
Wind (w)	100,000	–	–	–	–
Transmission	$C_{r,r'}^{\text{tr}}$				
Region 1–5	150,000	–	–	–	–
Other	100,000	–	–	–	–
Storage		C^{sto}		e^{sto}	l^{sto}
Storage	–	1000	–	0.95	0.00001

Appendix B. Planning models: mathematical details

Our planning model's generation, transmission and storage technologies are detailed in Table B.4. Time series inputs (hourly demand levels in Regions 2, 4 and 5; wind *generation potentials* (capacity factors) in Regions 2,5 and 6) contain data across Europe for 1980–2017, as introduced by Bloomfield et al. [51] and available at [52]. Let $\mathcal{I} = \{b, p, w\}$ and $\mathcal{R} = \{1, 2, 3, 4, 5, 6\}$ be the generation technologies (baseload, peaking, wind) and regions respectively. Then $\xi_t = [d_{2,t}, d_{4,t}, d_{5,t}, \lambda_{w,2,t}, \lambda_{w,5,t}, \lambda_{w,6,t}]$ is the demand and weather data at time t. The planning problem is to minimise

$$\sum_{r\in\mathcal{R}}\left[\frac{T}{8760}\Bigg(\underbrace{\sum_{i\in\mathcal{I}} C_i^{\text{gen}}\text{cap}_{i,r}^{\text{gen}}}_{\substack{\text{install cost,}\\ \text{generation}}} + \underbrace{\sum_{r'\in\mathcal{R}} C_{r,r'}^{\text{tr}}\text{cap}_{r,r'}^{\text{tr}}}_{\substack{\text{install cost,}\\ \text{transmission}}} + \underbrace{C^{\text{sto}}\text{cap}_r^{\text{sto}}}_{\substack{\text{install cost,}\\ \text{storage}}}\Bigg) + \underbrace{\sum_{t\in\mathcal{T}}\sum_{i\in\mathcal{I}} F_i^{\text{gen}}\text{gen}_{i,r,t}}_{\text{generation cost}}\right] \tag{B.1}$$

by optimising over design $\mathbf{D}$ and operation $(\mathbf{O}_t)_{t\in\mathcal{T}}$, where

$$\mathbf{D} = [\text{cap}_{i,r}^{\text{gen}},\ \text{cap}_{r,r'}^{\text{tr}},\ \text{cap}_r^{\text{sto}} \mid i\in\mathcal{I};\ r\in\mathcal{R};\ r'\in\mathcal{R}] \tag{B.2}$$

$$\mathbf{O}_t = [\text{gen}_{i,r,t},\ \text{tr}_{r,r',t},\ \text{ch}_{r,t} \mid i\in\mathcal{I};\ r\in\mathcal{R};\ r'\in\mathcal{R}] \tag{B.3}$$

subject to

$$\text{cap}_{b,r}^{\text{gen}}\Big|_{r\notin\{1,3,6\}} = \text{cap}_{p,r}^{\text{gen}}\Big|_{r\notin\{1,3,6\}} = \text{cap}_{w,r}^{\text{gen}}\Big|_{r\notin\{2,5,6\}} = 0 \tag{B.4}$$

$$\text{cap}_{r,r'}^{\text{tr}}\Big|_{(r,r')\notin\{(1,2),(1,5),(1,6),(2,3),(3,4),(4,5),(5,6)\}} = 0 \tag{B.5}$$

$$\text{cap}_r^{\text{sto}}\Big|_{r\notin\{2,5,6\}} = 0 \tag{B.6}$$

$$\sum_{i\in\mathcal{I}}\text{gen}_{i,r,t} + \sum_{r'\in\mathcal{R}}\text{tr}_{r',r,t} = d_{r,t} + \text{ch}_{r,t} \quad \text{for all } r,t \tag{B.7}$$

$$\text{tr}_{r,r',t} + \text{tr}_{r,'r_t} = 0 \quad \text{for all } r,r',t \tag{B.8}$$

$$\text{sto}_{r,0} = 0 \quad \text{for all } r \tag{B.9}$$

$$\text{sto}_{r,t+1} = (1-l^{\text{sto}})\text{sto}_{r,t} + \begin{cases} e^{\text{sto}}\text{ch}_{r,t} & \text{if } \text{ch}_{r,t}\geq 0 \\ \frac{1}{e^{\text{sto}}}\text{ch}_{r,t} & \text{if } \text{ch}_{r,t}<0 \end{cases} \quad \text{for all } r,t \tag{B.10}$$

$$0 \leq \text{gen}_{i,r,t} \leq \text{cap}_{i,r}^{\text{gen}} \quad \text{for all } i,r,t \tag{B.11}$$

$$0 \leq \text{gen}_{w,r,t} \leq \text{cap}_{w,r}^{\text{gen}}\lambda_{w,r,t} \quad \text{for all } r,t \tag{B.12}$$

$$|\text{tr}_{r,r',t}| \leq \text{cap}_{r,r'}^{\text{tr}} + \text{cap}_{r',r}^{\text{tr}} \quad \text{for all } r,r',t \tag{B.13}$$

$$0 \leq \text{sto}_{r,t} \leq \text{cap}_r^{\text{sto}} \quad \text{for all } r,t. \tag{B.14}$$

For definitions of terms and parameter values, see Tables B.3 and B.4 respectively. The factor $\frac{T}{8760}$ normalises install costs to the same temporal scale as generation costs, since C_i^{gen} and $C_{r,r'}^{tr}$ are costs per year of plant lifetime and there are 8760 hours (time steps) in a year.

The constraints have the following meanings. (B.4)–(B.6) are the model's generation, transmission and storage topology. (B.7) indicates that generation plus transmission into a region equals demand plus storage charging. (B.8) is the transmission balance. (B.9) specifies empty initial storage and (B.10) indicates how storage levels change with storage (dis)charging and self-loss. (B.11)–(B.12) ensure generation does not exceed installed capacity (for thermal technologies) or installed capacity times generation potential (for wind). (B.13) limits transmitted power to installed transmission capacity. (B.14) constrains storage levels to lie within storage (energy) bounds.

References

[1] Kools L, Phillipson F. Data granularity and the optimal planning of distributed generation. Energy 2016;112:342–52. http://dx.doi.org/10.1016/j.energy.2016.06.089.

[2] 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;162:631–43. http://dx.doi.org/10.1016/j.apenergy.2015.10.100.

[3] Collins S, Deane JP, Poncelet K, Panos E, Pietzcker RC, Delarue E, Ó Gallachóir BP. Integrating short term variations of the power system into integrated energy system models: A methodological review. Renew Sustain Energy Rev 2017;76:839–56. http://dx.doi.org/10.1016/j.rser.2017.03.090.

[4] Bloomfield HC, Brayshaw DJ, Shaffrey LC, Coker PJ, Thornton HE. Quantifying the increasing sensitivity of power systems to climate variability. Environ Res Lett 2016;11(12):124025. http://dx.doi.org/10.1088/1748-9326/11/12/124025.

[5] Staffell I, Pfenninger S. The increasing impact of weather on electricity supply and demand. Energy 2018;145:65–78. http://dx.doi.org/10.1016/j.energy.2017.12.051.

[6] Zeyringer M, Price J, Fais B, Li PH, Sharp E. Designing low-carbon power systems for great britain in 2050 that are robust to the spatiotemporal and inter-annual variability of weather. Nat Energy 2018;3(5):395–403. http://dx.doi.org/10.1038/s41560-018-0128-x.

[7] Collins S, Deane P, Ó Gallachóir B, Pfenninger S, Staffell I. Impacts of inter-annual wind and solar variations on the European power system. Joule 2018;2(10):2076–90. http://dx.doi.org/10.1016/j.joule.2018.06.020.

[8] Hilbers AP, Brayshaw DJ, Gandy A. Importance subsampling: improving power system planning under climate-based uncertainty. Appl Energy 2019;251:113114. http://dx.doi.org/10.1016/j.apenergy.2019.04.110.

[9] Kumler A, Carreño IL, Craig MT, Hodge BM, Cole W, Brancucci C. Inter-annual variability of wind and solar electricity generation and capacity values in Texas. Environ Res Lett 2019;14(4):44032. http://dx.doi.org/10.1088/1748-9326/aaf935.

[10] Bryce R, Losada Carreño I, Kumler A, Hodge BM, Roberts B, Brancucci Martinez-Anido C. Consequences of neglecting the interannual variability of the solar resource: A case study of photovoltaic power among the Hawaiian Islands. Sol Energy 2018;167(January):61–75. http://dx.doi.org/10.1016/j.solener.2018.03.085.

[11] Shaner MR, Davis SJ, Lewis NS, Caldeira K. Geophysical constraints on the reliability of solar and wind power in the United States. Energy Environ Sci 2018;11(4):914–25. http://dx.doi.org/10.1039/c7ee03029k.

[12] Hilbers AP, Brayshaw DJ, Gandy A. Efficient quantification of the impact of demand and weather uncertainty in power system models. IEEE Trans Power Syst 2021;36(3):1771–9. http://dx.doi.org/10.1109/TPWRS.2020.3031187.

[13] Cao KK, Von Krbek K, Wetzel M, Cebulla F, Schreck S. Classification and evaluation of concepts for improving the performance of applied energy system optimization models. Energies 2019;12(24). http://dx.doi.org/10.3390/en12244656.

[14] Goderbauer S, Comis M, Willamowski FJ. The synthesis problem of decentralized energy systems is strongly NP-hard. Comput Chem Eng 2019;124:343–9. http://dx.doi.org/10.1016/j.compchemeng.2019.02.002.

[15] Bloomfield HC, Gonzalez PLM, Lundquist JK, Stoop LP, Kies A, Browell J, Dargaville R, De Felice M, Gruber K, Hilbers AP, Panteli M, Thornton HE, Wohland J, Zeyringer M, Brayshaw DJ. The importance of weather and climate to energy systems: A workshop on next generation challenges in energy-climate modelling. Bull Am Meteorol Soc 2021;1–23. http://dx.doi.org/10.1175/bams-d-20-0256.1.

[16] Craig MT, Wohland J, Stoop LP, Kies A, Pickering B, Bloomfield HC, Browell J, De Felice M, Dent CJ, Deroubaix A, Frischmuth F, Gonzalez PL, Grochowicz A, Gruber K, Härtel P, Kittel M, Kotzur L, Labuhn I, Lundquist JK, Pflugradt N, van der Wiel K, Zeyringer M, Brayshaw DJ. Overcoming the disconnect between energy system and climate modeling. Joule 2022;6(7):1405–17. http://dx.doi.org/10.1016/j.joule.2022.05.010.

[17] Pfenninger S, Hawkes A, Keirstead J. Energy systems modeling for twenty-first century energy challenges. Renew Sustain Energy Rev 2014;33:74–86. http://dx.doi.org/10.1016/j.rser.2014.02.003.

[18] Koltsaklis NE, Dagoumas AS. State-of-the-art generation expansion planning: A review. Appl Energy 2018;230(August):563–89. http://dx.doi.org/10.1016/j.apenergy.2018.08.087.

[19] Hoffmann M, Kotzur L, Stolten D, Robinius M. A review on time series aggregation methods for energy system models. Energies 2020;13(3):641. http://dx.doi.org/10.3390/en13030641.

[20] Teichgraeber H, Brandt AR. Time-series aggregation for the optimization of energy systems: Goals, challenges, approaches, and opportunities. Renew Sustain Energy Rev 2022;157:111984. http://dx.doi.org/10.1016/j.rser.2021.111984.

[21] Welsch M, Howells M, Bazilian M, DeCarolis JF, Hermann S, Rogner HH. Modelling elements of smart grids - enhancing the OSeMOSYS (Open Source Energy Modelling System) code. Energy 2012;46(1):337–50. http://dx.doi.org/10.1016/j.energy.2012.08.017.

[22] De Sisternes FJ, Webster MD. Optimal selection of sample weeks for approximating the net load in generation planning problems. MIT ESD working paper series 0, Massachusetts Institute of Technology; 2013.

[23] Poncelet K, Hoschle H, Delarue E, Virag A, Drhaeseleer W. Selecting representative days for capturing the implications of integrating intermittent renewables in generation expansion planning problems. IEEE Trans Power Syst 2017;32(3):1936–48. http://dx.doi.org/10.1109/TPWRS.2016.2596803.

[24] 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.

[25] Härtel P, Kristiansen M, Korpås M. Assessing the impact of sampling and clustering techniques on offshore grid expansion planning. Energy Procedia 2017;137:152–61. http://dx.doi.org/10.1016/j.egypro.2017.10.342.

[26] 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. http://dx.doi.org/10.1016/j.apenergy.2017.03.051.

[27] Kotzur L, Markewitz P, Robinius M, Stolten D. Impact of different time series aggregation methods on optimal energy system design. Renew Energy 2018;117:474–87. http://dx.doi.org/10.1016/j.renene.2017.10.017.

[28] Kittel M, Hobbie H, Dierstein C. Temporal aggregation of time series to identify typical hourly electricity system states: A systematic assessment of relevant cluster algorithms. Energy 2022;247:123458. http://dx.doi.org/10.1016/j.energy.2022.123458.

[29] Merrick JH. On representation of temporal variability in electricity capacity planning models. Energy Econ 2016;59:261–74. http://dx.doi.org/10.1016/j.eneco.2016.08.001.

[30] Teichgraeber H, Lindenmeyer CP, Baumgärtner N, Kotzur L, Stolten D, Robinius M, Bardow A, Brandt AR. Extreme events in time series aggregation: A case study for optimal residential energy supply systems. Appl Energy 2020;275:115223. http://dx.doi.org/10.1016/j.apenergy.2020.115223.

[31] Pineda S, Morales JM. Chronological time-period clustering for optimal capacity expansion planning with storage. IEEE Trans Power Syst 2018;33(6):7162–70. http://dx.doi.org/10.1109/TPWRS.2018.2842093.

[32] Tso WW, Demirhan CD, Heuberger CF, Powell JB, Pistikopoulos EN. A hierarchical clustering decomposition algorithm for optimizing renewable power systems with storage. Appl Energy 2020;270:115190. http://dx.doi.org/10.1016/j.apenergy.2020.115190.

[33] De Guibert P, Shirizadeh B, Quirion P. Variable time-step: A method for improving computational tractability for energy system models with long-term storage. Energy 2020;213:119024. http://dx.doi.org/10.1016/j.energy.2020.119024.

[34] 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.

[35] 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;125:747–70. http://dx.doi.org/10.1016/j.energy.2017.02.081.

[36] Gabrielli P, Gazzani M, Martelli E, Mazzotti M. Optimal design of multi-energy systems with seasonal storage. Appl Energy 2018;219:408–24. http://dx.doi.org/10.1016/j.apenergy.2017.07.142.

[37] Tejada-Arango DA, Domeshek M, Wogrin S, Centeno E. Enhanced representative days and system states modeling for energy storage investment analysis. IEEE Trans Power Syst 2018;33(6):6534–44. http://dx.doi.org/10.1109/TPWRS.2018.2819578.

[38] Kotzur L, Markewitz P, Robinius M, Stolten D. Time series aggregation for energy system design: Modeling seasonal storage. Appl Energy 2018;213:123–35. http://dx.doi.org/10.1016/j.apenergy.2018.01.023.

[39] 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. http://dx.doi.org/10.1016/j.apenergy.2019.04.030.

[40] Novo R, Marocco P, Giorgi G, Lanzini A, Santarelli M, Mattiazzo G. Planning the decarbonisation of energy systems: The importance of applying time series clustering to long-term models. Energy Convers Manage: X 2022;15(August):100274. http://dx.doi.org/10.1016/j.ecmx.2022.100274.

[41] Gonzato S, Bruninx K, Delarue E. Long term storage in generation expansion planning models with a reduced temporal scope. Appl Energy 2021;298:117168. http://dx.doi.org/10.1016/j.apenergy.2021.117168.

[42] Wogrin S. Time series aggregation for optimization: One-size-fits-all? 2022, arXiv preprint arXiv:2206.03186.

[43] 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(4):2925–36. http://dx.doi.org/10.1109/TPWRS.2019.2892619.

[44] Zhang Y, Cheng V, Mallapragada DS, Song J, He G. A model-adaptive clustering-based time aggregation method for low-carbon energy system optimization. IEEE Trans Sustain Energy 2022;1–11. http://dx.doi.org/10.1109/tste.2022.3199571.

[45] 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(January):1–13. http://dx.doi.org/10.3389/fenrg.2017.00035.

[46] Teichgraeber H, Küpper LE, Brandt AR. Designing reliable future energy systems by iteratively including extreme periods in time-series aggregation. Appl Energy 2021;304:117696. http://dx.doi.org/10.1016/j.apenergy.2021.117696.

[47] Hilbers AP, Brayshaw DJ, Gandy A. Importance subsampling for power system planning under multi-year demand and weather uncertainty. In: 2020 international conference on probabilistic methods applied to power systems (PMAPS 2020). IEEE; 2020, p. 1–6. http://dx.doi.org/10.1109/PMAPS47429.2020.9183591.

[48] Li C, Conejo AJ, Siirola JD, Grossmann IE. On representative day selection for capacity expansion planning of power systems under extreme operating conditions. Int J Electr Power Energy Syst 2022;137(September 2021). http://dx.doi.org/10.1016/j.ijepes.2021.107697.

[49] Kamalinia S, Shahidehpour M, Khodaei A. Security-constrained expansion planning of fast-response units for wind integration. Electr Power Syst Res 2011;81(1):107–16. http://dx.doi.org/10.1016/j.epsr.2010.07.017.

[50] Elexon. Value of lost load review process. Technical report, Elexon; 2018, URL: https://www.elexon.co.uk/wp-content/uploads/2017/09/33_278_10_VoLL-Review-Process-Paper-v1.0.pdf.

[51] Bloomfield HC, Brayshaw DJ, Charlton-Perez AJ. Characterizing the winter meteorological drivers of the European electricity system using targeted circulation types. Meteorol Appl 2019. http://dx.doi.org/10.1002/met.1858.

[52] Bloomfield HC, Brayshaw DJ, Charlton-Perez A. MERRA2 derived time series of European country-aggregate electricity demand, wind power generation and solar power generation. University of Reading; 2020, http://dx.doi.org/10.17864/1947.239, Dataset.