



# The impact of short-term variability and uncertainty on long-term power planning

Henrik C. Bylling<sup>1</sup> · Salvador Pineda<sup>2</sup> · Trine K. Boomsma<sup>1</sup>

Published online: 19 November 2018

© Springer Science+Business Media, LLC, part of Springer Nature 2018

## Abstract

Traditionally, long-term investment planning models have been the apparent tool to analyse future developments in the energy sector. With the increasing penetration of renewable energy sources, however, the modelling of short-term operational issues becomes increasingly important in two respects: first, in relation to variability and second, with respect to uncertainty. A model that includes both may easily become intractable, while the negligence of variability and uncertainty may result in sub-optimal and/or unrealistic decision-making. This paper investigates methods for aggregating data and reducing model size to obtain tractable yet close-to-optimal investment planning decisions. The aim is to investigate whether short-term variability or uncertainty is more important and under which circumstances. In particular, we consider a generation expansion problem and compare various representations of short-term variability and uncertainty of demand and renewable supply. The main results are derived from a case study on the Danish power system. Our analysis shows that the inclusion of representative days is crucial for the feasibility and quality of long-term power planning decisions. In fact, we observe that short-term uncertainty can be ignored if a sufficient number of representative days is included.

**Keywords** Power planning problems · Investment · Variability · Uncertainty · Generation expansion

## List of symbols

### Sets

- $\mathcal{G}$  Set of production units  
 $\mathcal{G}^w$  Set of wind production units  
 $T$  Set of time periods  
 $T_d$  Set of time periods, except the last, within an aggregation period (e.g.  $\{1, \dots, 23\}$  for a day)  
 $S$  Set of scenarios for short-term uncertainty

---

✉ Henrik C. Bylling  
bylling@math.ku.dk

<sup>1</sup> Department of Mathematics, University of Copenhagen, Copenhagen, Denmark

<sup>2</sup> Department of Electrical Engineering, University of Malaga, Málaga, Spain

## Parameters

$c_g^I$	Linear investment cost of unit $g$ (€/MW)
$c_g$	Linear production cost of unit $g$ (€/MWh)
$c_g^+$	Additional cost of upward balancing of unit $g$ (€/MWh)
$c_g^-$	Opportunity cost of downward balancing of unit $g$ (€/MWh)
$r_g^D$	Ramp down rate of unit $g$ (p.u.)
$r_g^U$	Ramp up rate of unit $g$ (p.u.)
$\rho_{gt}$	Predicted production factor of unit $g$ at time $t$ (p.u.)
$\hat{\rho}_{gts}$	Realised production factor of unit $g \in \mathcal{G}^w$ at time $t$ in scenario $s$ (p.u.)
$\kappa$	Minimum wind penetration (%)
$v^L$	Cost of load shedding (€/MWh)
$v^S$	Cost of wind curtailment (€/MWh)
$v_t$	Load factor at time $t$ (p.u.)
$\bar{d}$	Maximum load (MWh)
$\tau_t$	Duration of time period $t$
$\pi_s$	Probability of short-term scenario $s$

## Variables

$\bar{p}_g$	Investment capacity of unit $g$
$p_{gt}$	Scheduled production of unit $g$ at time $t$
$k_t$	Scheduled load shedding at time $t$
$l_t$	Scheduled wind curtailment at time $t$
$p_{gts}^+$	Real-time upward balancing of unit $g \in \mathcal{G} \setminus \mathcal{G}^w$ at time $t$ in scenario $s$
$p_{gts}^-$	Real-time downward balancing of unit $g \in \mathcal{G} \setminus \mathcal{G}^w$ at time $t$ in scenario $s$
$\hat{p}_{gts}$	Real-time production of unit $g \in \mathcal{G} \setminus \mathcal{G}^w$ at time $t$ in scenario $s$
$\Delta k_{ts}$	Real-time load shedding at time $t$ in scenario $s$
$\Delta l_{ts}$	Real-time regulating wind curtailment at time $t$ in scenario $s$

## 1 Introduction

Long-term planning problems related to the electricity market, system and/or network arise in multiple contexts: generation expansion (Baringo and Conejo 2012; Jin et al. 2011; Pineda and Morales 2016), transmission expansion (Orfanos et al. 2013; Hemmati et al. 2014; Pozo et al. 2013), storage investment (Babrowski et al. 2015; Jabr et al. 2015; Ghofrani et al. 2013) etc. Fundamental to all of these problems is the modelling of short-term system operation, ideally accounting for both dynamics and uncertainty. With the penetration of renewable energy sources in many power systems, not only demand but also part of supply vary over time and stochastically. For instance, wind and solar power production is driven by weather conditions and thereby varies from hour to hour and from day to day and is difficult to accurately predict. To maintain the balance between demand and supply at all times, the system should be sufficiently flexible. The increasing penetration of renewables implies a greater need for flexibility in conventional generation and accentuates the effects of inter-temporal constraints and balancing costs (Poncelet et al. 2016). In particular, if demand is higher or lower than renewable production, conventional generation sources must be able to increase or decrease production accordingly. To handle variations over time, production must be able to adjust from one hour to the next. This type of flexibility is restricted by the technical specifications of the operating units, usually modelled by so-called ramping

constraints. Stochasticity is often handled through the division of the market into a day-ahead market for commitment of predicted demand and supply and a real-time balancing market that allows for adjustments at additional costs. In theory, these short-term characteristics could be explicitly modelled in the long-term planning. In practise, however, the computational effort to solve the planning problem becomes excessive (Poncelet et al. 2016). In fact, a complete representation of ramping abilities requires an hourly discretisation of a multi-decade time horizon, whereas the modelling of balancing decisions involves discretising the continuous distribution of demand and renewable supply. As a result, the model size increases with the number of time periods and the number of scenarios describing uncertainty.

The negligence of variability and uncertainty may result in sub-optimal and/or unrealistic decision-making. Indeed, failure to account for ramping restrictions and balancing costs may significantly overestimate flexibility and suggest investments in renewable energy sources beyond what is the physically and/or economically feasible. A compromise between computational effort and accuracy of the model results is provided by aggregated representations of time and uncertainty. The present paper investigates and compares methods for aggregating data to obtain tractable yet close-to-optimal investment planning decisions. We consider the following types of data aggregation:

- Representative hours: Hours are divided into a number of groups, each representing a given number of hours. The division is based on clustering of data and carried out for each hour independently.
- Representative days: Days are divided into a number of groups, each representing a given number of days. The division is likewise based on clustering of data but carried out for a day at the time, respecting the chronology of the hours.
- With short-term uncertainty: The distributions of unknown parameters are discretised, using a limited number of scenarios.
- Without short-term uncertainty: The distributions of unknown parameters are replaced by their expected values.

By disentangling short-term variability and uncertainty, we investigate which is more important, under which circumstances and how to obtain suitable data representations. To the best of our knowledge and as evidenced by the following review, the comparison of such modelling aspects cannot be found in the existing literature.

In evaluating the impact of short-term dynamics and uncertainty, we use a family of generation expansion models. All models take the perspective of a central planner, minimising total costs of meeting demand and a minimum requirement for renewable supply by investing and operating accordingly. We consider a planning horizon of a single year and with an hourly discretisation. Investments are one-time installations, whereas production decisions are made for every time period. Operation is subject to a number of technical constraints, including ramping restrictions, and the structure of the market, including a day-ahead market for commitment of predicted demand and supply and a real-time balancing market. We consider energy-only markets and the implications of short-term uncertainty and dynamics in these. Other related markets, e.g. capacity markets, are not considered.

The performance of these models is compared with respect to both the quality of the expansion plan and the computation time. The model results are illustrated for a case study on the Danish power system.

## 2 Literature review

Various methods have been suggested to discretise the time horizon of long-term planning problems in a way that enables computational tractability (Haydt et al. 2011). Most of these aim at aggregating hours, days and years to achieve an acceptable model size.

With an hourly representation (often referred to as 'time slices'), time periods are represented by the values of their state variables (demand, wind power production, etc.) and grouped according to these. A traditional example is the load duration curve for which time periods are sorted with respect to demand level and grouped into blocks of a given duration (Stoft 2002). This approach is used in the generation expansion planning of Pineda et al. (2014), Chaton and Doucet (2003) and Jin et al. (2011). Bertsch and Fichtner (2015) likewise use the PERSEUS-NET model with a load duration curve in multi-criteria analysis of power generation and transmission planning. As an alternative to sorting the hours throughout the year, demand can be clustered according to additional information such as seasonal variation (Pozo et al. 2013). Baringo and Conejo (2013) include both wind production and demand profiles in the clustering, and the correlation between the two variables is taken into account. With the same purpose, Wogrin et al. (2014) introduce another method based on a duration curve and chronological transitions between states. The method estimates transitions from one state to another and incorporates inter-connecting constraints on the significant transitions.

Representative days consists of choosing a number of days, or connected time periods in general, to represent the planning horizon. In this way, inter-temporal dynamics can be preserved within the time periods. An example is given by Babrowski et al. (2015) who investigate long-term storage planning. Another example is by Ghofrani et al. (2013) who use a representative scheduling period of 24 h to optimize storage placement. A fully dynamic model including all hours of the entire planning horizon have been proposed by Jabr et al. (2015). The model relates to storage investment and relies on robust optimisation. As the fully dynamic setup may very well be intractable for larger problems, Pina et al. (2013) use 12 representative days, 3 of each season in a year, for generation expansion in electricity systems with high penetrations of renewable energy. In a similar fashion, Ma et al. (2013) select five whole weeks to represent demand variations throughout the year in a unit construction and commitment problem and use this to analyse power system flexibility. Representative weeks are also used by Sisternes and Webster (2013) to approximate the net load in a generation expansion problem. The paper demonstrates how the quality of investments depends on the choice of representative weeks. In contrast, however, Nahmmacher et al. (2016) present a clustering technique that determines the representative days and show that using representative weeks instead of days increases the required number of hours to obtain a sufficiently good representation of the data.

The main difference between hourly and daily aggregation is the ability to include short-term operational flexibility. A daily representation may account for short-term flexibility by including inter-temporal constraints. Such constraints cannot be incorporated with an hourly representation. With an increasing penetration of non-dispatchable renewable energy sources, however, the representation of short-term dynamics in long-term models becomes increasingly important (Pfenninger et al. 2014). This is supported by Poncelet et al. (2016) who study and compare the effect of using the hourly and daily representations. More specifically, the results confirm that the need for inter-temporal techno-economic constraints increases with the penetration of renewable energy. Slednev et al. (2017) consider a combination of representative days and hours. These are determined in a k-means clustering method, using an error measure that measures grid-impact. The time resolution for both the hourly and the daily aggregation is usually hours. An example of a finer time resolution (such as 15 or

30 minutes) is provided by Schwarz et al. (2018) who analyse residential heat storage with photo-voltaic power generation.

In addition to variability over time, long-term planning naturally involves uncertainty of key parameters. Long-term uncertainty relates to the future development of demand and costs parameters. Nevertheless, uncertainty also arises in the short-term. Traditionally, the main concern has been the stochastic variability of demand. At present, however, demand can be rather accurately predicted 24 h in advance (Aneiros et al. 2016). On the other hand, a high penetration of renewable electricity sources in modern power systems introduces a significant source of short-term uncertainty. Some authors ignore this short-term uncertainty by assuming perfect information of future power production and model system operation as deterministic. This approach can be found in the generation expansion problem of Jin et al. (2011), who assume long-term demand and price levels to be uncertain but solve the short-term scheduling problem with perfect information. The same approach is taken by Pozo et al. (2013) and Ludig et al. (2011).

In contrast, Baringo and Conejo (2013, 2011) and Ma et al. (2013) model the system operation problem as a two-stage stochastic program with a day-ahead market as the first stage and a balancing market as the second stage. In the day-ahead market, production decisions are made according to expected demand and wind production. Uncertainty in the real-time balancing market is modelled by scenarios for wind production. In each scenario, day-ahead commitments can be adjusted to realised production by making balancing decisions (with potentially additional costs). Further details are provided by Pineda and Morales (2016).

The remainder of this paper is organised in the following way. The generation expansion problem is presented in Sect. 3.1 and the different approaches to including short-term characteristics are discussed in Sect. 3.2. Section 4 provides a small example which serves as a basis for analysing the effects of short-term variability and uncertainty on the solutions to the generation expansion problem. A larger case study further elaborates on this in Sect. 5. Section 6 concludes the paper.

### 3 Investment optimisation and aggregation of data

The purpose of this paper is to compare different approaches to represent short-term dynamics and uncertainty in long-term planning problems. In particular, we consider four different approaches for aggregation of data in a generation expansion problem. All aggregation approaches are used in combination with the same optimisation model. The model is presented in Sect. 3.1 whereas the data aggregation approaches are defined in Sect. 3.2.

#### 3.1 Model

The model takes the perspective of a central planner, with the objective of minimising total costs of meeting a fixed demand by investing and operating accordingly. The decisions obtained by this model coincide with those of a generation expansion equilibrium with perfectly competitive and risk-neutral power producers (Ehrenmann and Smeers 2011). Furthermore, since demand is assumed to be inelastic, minimising costs is equivalent to maximising social welfare (Gabriel et al. 2010). The model includes a minimum wind penetration constraint that represents a political target for the penetration of renewables, like those imposed as part of the political agenda in the European Union (European Commission 2014). Note that the minimum wind penetration is given as an exogenous parameter, whereas

the decision to invest in wind capacity is endogenous to the model. For simplicity, we focus on short-term uncertainty of wind power production, although the model could easily be extended to include demand uncertainty, stochastic capacity availability etc.

The modelling of generation expansion is divided into two: investment and market clearing.

### 3.1.1 Investment

Generally, generation expansion models are classified as either static (single-year) or dynamic (multiple-year) models (Akbari et al. 2012). For simplicity and as is common practice in the literature (Baringo and Conejo 2012; Wang et al. 2009; Murphy and Smeers 2005), our investment model is static with a single-year planning horizon. Thus, investment variables relate to a one-time installation of generation capacity while system operation involves production decisions for every time period (e.g. hour) throughout the year (in the following referred to as the target year). We assume that at the beginning of the year, there is no existing capacity in the market, i.e. we take a greenfield approach. We also assume that new generation capacity is available once installed, meaning that construction time is zero.

### 3.1.2 Market clearing

Our market model consists of a day-ahead market and a real-time balancing market (Pineda et al. 2016). For each time period, the day-ahead market is modelled as an economic dispatch problem in which the generating units are dispatched to meet demand at minimal costs given the forecasted wind power production. In the balancing market, stochastic wind power production is realised, and the demand must be met given this realised wind power production. This may require re-dispatch of power generation and incurs an additional (positive) balancing cost. Such costs may be justified by increased stress on the generation units. Balancing costs are further discussed in Sect. 5.1.

Our techno-economical constraints include ramping constraints of the generation units, but for simplicity, we do not consider a unit commitment problem (Poncelet et al. 2016). This simplification is likewise discussed in Sect. 3.2.4.

To ensure that the expansion problem is feasible irrespective of the investment plan, we include the possibility of load shedding and wind curtailment during economic dispatch. If installed capacity is insufficient to meet demand, load shedding occurs. Likewise, if the realised wind power production exceeds demand, wind curtailment occurs. Load shedding and wind curtailment are present in both the day-ahead and the balancing market. The realised load shedding or wind curtailment is given as the sum of the scheduled load shedding or wind curtailment and the adjustments to these. Only the realised load shedding or wind curtailment is penalised in the objective function. As an estimate for load shedding and wind curtailment costs, we use the maximum and minimum price caps for the market in question (Stoft 2002). These costs serve as compensation to the consumer and the wind power producer, respectively, should load shedding or wind curtailment occur.

The static generation expansion problem is formulated as follows:

$$\begin{aligned} \min \quad & \sum_{g \in \mathcal{G}} c_g^I \bar{p}_g + \sum_{t \in T} \tau_t \left( \sum_{g \in \mathcal{G}} c_g P_{gt} + \sum_{g \in \mathcal{G} \setminus \mathcal{G}^w} \sum_{s \in S} \pi_s \left( (c_g + c_g^+) P_{gts}^+ - (c_g - c_g^-) P_{gts}^- \right) \right) \\ & + \sum_{t \in T} \tau_t \sum_{s \in S} \pi_s \left( v^L (k_t + \Delta k_{ts}) + v^S (l_t + \Delta l_{ts}) \right) \end{aligned} \tag{1a}$$

$$\text{s.t. } \sum_{g \in \mathcal{G} \setminus \mathcal{G}^w} p_{gt} + k_t - l_t = v_t \bar{d} - \sum_{g \in \mathcal{G}^w} \rho_{gt} \bar{p}_g, \quad \forall t \in T \tag{1b}$$

$$0 \leq p_{gt} \leq \bar{p}_g, \quad \forall g \in \mathcal{G} \setminus \mathcal{G}^w, \forall t \in T \tag{1c}$$

$$-r_g^D \bar{p}_g \leq p_{gt} - p_{g(t-1)} \leq r_g^U \bar{p}_g, \quad \forall g \in \mathcal{G} \setminus \mathcal{G}^w, t \in T_d \tag{1d}$$

$$0 \leq k_t \leq v_t \bar{d}, \quad \forall t \in T \tag{1e}$$

$$0 \leq l_t \leq \sum_{g \in \mathcal{G}^w} \rho_{gt} \bar{p}_g, \quad \forall t \in T \tag{1f}$$

$$\sum_{g \in \mathcal{G} \setminus \mathcal{G}^w} \tilde{p}_{gts} + \Delta k_{ts} - \Delta l_{ts} = v_t \bar{d} - \sum_{g \in \mathcal{G}^w} \tilde{\rho}_{gts} \bar{p}_g, \quad \forall t \in T, s \in S \tag{1g}$$

$$0 \leq \tilde{p}_{gts} \leq \bar{p}_g, \quad \forall g \in \mathcal{G} \setminus \mathcal{G}^w, t \in T, s \in S \tag{1h}$$

$$-r_g^D \bar{p}_g \leq \tilde{p}_{gts} - \tilde{p}_{g(t-1)s} \leq r_g^U \bar{p}_g, \quad \forall g \in \mathcal{G} \setminus \mathcal{G}^w, s \in S, t \in T_d \tag{1i}$$

$$0 \leq k_t + \Delta k_{ts} \leq v_t \bar{d}, \quad \forall t \in T, s \in S \tag{1j}$$

$$0 \leq l_t + \Delta l_{ts} \leq \sum_{g \in \mathcal{G}^w} \tilde{\rho}_{gts} \bar{p}_g, \quad \forall t \in T, s \in S \tag{1k}$$

$$\tilde{p}_{gts} = p_{gt} + p_{gts}^+ - p_{gts}^-, \quad \forall g \in \mathcal{G} \setminus \mathcal{G}^w, t \in T, s \in S \tag{1l}$$

$$\sum_{t \in T} \tau_t \sum_{s \in S} \pi_s \left( \sum_{g \in \mathcal{G}^w} \tilde{\rho}_{gts} \bar{p}_g - (l_t + \Delta l_{ts}) \right) \geq \kappa \sum_{t \in T} \sum_{s \in S} \pi_s \tau_t (v_t \bar{d} - (k_{ts} + \Delta k_{ts})) \tag{1m}$$

$$p_{gts}^+, p_{gts}^- \geq 0, \quad \forall g \in \mathcal{G} \setminus \mathcal{G}^w, t \in T, s \in S \tag{1n}$$

$$\bar{p}_g \geq 0, \quad \forall g \in \mathcal{G} \tag{1o}$$

The objective function in (1a) accumulates investment costs, day-ahead planning costs and expected real-time balancing costs, including load shedding and wind curtailment costs. Day-ahead planning costs of time period  $t$  are weighted by the parameter  $\tau_t$  due to the aggregation of data (see Sect. 3.2.1) and balancing costs of scenario  $s$  are weighted by the probability  $\pi_s$ . The day-ahead operational constraints (1b), (1c), (1d), (1e) and (1f) cover demand satisfaction, capacity limits, ramp rate restrictions and bounds on load shedding and wind curtailment, respectively. Note that there are no ramping constraints between aggregation periods (e.g. days), as indicated by the set  $T_d$ , see also Sect. 3.2.2. Similar operating constraints apply to the balancing market in (1g), (1h), (1i), (1k) and (1j). Moreover, the realised dispatch of generation is defined in (1l). Finally, (1m) requires a percentage of  $\kappa$  of the annual demand to be covered by wind power production, while (1n), (1o) ensure non-negativity of the relevant variables.

### 3.2 Variability and uncertainty

We consider different representations of data with respect to two major short-term aspects: the aggregation over time and the representation of uncertainty.

#### 3.2.1 Aggregation over time

We consider two approaches to aggregation of data over time: Representative hours and representative days.

- *Representative hours* Aggregation by hours means that hours are evaluated separately with respect to their state values, e.g. demand and wind production. Hours are clustered into a number of groups, each representing a number of "similar" hours in a year. The index of a time period  $t$  therefore refers to a group. The duration of a group is given by  $\tau_t$ , indicating the number of hours represented. Due to the loss of chronology, ramping constraints cannot be considered, and hence, the constraints (1d) and (1i) are omitted from the model.
- *Representative days* Aggregation by days means that hours are evaluated while respecting the order of their state values throughout a day. Days are likewise clustered into a given number of groups, each representing a number of "similar" days in a year. A representative day has an associated weight, referring to the number of days represented by the group and given by  $\tau_t$ . This weight applies to time periods and is the same for all time periods  $t$  of the same representative day. The index of the hourly time periods  $t$  runs from 1 to  $24 * N$ , where  $N$  is the number of representative days. Moreover, the set  $T_d$  contains all hours except the last of each day, i.e.  $T_d$  includes indices that are not multiples of 24. With preservation of chronology within a day, ramping constraints (1d) and (1i) are included in the model, although not between days.

Alternative choices of aggregation over time, such as a combination of representative hours and representative days or time periods of more or less than an hour, could be considered, see for example Slednev et al. (2017). We briefly discuss the combination of representative hours and representative days in Sect. 3.2.4.

### 3.2.2 Representation of uncertainty

To evaluate the importance of including short-term uncertainty, two different approaches are considered: one with short-term uncertainty and another without short-term uncertainty.

- *With short-term uncertainty* Section 3.1 describes the two-stage day-ahead and balancing market clearing in the presence of short-term uncertainty. In the following, we refer to the model with stochastic market clearing. The distribution of the state values, e.g. wind power production, is discretised, using a limited number of scenarios ( $|S| > 1$ ) with corresponding probabilities. Depending on the representation of data over time, scenarios either consist of hourly values of production or of daily production schedules.
- *Without short-term uncertainty* In the absence of short-term uncertainty, the balancing market serves no purpose and the day-ahead market clearing is sufficient. We refer to this as conventional market clearing. The distribution is replaced by the expected hourly wind power production ( $|S| = 1$ ).

### 3.2.3 Overview

We consider four combinations of data aggregation over time and representation of uncertainty and the resulting four models for the generation expansion problem: daily representation and conventional market clearing (DC), daily representation and stochastic market clearing (DS), hourly representation and conventional market clearing (HC), and hourly representation and stochastic market clearing (HS). These four models are found in Table 1, where an acronym indicates the model and the number of days or hours included.



**Table 1** Overview of models for generation expansion planning, categorised by data aggregation over time and representation of uncertainty

S–T uncert.	Data agg.	Acronym	Optimisation problem		
Without	Rep. days	DC-(# of days)	min	(1a)	$ S  = 1$
			s.t.	(1b)–(1o)	
With	Rep. days	DS-(# of days)	min	(1a)	$ S  > 1$
			s.t.	(1b)–(1o)	
Without	Rep. hours	HC-(# of hours)	min	(1a)	$ S  = 1$
			s.t.	(1b), (1c), (1g)–(1h), (1j)–(1o)	
With	Rep. hours	HS-(# of hours)	min	(1a)	$ S  > 1$
			s.t.	(1b), (1c), (1g)–(1h), (1j)–(1o)	

### 3.2.4 Limitations of the methodology

In our analysis, we use the most simple model that includes both short-term variability and uncertainty. Focus is on whether variability or uncertainty is more important and in which situations. Our simplifications, however, do introduce limitations to the scope of the paper.

A main simplification is to represent flexibility using ramp rates only and not include the typical features of a unit commitment problem such as start-up costs, minimum up- and down-time constraints etc. However, we expect that the effects of short-term variability and uncertainty will be more pronounced with less flexibility in the power system, and thus, in the presence of unit commitment constraints.

Furthermore, we confine ourselves to the temporal dimension and do not consider the spacial dimension of a power network. The representation of the network could provide both flexibility and restrictions to the optimisation model, and thus, both reduce and amplify the effects of variability and uncertainty. When clustering days or hours, for example, the effect of the peak flow on the network should ideally be considered (Schwarz et al. 2018).

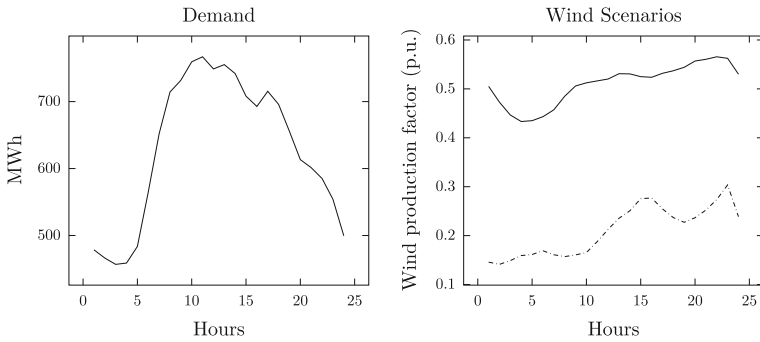
For simplicity, we use either representative days or representative hours and not a combination of days and hours. A hybrid approach is proposed by Slednev et al. (2017) who report promising computational results. The number and selection of representative days and hours, however, are critical to the performance.

Finally, we consider a greenfield system to highlight the differences in the expansion plans resulting from the four models of variability and uncertainty. Such differences would be diluted if existing capacities were considered. In other words, if considering a brownfield system, the differences between the four models would be much less.

## 4 Illustrative example

We start by illustrating the effects of short-term uncertainty and variability on a stylised example.

Demand and wind power data is from the pricing zone DK1 (Nordpool 2017). We assume deterministic demand, using a single representative day, and stochastic wind production, characterised by two scenarios with the same probability. These two scenarios correspond to the wind capacity factor in DK1 for two representative days of 2017 and were determined by scenario reduction techniques, see Sect. 5.1.1. Demand and wind production profiles are



**Fig. 1** Demand and wind production profiles of a representative day and two wind power scenarios, respectively

**Table 2** Generation unit data for an illustrative example . *Source:* Ea Energianalyse (2014) and Schröder et al. (2013)

$g$	Wind	Coal	GasDA	GasBal	Nuclear
$c_g^I$ (T€/MWh)	124	106	51	51	150
$c_g$ (€/MWh)	0	31.4	63.1	63.1	15.4
$c_g^+, c_g^-$ (€/MWh)	0	5.23	500	4.51	500
$r_g^D, r_g^U$ (p.u.)	1	0.3	0.7	0	0

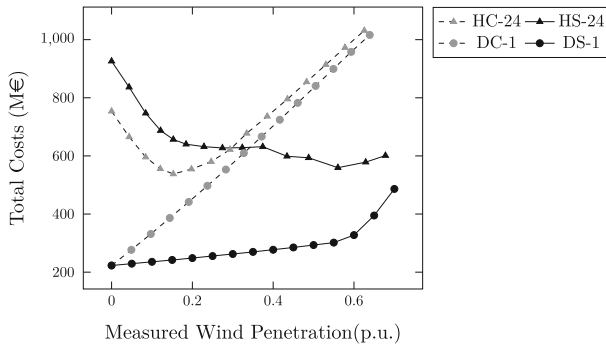
shown in Fig. 1. We consider the models DS-1, including ramping and stochastic market clearing, DC-1 with ramping only, HS-24 with stochastic market clearing only, and HC-24 excluding both ramping and stochastic market clearing. For the notation, see Table 1.

We consider the following generation units named according to the technology with most similar characteristics: wind turbines, nuclear, gas and coal. To illustrate the differences resulting from choice of modelling, we divide the flexible gas units into two different types: a gas unit that is flexible in the day-ahead market (with ramping ability, but high balancing costs) and a gas unit that is flexible in the balancing market (with low balancing costs, but no ramping ability). In reality, as in the case study of Sect. 5, however, most gas units are flexible in both markets. The nuclear unit is assumed inflexible in both the day-ahead and in the balancing market. The data for these units is shown in Table 2. Furthermore, load shedding and wind curtailment costs are set to  $v^L = v^S = 500$  €/MWh.

To evaluate and compare investment decisions across the four models, the following procedure is used:

1. Solve each of the problems DC-1, DS-1, HC-24 and HS-24, see Table 1. for their definition.
2. For each of the optimal solutions to DC-1, HC-24 and HS-24, fix the investment decision and solve the generation expansion problem DS-1 [without minimum wind penetration constraints (1m)]. Record the objective function value.

Since the DS-1 model includes both short-term variability and uncertainty, we use this as the baseline for evaluation and comparison. Thus, by definition this model provides the optimal investment decisions and the minimal costs. We evaluate the objective function value of using the (feasible, but not necessarily optimal) investment capacities from DC-1, HC-24 and HS-24, which are at least as high as those of DS-1. The difference in objective function values can be interpreted as the costs of disregarding variability and/or uncertainty in the optimisation. Figure 2 show the objective function values of the DS-1 model plotted as functions of the



**Fig. 2** Total costs from fixing the optimal investment capacities from the problems DC-1, DS-1, HC-24 and HS-24 in DS-1. As functions of the measured wind penetration

**Table 3** Investment decisions in MW for  $\kappa = 0.2$  and  $\kappa = 0.6$

Model	$\kappa$	Wind	Nuclear	GasDA	GasBal	Coal	LS	WC
DC-1	0.2	352	355	288	0	0	0.08	0.05
DS-1	0.2	352	300	279	103	24	0	0
HC-24	0.2	352	577	66	0	0	0.07	0.24
HS-24	0.2	352	528	0	173	0	0.12	0.24
DC-1	0.6	1060	134	266	0	0	0.2	0.09
DS-1	0.6	1076	0	208	288	86	0	0.02
HC-24	0.6	1056	326	76	0	0	0.19	0.12
HS-24	0.6	1071	162	48	372	0	0.06	0.08

LS is load shedding as a ratio of demand and WC is wind curtailment as a ratio of realised wind production

measured wind penetration. Note that whereas the required wind penetration is exogenous, the measured wind penetration is a function of capacity, and hence, endogenous. For  $\kappa = 0.2$  and  $\kappa = 0.6$ , Table 3 shows optimal investment capacities for each of the problems DS-1, DC-1, HS-24 and HC-24.

The results show that since the HC-24 model disregards both variability and uncertainty, investments are mainly in the inflexible base nuclear generation, meaning that load shedding costs are high, cf. Table 3 and Fig. 2. The minor investment in day-ahead flexible gas serves to cover peak load hours and is almost the same for all wind penetrations. With higher wind penetration, the major change in investments is substitution of wind for nuclear. Total costs decrease with wind penetrations up to  $\kappa = 0.2$  since wind power provides some flexibility through curtailment. For wind penetrations from  $\kappa = 0.2$  and up, total costs increase, as the cost savings of wind power are out-weighted by the costs insufficient balancing capacity and the resulting load shedding.

As with the HC-24 model, the HS-24 model invests in the nuclear unit to cover base load. Moreover, when accounting for uncertainty, the model also invests in the gas unit flexible in the balancing market to provide peak load capacity in some hours. The choice of gas unit, however, means that the total costs of the HS-24 model are higher than those of the HC-24 model for low wind penetrations. As wind penetration increases, total costs of the HS-24 model first decreases and then stabilises. The reason for decreasing costs is that wind power provides cost savings through the flexibility to curtail, but also that the installed balancing

capacity handles the uncertainty with minimal load shedding. As with the HC-24 model, minor installations in day-ahead flexible gas serve the peak load.

The DC-1 model includes variability and thus invests in day-ahead flexible gas in addition to nuclear. The deterministic model, however, disregards balancing and therefore investment capacities neither include the gas units flexible in this market nor coal units. This leads to expensive wind curtailment and/or load shedding as wind penetration increases, and thus, increasing total costs.

By definition, the DS-1 model provides the lowest costs for all wind penetration levels. By accounting for both variability and uncertainty, this model avoids significant wind curtailment and load shedding costs. The higher wind penetrations, the higher the total costs. The reason is that higher requirements of wind penetration leads to higher investment costs of wind investments and, for very high wind penetration, wind curtailment costs in some hours and scenarios. For low penetrations, the DS-1 model produces a combination of all generation technologies to serve flexibility needs both in terms of ramping and balancing. For wind penetrations of  $\kappa = 0.6$  and up, however, nuclear is substituted by the other technologies.

We conclude this example by noting that our model clearly captures the impact of the two short-term effects: uncertainty and variability. The results in Fig. 2 show that representative days are very valuable for incorporating the short-term variability, although cost savings are less for high wind penetration levels. In contrast, for wind penetrations above a certain level, the inclusion of the stochastic market clearing provides significant cost savings.

## 5 Case study

We continue by applying the modelling framework introduced in Sect. 3 to data from the Nordpool bidding region DK1 covering the Western part of Denmark (Nordpool 2017). The wind penetration target is set to 30%, as is the Danish 2020 renewable energy target (The Danish Government 2013).

### 5.1 Data

We use historical market data from 2014. The data includes aggregated demand, wind power forecasts and realised wind power production for the entire region. The data is available on an hourly basis and is normalised by total capacity.

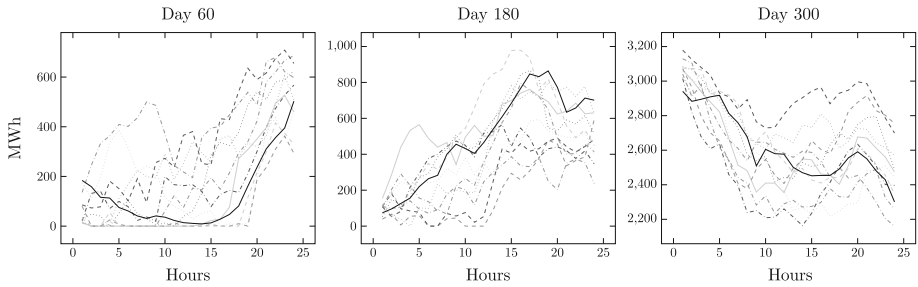
With data available for both forecasted and realised wind power production, we model the hourly forecast error:

$$\tilde{\rho}_t = \rho_t + e_t, \quad (2)$$

where  $\tilde{\rho}_t$  is the realised wind production,  $\rho_t$  is the wind production forecast and  $e_t$  is the forecast error, all given as capacity factors. Recall that only the forecast is known when the day-ahead market clears, whereas the forecast error is realised at the time of clearing the balancing market.

For simplicity, we fit the wind forecast errors to an ARMA time series model, assuming that the process is stationary with decaying autocorrelations. More detailed approaches to modelling wind forecasting errors are given by Bludszweit et al. (2008) and Box et al. (1994). By inspection of the autocorrelation functions, we choose an AR(2) model on the following form:

$$e_t = \phi_1 e_{t-1} + \phi_2 e_{t-2} + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma^2). \quad (3)$$



**Fig. 3** Scenarios of wind production (dashed lines) and historical wind production (solid line) for 3 selected days

**Table 4** Generation unit data for the case study. *Source:* Ea Energianalyse (2014) and Schröder et al. (2013)

$g$	Wind	Coal	Gas	Nuclear
$c_g^I$ (T€/MWh)	124	106	51	150
$c_g$ (€/MWh)	0	31.4	63.1	15.4
$c_g^+, c_g^-$ (€/MWh)	0	5.23	4.51	25.67
$r_g^D, r_g^U$ (p.u.)	1	0.3	0.7	0.03

here the error terms  $\epsilon_t$  capture variations in the historical data that are not explained by previous observations, and are assumed independent and identically normally distributed around zero. Fitting this model to the forecast errors from DK1 in 2014, we obtain the estimates  $\phi_1 = 1.186$  and  $\phi_2 = -0.294$  and the z-test statistics indicate that the coefficients are statistically significant ( $\Pr(> |z|) < 2.2 \times 10^{-16}$  for both coefficients). The assumption of normally distributed residuals is confirmed to a satisfying extent by histograms and QQplots.

The time series model is used to generate scenarios for realised wind power production for each day of the target year. The model takes as input the observed forecast errors of the last 2 h from the previous day. For the following 24 h, we sample the error term and recursively compute the forecast errors. We generate 1000 scenarios of wind forecast errors and reduce these to 10 by the scenario reduction technique of Dupačová et al. (2003). The aim is to accurately represent uncertainty while the model remains computationally tractable (Morales et al. 2009). Using (2), these scenarios are translated into realised production. The result is 10 24-dimensional scenario vectors,  $(\tilde{\rho}_{1s}, \dots, \tilde{\rho}_{24s}), s = 1, \dots, 10$  of realised wind power production for each of the 365 days of the year. In Fig. 3 we plot the scenarios and the observed historical wind power production of three selected days.

The data for conventional generation taken from Ea Energianalyse (2014) and Schröder et al. (2013), and chosen to represent a diverse selection of production units. All costs taken from Ea Energianalyse (2014) are 2020 predictions and investment costs are annualised with expected lifetime of the technology and using a discount rate of 4%. The expected lifetime is defined as the minimum of the technical and the economical lifetime of a unit, where the technical lifetime is taken from Danish Energy Agency (2012) and the economical lifetime captures the number of years operation is profitable, taking future discounted fuel and CO2 prices into account. The four production units are: wind power (wind), coal-fired pulverised fuel power plant (coal), combined-cycle gas turbine (gas) and nuclear. Table 4 provides the data.

Ramp rates of the *gas* and *coal* units are not publicly available. We, therefore, derive the ramp rates from the aggregated hourly output for each technology, collected from Entsoe

(2016) and with outliers removed. More specifically, we use the maximum hourly change in aggregate output to approximate the aggregate ramp rate. The ramp rate is finally normalised by the maximum hourly output. For simplicity and supported by the data, upward and downward ramp rates are the same.

Balancing costs are modelled as follows. We assume that the balancing costs are increasing in production costs and decreasing in ramp rates and consider the following relation for  $c_g^+$  and  $c_g^-$ :

$$c_g^+ = M \cdot \frac{c_g}{r_g^u} \quad \text{and} \quad c_g^- = M \cdot \frac{c_g}{r_g^d}, \quad (4)$$

where  $M$  is an adjustment factor. This  $M$  is estimated from historical balancing prices, cf. Nordpool (2017). The average balancing price in DK1 for 2014 is 6.30 €/MWh, and thus, we set  $M = 0.05$  to achieve the balancing costs in Table 4. We further consider the case of a zero balancing cost for all units and present both cases in the results.

The load shedding costs are estimated by the maximum price of electricity. From Nordpool (2017) we note that the maximum price in DK1 is 3000 €/MWh, and thus, we set  $v^L = 3000$  €/MWh. Similarly, we estimate the wind curtailment costs by the minimum price. The minimum price of electricity in DK1 is  $-500$  €/MWh, and the wind curtailment costs are therefore set to  $v^S = 500$  €/MWh.

### 5.1.1 Data aggregation

The technical literature includes several methods to select representative days or hours. In Hastie et al. (2009) and ElNozahy et al. (2013), representative days or weeks are chosen using classical clustering techniques such as K-means or hierarchical clustering. New methods to select representative days have recently been proposed. For instance, Poncelet et al. (2017) provide a novel optimisation-based approach to select representative periods. Similarly, Liu et al. (2017) propose a modified hierarchical clustering procedure to choose a reduced set of representative days that retains important statistical features of the input data such as correlation.

Our data aggregation is carried out using the GAMS/SCENRED tool (Römisch 2002). Although this tool is intended for scenario reduction, the clustering algorithm may likewise apply for the reduction of hours or days to a smaller subset, with each day or hour of a year being equally probable. The GAMS/SCENRED tool is an out-of-the-box tool and the reduction selects a specific hour or day as representative. When clustering by hour, we consider all 8760 h of historical wind production and demand data and reduce to the required number of representative hours, as indicated by the suffix of the model name, e.g. HC-24. When clustering by day, we likewise use all 365 days of historical data and reduce to the required number of days, likewise revealed by the suffix of the model name, e.g. DC-1.

## 5.2 Results

We consider the four combinations of data aggregation over time and representation of uncertainty and the resulting models for the generation expansion problem, cf. Sect. 3.2. The results are divided into two sections: First, we analyse these models using the full data set (we refer to the models HC-8760, HS-8760, DC-365 and DS-365 as full models). Secondly, we include only a subset of the data obtained via aggregation and benchmark against the full DS-365 model, using the procedure of Sect. 4. The full results are included in Appendix A.

**Table 5** Optimal investment decisions and model runtimes for the different full models

Model	Wind	Coal	Gas	Nuclear	Runtime (s)	TC	IC	OC	LSC	WCC
(a) $c_g^+ = 0.05 \frac{c_g}{r_g^u}$ and $c_g^- = 0.05 \frac{c_g}{r_g^d}$ .										
DS-365	2562	983	796	1354	30,859	990	665	316	3	7
HS-8760	2559	339	939	1841	1829	1020	676	322	4	18
DC-365	2561	935	747	1417	217	990	666	312	4	8
HC-8760	2559	241	902	1929	32	1035	677	326	7	25
(b) $c_g^+ = c_g^- = 0$										
DS-365	2561	950	777	1406	24,512	981	668	304	3	6
HS-8760	2559	244	937	1937	1162	1019	680	315	3	20
DC-365	2561	935	747	1417	88	981	666	303	4	7
HC-8760	2559	241	902	1929	32	1022	677	316	7	23

TC total costs, IC investment costs, OC operating costs, LSC load shedding costs, WCC wind curtailment costs, all in M€, from evaluating the investment decisions in DS-365

### 5.2.1 Technical details

Our model is implemented using GAMS 24.7.4 and solved using CPLEX 12.6.3.0 on a HP ProLiant server with 4 AMD 2.50 GHz CPUs, with a total of 64 cores and 256 GB RAM. The reported runtimes are as measured by GAMS (Rosenthal 2014).

### 5.2.2 Results from the full models

The optimal investment decisions and resulting costs of the four models are provided in Table 5 with respectively non-zero and zero balancing costs.

Regarding the investment decisions, we note that all models in Table 5 install approximately the same wind capacity (around 2560 MW) due to the minimum wind penetration constraint. The small differences in wind investments is due to load shedding and wind curtailment.

When comparing representative days and hours in Table 5 the main difference is in nuclear investment capacities. Representative hours results in approximately 35% larger nuclear capacities; 1929 MW versus 1417 MW in the deterministic models (HC-8760 and DC-365) and 1841 MW versus 1354 MW in the stochastic models (HS-8760 and DS-365) including balancing costs. The reason is that ramping needs are ignored and nuclear is inexpensive baseload. For representative days accounting for ramping, nuclear is replaced by coal, the capacity of which is 2–3 times larger than for representative hours. Somewhat surprisingly gas investment capacity is around 20% less for representative days than for representative hours; 747 MW versus 902 MW in the deterministic models and 796 versus 939 MW in the stochastic models. This can be explained by the larger installation of coal that to some extent covers the need for flexibility.

Note that the optimal investment decisions from the deterministic models are the same with non-zero or zero balancing costs while they differ for the stochastic models, e.g. coal investment in HS-8760 is 339 MW with balancing costs and 244 MW without balancing costs. The reason is that in the deterministic models it is never optimal to use the balancing

**Table 6** The number of variables and constraints and the runtimes for the four different models with non-zero balancing costs

Model	# of variables	# of constraints	Runtime (s)
DS-365	$13 \times 10^5$	$12 \times 10^5$	30,859
HS-8760	$13 \times 10^5$	$7 \times 10^5$	1829
DC-365	$2 \times 10^5$	$2 \times 10^5$	217
HC-8760	$2 \times 10^5$	$1 \times 10^5$	32

market. Nevertheless, balancing costs do influence the costs of the investment decisions when evaluated in the DS-365 model.

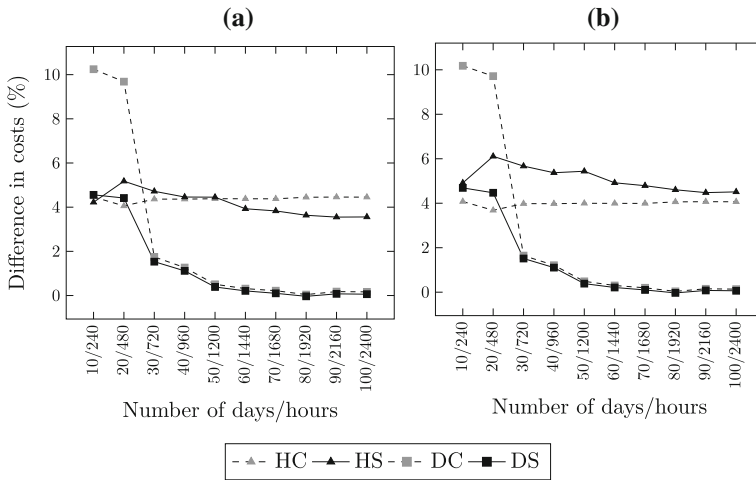
The differences between the deterministic and stochastic models are less than 5%, except when comparing the coal investments for representative hours (HC-8760 and HS-8760) with non-zero balancing costs. These small differences in the case study are in contrast to the example in Sect. 4, for which we observed significant differences between the investments from the stochastic and the deterministic models. Since the example in Sect. 4 is a stylised, illustrative example, this is not surprising. The units of the example are either flexible with respect to ramping and balancing, whereas this is rarely the case in reality. In contrast, some units of the case study such as coal and gas are flexible with respect to both ramping and balancing, with relatively high ramp rates and low balancing costs. Thus, the flexibility needs in a stochastic market clearing are already partly covered by the flexible units installed to cope with variability of demand and wind power production. In fact, the only significant difference between the deterministic and stochastic models is for the representative hours in Table 5a. Here, the HS-8760 model results in 40% higher investment capacities in the flexible coal than the HC-8760 model, that is, 339 MW versus 241 MW. The same does not apply for the results in Table 5b since the assumption of zero balancing costs produces the less realistic conclusion that nuclear is the best option for balancing. To summarise, the inclusion of stochastic market clearing improves the results for representative hours when balancing costs are non-zero but not much for representative days.

The same tendency is observed for the costs, for which the main differences are between representative days and representative hours and not between the deterministic and stochastic models. The higher nuclear capacities lead to higher investment costs for representative hours than representative days. The higher nuclear capacities, however, also generate higher realised operating costs as gas satisfy the flexibility needs for representative hours whereas coal serves this purpose for representative days. Moreover, wind curtailment costs are very different for representative days and hours because the large inflexible nuclear capacities act as baseload and peaks in wind production must be curtailed. Finally, since the objective functions are based on expected costs, security of supply is only accounted for through expected load shedding penalties. Lower penalties reveal that the stochastic models have slightly higher reliability rates.

The number of variables and constraints in the different models is specified in Table 6. When comparing the deterministic models to the stochastic models with 10 scenarios, we observe that the runtimes increase with a factor between 60 and 200. The number of variables and constraints increase with a factor slightly less than 10, indicating that the computational burden increases more than linearly in the number of variables and constraints. When comparing the representative days and hours, the increase in runtimes is with a factor between 3 and 11 and is due to the additional constraints for representative days.

We conclude this section by comparing the optimal investment decisions, the total costs and the model runtimes of the four full models. The trade-off between runtime and total costs





**Fig. 4** Total costs differences between the models with aggregated data and the full model. The x-axis refers to the number of representative days and hours. **a**  $c_g^+ = 0.05 \frac{c_g}{r_g}$  and  $c_g^- = 0.05 \frac{c_g}{r_g}$ , **b**  $c_g^+ = c_g^- = 0$

clearly points at DC-365 as the preferred model. Representative hours do not perform as well as representative days and DS-365 does not perform significantly better than DC-365, even with a runtime significantly larger. Hence, when faced with the choice between modelling variability or uncertainty, the inclusion of dynamics is preferable.

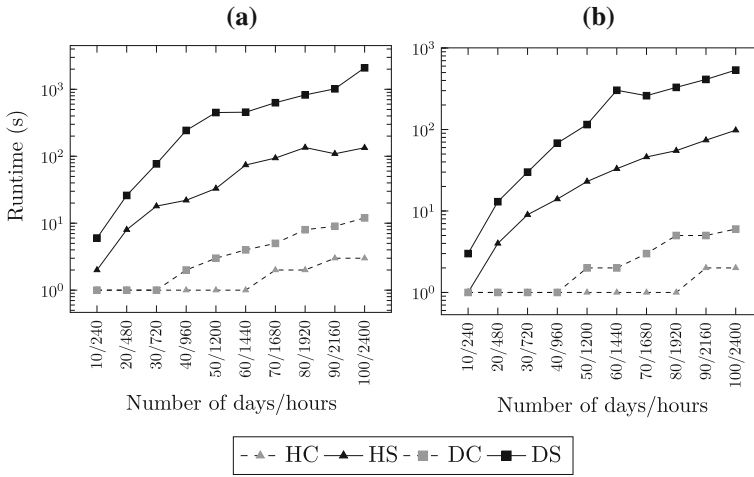
### 5.2.3 Results from the models with aggregated data

We evaluate the performance of the different models for an increasing number of days and hours. For example, we solve each of the problems DC-10, DS-10, HC-240 and HS-240. For each of the optimal solutions to DC-10, DS-10, HC-240 and HS-240, we fix the investment decision and solve the full DS-365. We do the same for higher numbers of days and hours. Figure 4 shows the differences between the resulting objective function values and the minimal costs of DS-365 in percentage.

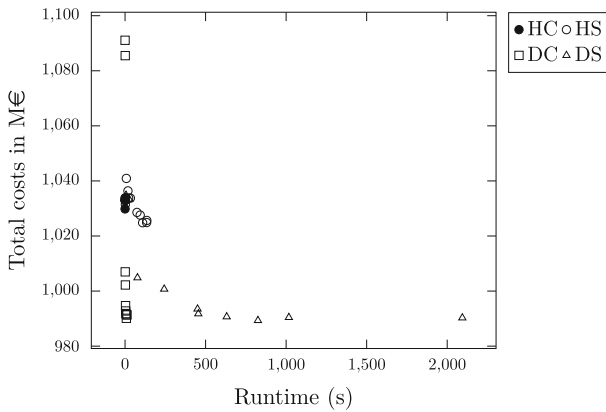
In both Fig. 4a, b, we observe that with representative days the cost differences approach zero when the number of days increases. As expected, the inclusion of more days results in investment decisions closer to those of the full model. The same is not observed with representative hours. The cost differences do not improve for an increasing number of hours, and thus, even for the highest number of hours included, 2400, the level of detail in representative hours is insufficient. When comparing the representative days and hours, the former outperform the latter when including 30 days or more. In fact, we confirm that the effect of taking short-term variability into account is crucial, even for a limited number of days.

When comparing the deterministic and stochastic models, we note that for 30 days or more, the models produce very similar cost differences from the full model. When including 30 representative days, the costs difference is already less than 2%, indicating that 30 representative days offset the effects of uncertainty in this specific case study. We, therefore, stipulate that you can ignore uncertainty by adding a sufficient number of representative days, which is computationally much less expensive than doing stochastic optimisation.

Analysing the differences between Fig. 4a, b, the main difference is that the stochastic model with representative hours performs better in Fig. 4a than in Fig. 4b. This is because the



**Fig. 5** Runtimes as a function of number of days/hours for all models. Note the y-axis is logarithmic and that runtimes have a lower bound of 1 second. **a**  $c_g^+ = 0.05 \frac{c_g^-}{r_g^u}$  and  $c_g^- = 0.05 \frac{c_g^-}{r_g^d}$ , **b**  $c_g^+ = c_g^- = 0$



**Fig. 6** Runtime versus total costs for all models (except DS-365)

non-zero balancing costs in Fig. 4a incentivise investment in more flexible units which in turn then reduce the difference in costs from the full model. Observe also that the the stochastic model with representative hours performs slightly worse than the deterministic counterpart in Fig. 4b which seems counter-intuitive. In the stochastic model with zero balancing costs, the inflexible nuclear unit can be used as a balancing unit at zero costs because there are no ramping constraints. Thus, the stochastic model invests more in nuclear than the deterministic, which is more costly when evaluated in the full model.

Figure 5 shows the runtimes of each model plotted against the number of hours or days. Note that the y-axis is logarithmic. The results are very similar with zero and non-zero balancing cost. In both cases, the stochastic models are by far the most computationally heavy. The reason is that the stochastic models are larger by an order of magnitude of 50-150 with representative days and 25–50 with representative hours.

To illustrate the trade-off between the quality of the investment decisions and the computational effort, Fig. 6 plots the runtimes against the total costs for all models and all days/hours.

With hourly representation, the points are all close, with small relative differences in both runtime and total costs. The stochastic models, however, always have lower total costs and higher runtime than the deterministic. With daily representation, all models have relative low runtime, whereas the best deterministic models also have relatively low total costs. The stochastic models have the lowest total costs but only for models with a very high runtime.

To summarise the findings of the case study, the DC-30 model yields investment decisions with less than a 2% difference in total costs to the full DS-365 model. Furthermore, the computational burden of the DC-30 model is far less than DS-365, with runtimes under 1 second for the DC-30 model and over 30,000 s for the DS-365 model when considering non-zero balancing costs.

## 6 Conclusion

With higher shares of renewable energy sources in many power systems, it is increasingly important to account for short-term variability and uncertainty in long-term power planning. Nevertheless, this often requires a level of techno-economical detail in modelling that significantly affects computational tractability. In this paper, we compare different approaches to represent variability and uncertainty in a model, while reducing runtime. We use an example to illustrate the effects of variability and uncertainty, whereas a Danish case study provides more realistic results.

Our example shows that accounting for short-term variability through ramping constraints and/or uncertainty via balancing costs has significant impact on the quality of investment decisions. In our more realistic case study, however, the inclusion of representative days and ramping constraints has the most significant effect, both regarding the quality of the solution and the computational burden of solving the model. In particular, we observe that the impact of short-term uncertainty is less important as the number of representative days increase.

Our model can be extended in various directions. For computational reasons, we capture inter-temporal restrictions through ramping constraints only. Our results may therefore underestimate the importance of including short-term techno-economical details in a long-term power planning problem. At the expense of longer runtimes, however, the model can be extended to account for unit commitment. Our model can likewise be extended to include network and transmission expansion. Network expansion may provide further system flexibility, whereas transmission constraints may impose restrictions on flexibility in generation. This trade-off may be subject of future research. Moreover, the market structure with perfect competition could be further investigated, from the perspective of both investors and policy makers. Allowing for market power, the model may become a mathematical programming problem with equilibrium constraints, for which computational tractability is of an even higher concern.

**Acknowledgements** T. K. Boomsma gratefully acknowledges support from the project Analyses of Hourly Electricity Demand (AHEAD) funded by ForskEl 2017.

## A: Results tables

All results listed in two Tables: one for non-zero balancing costs and one for balancing costs equal to zero.

See Tables 7 and 8.

**Table 7** Investment decisions and runtimes for the different models in the case of  $c_g^+ = 0.05c_g/r_g^u$  and  $c_g^- = 0.05c_g/r_g^d$

Model	Wind	Coal	Gas	Nuclear	Runtime	TC	IC	OC	LSC	WCC	Meas. wind pen.
DC-365	2561	935	747	1417	217	990.22	666.45	311.59	4.41	7.76	0.3
DS-365	2562	983	796	1354	30,859	989.68	664.57	315.65	2.66	6.8	0.3
HC-240	2575	259	963	1971	0	1033.85	690.07	322.76	1.26	19.76	0.3
HC-480	2550	276	979	1978	0	1029.83	690.81	321.24	0.72	17.05	0.298
HC-720	2557	262	977	1994	0	1032.87	692.44	321.73	0.76	17.95	0.298
HC-960	2564	270	972	1990	1	1032.87	693.34	320.79	0.75	17.98	0.299
HC-1200	2567	271	965	1989	1	1033.06	693.23	320.63	0.83	18.38	0.299
HC-1440	2566	271	965	1989	1	1033.03	693.19	320.64	0.83	18.37	0.299
HC-1680	2566	271	965	1989	2	1033.01	693.17	320.65	0.83	18.37	0.299
HC-1920	2566	266	965	1994	2	1033.71	693.41	320.88	0.83	18.58	0.299
HC-2160	2566	266	965	1995	3	1033.8	693.45	320.91	0.83	18.61	0.299
HC-2400	2567	267	964	1995	3	1033.76	693.52	320.8	0.83	18.61	0.299
HC-8760	2559	241	902	1929	32	1035.44	676.94	326.48	6.94	25.08	0.298
HS-240	2693	383	976	1796	2	1031.47	692.38	317.77	2.01	19.31	0.314
HS-480	2721	344	955	1856	8	1040.91	699.57	316.61	2.07	22.65	0.317
HS-720	2695	346	967	1842	18	1036.38	695.16	318.14	2.05	21.03	0.314
HS-960	2670	335	970	1844	22	1033.81	691.27	319.92	2.25	20.37	0.311
HS-1200	2656	328	944	1861	33	1033.75	690.09	319.59	2.88	21.19	0.31
HS-1440	2621	335	933	1857	74	1028.56	685.26	319.98	3.27	20.05	0.306
HS-1680	2613	338	919	1847	94	1027.59	682.42	320.39	4.37	20.41	0.305
HS-1920	2593	340	902	1846	135	1025.64	679.14	320.57	5.42	20.51	0.302
HS-2160	2594	333	935	1847	109	1024.82	680.31	321.6	3.72	19.19	0.302
HS-2400	2594	331	941	1845	134	1024.89	680.23	321.98	3.61	19.07	0.303
HS-8760	2559	339	939	1841	1829	1019.59	675.99	322.4	3.54	17.66	0.299

Table 7 continued

Model	Wind	Coal	Gas	Nuclear	Runtime	TC	IC	OC	LSC	WCC	Meas. wind pen.
DC-10	2653	843	420	1511	0	1091.08	665.52	300.38	101.18	23.99	0.309
DC-20	2652	925	447	1402	1	1085.5	659.12	306.6	100.91	18.87	0.31
DC-30	2685	871	732	1464	1	1006.98	681.19	306.31	6.78	12.71	0.314
DC-40	2652	870	779	1420	2	1002.21	672.61	311.89	6.7	11.01	0.31
DC-50	2590	876	790	1404	3	994.71	663.82	315.31	6.63	8.96	0.303
DC-60	2576	910	763	1396	4	992.77	663.2	314.54	6.61	8.41	0.301
DC-70	2566	892	766	1413	5	991.84	662.59	314.28	6.6	8.36	0.3
DC-80	2552	908	763	1400	8	990.11	660.54	315.11	6.59	7.87	0.299
DC-90	2562	904	746	1420	9	991.44	663.57	312.94	6.6	8.33	0.3
DC-100	2561	911	735	1424	12	991.32	664.27	312.12	6.6	8.34	0.3
DS-10	2654	904	537	1436	6	1034.79	666.77	306.22	45.8	16.01	0.31
DS-20	2652	944	551	1383	26	1033.37	663.53	309.48	45.74	14.62	0.31
DS-30	2686	926	780	1382	77	1004.84	677.28	311.25	5.06	11.25	0.314
DS-40	2653	933	816	1338	243	1000.73	669.19	316.73	5.08	9.72	0.31
DS-50	2591	939	830	1327	450	993.5	661.17	319.93	4.6	7.8	0.303
DS-60	2576	959	808	1329	455	991.76	660.66	319.08	4.62	7.4	0.302
DS-70	2566	945	809	1341	631	990.67	659.89	318.86	4.63	7.29	0.3
DS-80	2552	961	794	1333	825	989.32	657.79	319.37	5.19	6.97	0.299
DS-90	2563	959	786	1344	1017	990.43	660.07	317.87	5.18	7.31	0.3
DS-100	2562	962	779	1347	2094	990.28	660.41	317.39	5.18	7.3	0.3

**Table 8** Investment decisions and runtimes for the different models in the case of zero balancing costs for all units

Model	Wind	Coal	Gas	Nuclear	Runtime	TC	IC	OC	LSC	WCC	Meas. wind pen.
DC-365	2561	935	747	1417	88	981.02	666.45	303.41	4.4	6.75	0.3
DS-365	2561	950	777	1406	24,512	980.7	667.92	303.91	2.67	6.2	0.3
HC-240	2575	259	963	1971	0	1020.68	690.07	311.68	1.04	17.9	0.3
HC-480	2550	276	979	1978	0	1016.82	690.81	310.27	0.56	15.18	0.298
HC-720	2557	262	977	1994	0	1019.75	692.44	310.68	0.57	16.06	0.298
HC-960	2564	270	972	1990	0	1019.73	693.34	309.71	0.57	16.11	0.299
HC-1200	2567	271	965	1989	1	1019.9	693.23	309.53	0.63	16.52	0.3
HC-1440	2566	271	965	1989	1	1019.88	693.19	309.54	0.63	16.51	0.3
HC-1680	2566	271	965	1989	1	1019.85	693.17	309.55	0.63	16.51	0.299
HC-1920	2566	266	965	1994	1	1020.53	693.41	309.77	0.63	16.71	0.3
HC-2160	2566	266	965	1995	2	1020.61	693.45	309.79	0.63	16.74	0.3
HC-2400	2567	267	964	1995	2	1020.57	693.52	309.68	0.63	16.74	0.3
HC-8760	2559	241	902	1929	32	1022.39	676.94	315.62	6.55	23.28	0.298
HS-240	2693	295	974	1886	1	1028.88	696.45	309.56	1.91	20.97	0.314
HS-480	2721	250	949	1956	4	1040.62	704.28	308.73	2.05	25.56	0.317
HS-720	2695	247	965	1943	9	1036.3	699.67	310.89	2	23.74	0.314
HS-960	2670	241	967	1941	14	1033.38	695.66	312.57	2.18	22.96	0.311
HS-1200	2656	232	943	1958	23	1034.04	694.39	312.49	2.82	24.34	0.309
HS-1440	2621	238	932	1955	33	1028.95	689.58	312.92	3.19	23.26	0.305
HS-1680	2613	244	917	1942	46	1027.67	686.66	313.01	4.28	23.72	0.304
HS-1920	2593	247	900	1941	55	1025.87	683.39	313.09	5.35	24.04	0.302
HS-2160	2594	242	932	1941	74	1024.61	684.61	314.24	3.63	22.14	0.302
HS-2400	2594	237	939	1941	98	1024.94	684.5	314.88	3.52	22.03	0.302
HS-8760	2559	244	937	1937	1162	1019.49	680.24	315.38	3.45	20.43	0.298

Table 8 continued

Model	Wind	Coal	Gas	Nuclear	Runtime	TC	IC	OC	LSC	WCC	Meas. wind pen.
DC-10	2653	843	420	1511	0	1080.53	665.52	290.84	100.97	23.21	0.31
DC-20	2652	925	447	1402	1	1075.97	659.12	298.05	100.62	18.18	0.31
DC-30	2685	871	732	1464	1	996.84	681.19	297.24	6.72	11.7	0.314
DC-40	2652	870	779	1420	1	992.59	672.61	303.31	6.68	9.99	0.31
DC-50	2590	876	790	1404	2	985.49	663.82	307.11	6.61	7.95	0.303
DC-60	2576	910	763	1396	2	983.66	663.2	306.44	6.6	7.42	0.302
DC-70	2566	892	766	1413	3	982.64	662.59	306.09	6.59	7.37	0.3
DC-80	2552	908	763	1400	5	981.07	660.54	307.09	6.57	6.87	0.299
DC-90	2562	904	746	1420	5	982.2	663.57	304.7	6.58	7.34	0.3
DC-100	2561	911	735	1424	6	982.05	664.27	303.86	6.58	7.34	0.3
DS-10	2653	868	505	1505	3	1026.69	671.44	292.4	45.59	17.26	0.31
DS-20	2652	918	531	1428	13	1024.54	666.52	297.57	45.49	14.97	0.31
DS-30	2685	903	749	1436	30	995.55	681.26	298.4	5.05	10.84	0.314
DS-40	2652	902	792	1394	68	991.55	672.9	304.28	5.08	9.29	0.31
DS-50	2590	910	809	1377	115	984.47	664.4	308.23	4.61	7.23	0.303
DS-60	2576	933	787	1376	303	982.77	663.85	307.5	4.62	6.81	0.302
DS-70	2566	918	788	1389	260	981.69	663.1	307.25	4.63	6.71	0.3
DS-80	2552	935	772	1381	329	980.39	661.09	307.73	5.18	6.39	0.299
DS-90	2562	931	764	1392	412	981.47	663.3	306.26	5.17	6.73	0.3
DS-100	2561	936	757	1395	536	981.32	663.66	305.77	5.17	6.72	0.3

## References

- Akbari, T., Rahimi-Kian, A., & Tavakoli Bina, M. (2012). Security-constrained transmission expansion planning: A stochastic multi-objective approach. *International Journal of Electrical Power and Energy Systems*, *43*(1), 444–453.
- Aneiros, G., Vilar, J., & Raña, P. (2016). Short-term forecast of daily curves of electricity demand and price. *International Journal of Electrical Power and Energy Systems*, *80*, 96–108.
- Babrowski, S., Jochem, P., & Fichtner, W. (2015). Electricity storage systems in the future German energy sector. *Computers and Operations Research*, *66*, 1–13.
- Baringo, L., & Conejo, A. J. (2011). Wind power investment within a market environment. *Applied Energy*, *88*(9), 3239–3247.
- Baringo, L., & Conejo, A. J. (2012). Transmission and wind power investment. *IEEE Transactions on Power Systems*, *27*(2), 885–893.
- Baringo, L., & Conejo, A. J. (2013). Correlated wind-power production and electric load scenarios for investment decisions. *Applied Energy*, *101*, 475–482.
- Bertsch, V., & Fichtner, W. (2015). A participatory multi-criteria approach for power generation and transmission planning. *Annals of Operations Research*, *245*, 177–207.
- Bludszuweit, H., Dominguez-Navarro, J. A., & Llombart, A. (2008). Statistical analysis of wind power forecast error. *IEEE Transactions on Power Systems*, *23*(3), 983–991.
- Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (1994). *Time series analysis: Forecasting and control* (3rd ed.). Hoboken: Wiley.
- Chaton, C., & Doucet, J. (2003). Uncertainty and investment in electricity generation with an application to the case of hydro-quebec. *Annals of Operations Research*, *120*, 59–80.
- Danish Energy Agency. (2012). Technology data for energy plants. *Energimtedk*, *1*, 212.
- Dupačová, J., Gröwe-Kuska, N., & Römisch, W. (2003). Scenario reduction in stochastic programming. *Mathematical Programming*, *95*(3), 493–511.
- Ea Energianalyse. (2014). Elproduktionsomkostninger. *Energistyrelsen*, *1*(1), 1–42.
- Ehrenmann, A., & Smeers, Y. (2011). Generation capacity expansion in a risky environment: A stochastic equilibrium analysis. *Operations Research*, *59*(6), 1332–1346.
- ElNozahy, M. S., Salama, M. M. A., & Seethapathy, R. (2013). A probabilistic load modelling approach using clustering algorithms. In *2013 IEEE power energy society general meeting* (pp. 1–5). IEEE <https://doi.org/10.1109/PESMG.2013.6672073>, <http://ieeexplore.ieee.org/document/6672073/>
- Entsoe. (2016). European network of transmission system operators for electricity.
- European Commission (2014) Communication from the commission to the european parliament, the council, the European economic and social committee and the committee of the regions—A policy framework for climate and energy in the period from 2020 to 2030. European Commission p Brussels.
- Gabriel, S. A., Conejo, A. J., Fuller, J. D., Hobbs, B. F., & Ruiz, C. (2010). *Complementarity modeling in energy markets* (Vol. 157). New York: Springer.
- Ghofrani, M., Arabali, A., Etezadi-Amoli, M., & Fadali, M. S. (2013). A framework for optimal placement of energy storage units within a power system with high wind penetration. *IEEE Transactions on Sustainable Energy*, *4*(2), 434–442.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). New York: Springer.
- Haydt, G., Leal, V., Pina, A., & Silva, C. A. (2011). The relevance of the energy resource dynamics in the mid/long-term energy planning models. *Renewable Energy*, *36*(11), 3068–3074.
- Hemmati, R., Hooshmand, R. A., & Khodabakhshian, A. (2014). Market based transmission expansion and reactive power planning with consideration of wind and load uncertainties. *Renewable and Sustainable Energy Reviews*, *29*, 1–10.
- Jabr, R. A., Džafić, I., & Pal, B. C. (2015). Robust optimization of storage investment on transmission networks. *IEEE Transactions on Power Systems*, *30*(1), 531–539.
- Jin, S., Ryan, S. M., Watson, J. P., & Woodruff, D. L. (2011). Modeling and solving a large-scale generation expansion planning problem under uncertainty. *Energy Systems*, *2*(3–4), 209–242.
- Liu, Y., Sioshansi, R., & Conejo, A. J. (2017). Hierarchical clustering to find representative operating periods for capacity-expansion modeling. *IEEE Transaction on Power Systems*, *99*, 1–8.
- Ludwig, S., Haller, M., Schmid, E., & Bauer, N. (2011). Fluctuating renewables in a long-term climate change mitigation strategy. *Energy*, *36*(11), 6674–6685.
- Ma, J., Silva, V., Belhomme, R., Kirschen, D. S., & Ochoa, L. F. (2013). Evaluating and planning flexibility in sustainable power systems. *IEEE Transactions on Sustainable Energy*, *4*(1), 200–209.
- Morales, J. M., Pineda, S., Conejo, A. J., & Carrión, M. (2009). Scenario reduction for futures market trading in electricity markets. *IEEE Transactions on Power Systems*, *24*(2), 878–888.



- Murphy, F. H., & Smeers, Y. (2005). Generation capacity expansion in imperfectly competitive restructured electricity markets. *Operations Research*, 53(4), 646–661.
- Nahmmacher, P., Schmid, E., Hirth, L., & Knopf, B. (2016). Carpe diem: A novel approach to select representative days for long-term power system modeling. *Energy*, 112, 430–442.
- Nordpool. (2017). Nordpool. [www.nordpoolspot.com/](http://www.nordpoolspot.com/).
- Orfanos, G. A., Georgilakis, P. S., & Hatzigiorgiou, N. D. (2013). Transmission expansion planning of systems with increasing wind power integration. *IEEE Transactions on Power Systems*, 28(2), 1355–1362.
- Pfenninger, S., Hawkes, A., & Keirstead, J. (2014). Energy systems modeling for twenty-first century energy challenges. *Renewable and Sustainable Energy Reviews*, 33, 74–86.
- Pina, A., Silva, C. A., & Ferrão, P. (2013). High-resolution modeling framework for planning electricity systems with high penetration of renewables. *Applied Energy*, 112, 215–223.
- Pineda, S., & Morales, J. M. (2016). Capacity expansion of stochastic power generation under two-stage electricity markets. *Computers and Operations Research*, 70, 101–114.
- Pineda, S., Morales, J. M., Ding, Y., & Stergaard, J. (2014). Impact of equipment failures and wind correlation on generation expansion planning. *Electric Power Systems Research*, 116, 451–458.
- Pineda, S., Morales, J. M., & Boomsma, T. K. (2016). Impact of forecast errors on expansion planning of power systems with a renewables target. *European Journal of Operational Research*, 248(3), 1113–1122.
- Poncelet, K., Delarue, E., Six, D., Duerinck, J., & D’haeseleer, W. (2016). Impact of the level of temporal and operational detail in energy-system planning models. *Applied Energy*, 162, 631–643.
- Poncelet, K., Hoeschle, H., Delarue, E., Virag, A., & D’haeseleer, W. (2017). Selecting representative days for capturing the implications of integrating intermittent renewables in generation expansion planning problems. *IEEE Transactions on Power Systems*, 32(3), 1936–1948. <https://doi.org/10.1109/TPWRS.2016.2596803>.
- Pozo, D., Sauma, E., & Contreras, J. (2013). A three-level static MILP model for generation and transmission expansion planning. *IEEE Transactions on Power Systems*, 28(1), 202–210.
- Römisch, W. (2002). Scenred2. GAMS Development Corporation
- Rosenthal, R. E. (2014). GAMS ÆT: A user’s guide. GAMS development corporation 1
- Schröder, A., Kunz, F., Meiss, J., Mendelevitch, R., & von Hirschhausen, C. (2013). Current and prospective costs of electricity generation until 2050 2013. Deutsches Institut für Wirtschaftsforschung
- Schwarz, H., Bertsch, V., & Fichtner, W. (2018). Two-stage stochastic, large-scale optimization of a decentralized energy system: A case study focusing on solar pv, heat pumps and storage in a residential quarter. *OR Spectrum*, 40(1), 265–310.
- Sisternes, F. J. D., & Webster, M. D. (2013). Optimal selection of sample weeks for approximating the net load in generation planning problems. *ESD Working Paper Series 1*
- Slednev, V., Bertsch, V., & Fichtner, W. (2017). A multi-objective time segmentation approach for power generation and transmission models. In K. F. Dörner, I. Ljubic, G. Pflug, & G. Tragler (Eds.), *Operations Research Proceedings 2015* (pp. 707–714). Cham: Springer.
- Stoft, S. (2002). Power system economics. *Journal of Energy Literature*, 8, 94–99.
- The Danish Government. (2013). The Danish climate policy plan, The Ministry of climate. *Energy and Building*, 1(August), 61.
- Wang, J., Shahidepour, M., Li, Z., & Botterud, A. (2009). Strategic generation capacity expansion planning with incomplete information. *IEEE Transactions on Power Systems*, 24(2), 1002–1010.
- Wogrin, S., Duenas, P., Delgadillo, A., & Reneses, J. (2014). A new approach to model load levels in electric power systems with high renewable penetration. *IEEE Transactions on Power Systems*, 29(5), 2210–2218.