

# Optimal Storage, Investment and Management under Uncertainty: It is Costly to Avoid Outages!

Joachim Geske<sup>a</sup> and Richard Green<sup>a\*</sup>

---

## ABSTRACT

We show how electricity storage is operated optimally when the load net of renewable output is uncertain. We estimate a diurnal Markov-process representation of how Germany's residual load changed from hour to hour and design a simple dynamic stochastic electricity system model with non-intermittent generation technologies and storage. We derive the optimal storage, generator output and capacity levels. If storage capacity replaces some generation capacity, the optimal storage strategy must balance arbitrage (between periods of high and low marginal cost) against precautionary storage to ensure energy is available throughout a long peak in net demand. Solving the model numerically under uncertainty (only the transition probabilities to future loads are known), we compare the results to perfect foresight findings. The latter over-estimate the cost-saving potential of energy storage by 27%, as storage can take up arbitrage opportunities that would not be chosen if there was a need for precautionary storage.

**Keywords:** Electricity Storage, Uncertainty, Stochastic Electricity System Model, Precautionary Storage

<https://doi.org/10.5547/01956574.41.2.jges>

## 1. INTRODUCTION

Electricity storage has the technical potential to increase the efficiency of electrical systems significantly—especially in the context of integrating intermittent renewable technologies. This is achieved by shifting energy from periods with a low valuation of electricity (due to low demand or high renewable generation) to periods of high valuation (high demand and low renewable generation). Thus, the share of electricity generated in expensive peak load power plants can be reduced and the share generated by cheaper medium-load power plants can be increased. Additional efficiency gains can come from adapting generation capacity to the post-storage load—with a higher base load and lower peak load share.

As capacities are adapted to high levels of electricity storage, the available generation capacity will sometimes fall below peak load level, because the storage is a substitute for peak load generation and cutting capacity reduces system costs. But since electricity storage technologies are limited both in the amount of power (MW) they can discharge and the amount of energy (MWh) that they can store, a prolonged demand peak coinciding with low renewable generation raises the risk of outages. The per-MWh costs of unanticipated load-shedding are far higher than the per-MWh gains

<sup>a</sup> Imperial College Business School, London, SW7 2AZ, United Kingdom.

\* Corresponding author. E-mail: [r.green@imperial.ac.uk](mailto:r.green@imperial.ac.uk).

from adjusting generation capacity and scheduling. Furthermore, electricity demand and renewable generation are uncertain.

Arbitrage from storage requires charging when prices are low and discharging when they are high. We show that it is optimal to limit discharging, or even to charge the store at times of high prices, if it is holding relatively low amounts of energy. This reduces the risk of a future power cut in the event of an extreme demand pattern, even if it forgoes the arbitrage profit and immediate cost saving that could be made in the short term. As an anonymous referee has pointed out, this could be seen as a strategy of patient arbitrage—holding back energy in the hope of the very high prices likely to accompany extreme demands—but we give it the name “precautionary storage”, since it is a response to the uncertainty of electricity demand that would not be seen in a deterministic model. We show the relevance of this precautionary storage in a stochastic model that integrates short-term operations and long-term investment decisions. We model explicit technologies with a step function for marginal cost, even though this makes derivative-based approaches to the solution of stochastic dynamic problems impossible.

While there are numerous stochastic modelling approaches in electricity systems (see overviews by Wallace and Fleten (2003), Kallrath et al. (2009), Möst and Keles (2010) or Rebennack and Kallrath (2017)), relatively few deal with storage. Sioshansi et al. (2009) and Teng et al. (2012) apply the price taker assumption thus abstracting from the endogenous character of precautionary storage in the energy system. A literature overview of the value of storage can be found in Zucker et al (2013). Cruise and Zachary (2015) and Durmaz (2016) explicitly deal with precautionary storage. However, the endogenous capacity component is not considered. Powell et al. (2012) do consider capacity expansion in a stochastic model, comparing solution techniques. They find that precautionary storage is optimal when using Approximate Dynamic Programming, but is not seen when using a scenario-based approach and taking advantage of perfect knowledge within each scenario.

We set up a stochastic welfare maximization model for the electricity system with storage and derive a jointly optimal strategy for storage, for the (perfectly coordinated) conventional generation outputs and for capacity decisions under residual load uncertainty. We solve the model numerically and analyse the optimal strategy and capacities. To keep things simpler, residual load is considered as the only stochastic driver.

This residual load exhibits diurnal, weekly and seasonal patterns which make the stochastic optimization time inhomogeneous. These problems are almost insolvable with numerical techniques, not to mention analytical ones. Commonly-used approaches to the solution focus on numerical solutions taking either the short- or the long-run perspective. As these exclusive approaches are not well suited to this analysis we abstract from seasonal, (e.g. Simonsen et al., 2004) time inhomogeneity and unit commitment modelling and apply a Markov Decision Process framework with a diurnal structure to a merit order stack. The Markovian approach and its efficient handling of stationarity opens a feasible way to interpret the long term as the “sum” of many short periods thereby neatly unifying short and long-term perspectives.

To be specific, the expected costs of generation with storage under capacity constraints to meet a stochastic residual load are approximated by an infinitely repeated “representative” 24-hour average cost minimization problem. This “representative” stage is scaled up to a year in length and the resulting total operating costs are combined with the annualised capacity cost. The model allows the expected system cost to be minimized under uncertainty, considering simultaneously the optimal generation investment and operating decisions for generation and storage.

Unfortunately, the optimization problem as stated here is non-convex. To determine the global optimum we propose the following algorithm: In the short-term stage, the storage strategy

and generation (by merit order dispatch of a set of conventional generation technologies) are optimised as a stationary Markov Decision Process given the generation capacities. The resulting expected load-duration curve is submitted to the long-term stage and capacities are incrementally updated according to a screening-curve approach.<sup>1</sup> These two steps are iterated until a fixed point is reached. A fixed point is a solution candidate for the global optimization problem; we compare it with any other fixed points and select the best. In this paper, we take the capacity of storage as fixed, but a further iterative process could find its optimal value, given costs and loads.

This algorithm is applied to solve two case studies. First, we specify the model to a simple case with only two generation technologies. We compare a strategy that takes every opportunity for arbitrage with the optimal strategy. In some states of the world, it is optimal to keep more energy in the store than under full arbitrage, in order to reduce the risk of losing load once the store is fully discharged. We show how the optimal strategy depends on the Value of Lost Load. If power cuts are cheap, it may not be worth missing out on arbitrage profits to reduce the risk of running out of energy, while if losing load is costly enough, it will be optimal to keep enough generating capacity to always meet demand in full. This in turn would allow storage to follow an arbitrage strategy with no risk of causing power cuts.

Our second case study represents Germany in 2011–15, with 300 GWh of storage capacity (6 hours average load), and five conventional generation technologies with empirical fixed and variable cost. The Markov process for the residual load is estimated from five years of hourly data. Under these more realistic conditions precautionary storage occurs as capacity is reduced below peak load. The strategies are qualitatively similar to those in the simple case but the quantitative extent could be determined more realistically. In comparison to perfect foresight analyses, uncertainty reduces the gain from storage by 27%, as energy is held back for precautionary reasons. This could be an overestimate of the difference, as some more information on future loads will be available due to weather forecasts, an issue we explore in follow-up work (Geske and Green, 2018); on the other hand, we ignore uncertainty over generator outages.

The article is structured as follows: First in section 2 the general setting of a stochastic dynamic electricity system model with fossil generation technologies and storage as a Markov Decision Process is introduced and the solution algorithm is proposed. We show in Section 3 that the residual load can be estimated as a Markov Process, using hourly load data for 2011–2015 in Germany. Based on the modelling environment of Section 2 and a simplified version of the estimated Markov process in Section 3, in Section 4 a two-technology storage and generation model is presented. This analysis is empirically refined in Section 5 with our calibration to Germany in 2011–2015. The model is solved numerically and the optimal storage strategy is presented. Sensitivities with respect to storage capacities and a comparison to perfect foresight deepen the understanding of the storage strategy and the impact of the uncertainty. In Section 6 we draw conclusions for the implementation of the optimal strategy in a market environment.

## **2. STOCHASTIC ELECTRICITY SYSTEM MODEL**

In the following section, welfare-maximizing capacities, outputs and storage decisions are determined to derive the social value of electricity storage. Welfare is interpreted in terms of minimising the annualised system cost of meeting the demand for electricity, or occasionally leaving some unserved at a cost of  $V_{oLL}$ , the Value of Lost Load. The demand to be met,  $D_t$ , is the load net of

1. Powell et al. (2012) use a related approach, breaking a large problem into a sequence of small linear programs, in solving their SMART model.

the output of variable renewable generators such as wind and solar plants. Electricity can be generated by a portfolio of non-intermittent technologies, which we model as a merit order stack, deploying them in order of increasing variable cost without considering dynamic constraints. The vector  $x_t \geq 0$  describes production levels for each non-intermittent generation technology in hour  $t$ .  $x_t$  is limited by capacities  $k$  that have to be set in advance, so that  $x_t \leq k$ .  $s_t$  corresponds to the charging (positive values) or discharging (negative) of the storage in hour  $t$ . We follow the common practice and include a lost load option thus demand need not be met in full if this would lower system costs.<sup>2</sup> The cost of any shortfall is modelled as a term in the objective function:  $VoLL \cdot (D_t + s_t - \sum x_t)^+$ .

Generators have fixed costs per unit of capacity, given by the vector  $c^{fix}$ , and variable costs per unit of output, given by the vector  $c^{var}$ . As we are interested in expected cost over a year, we use annualised fixed cost plus a year's worth of the expectation of variable costs over an infinite period.<sup>3</sup>

The residual load is assumed to evolve according to a Markov process represented by the diurnal transition probabilities  $Prob(D'_i | D_j, h)$ . To introduce a diurnal residual load structure an additional "time of the day" state variable  $h(t)$  is introduced, making the transition matrix vary with the time of day. This reflects the way in which diurnal patterns of load and renewable generation will affect the optimal storage strategy. Its quantification and the adequacy of this assumption will be discussed in the following Section 3. Thus, the time homogenous laws of motion of the dynamic state variables are described as transition probabilities

$$Prob((D'_i, S'_i, h') | (D_j, S_j, h), s) = Prob(D'_i | D_j, h) \begin{cases} 1 & S'_i = S_j + s, h' = 1 + h \bmod 24 \\ 0 & \text{else} \end{cases} \quad (1)$$

Unlike in the perfect foresight model, in the case of uncertainty after retrieving information on the latest residual load many different following residual load states have to be considered. These requirements are met in the Markov Decision Process approach by the determination of an optimal strategy (policy).<sup>4</sup> This will depend upon the state of the system  $(S_t, D_t, k, h, \hat{S})$ , described by the state of charge  $S_t$  [GWh] ( $\hat{S} \geq S_t \geq 0$ ), the residual load level  $D_t$ , the initially determined technology-specific generation capacities  $k$  and the capacity of storage  $\hat{S}$ . The hour of a day is associated to the period number  $t$  via the function  $h(t) = (t-1) \bmod 24 + 1$ .<sup>5</sup>

A strategy  $\pi$  consists of a set of decision rules  $f_t$ , one for each time period,  $\pi = (f_1, f_2, \dots)$ . Each decision rule assigns in practice a single action—the level of non-intermittent generation  $x_t$  and the change of the state of charge (SOC)  $s_t(s_t, x_t) = f_t(S_t, D_t, k, h, \hat{S})$  at each state  $(S_t, D_t, k, h, \hat{S})$ . The formalism follows Ross (1983) and Puttermann (1994) including a distribution (probably degenerate) over the admissible actions. MDP analyses focus on finding a time homogenous (stationary) decision rule  $\pi = (f, f, \dots)$ .<sup>6</sup> Using these definitions, the optimization of the energy system, with residual load modelled by the Markov Chain  $P(D_{t+1} | D_t)$  can be described as a Markov Deci-

2. Modeling the VoLL reminds us of the "penalty" approach to the numerical solution of constrained nonlinear optimization problems. This analogy can be used to apply an (intuitive) theorem that sheds light on the modelling approach: The penalty solution converges to the solution of the constrained problem, as VoLL tends to infinity. Proof: penalty methods e.g. Luenberger (1984).

3. Tijms (2003, p.234) explains: "For many applications of Markov decision theory this criterion is the most appropriate optimality criterion. The average cost criterion is particularly appropriate when many state transitions occur in a relatively short time."

4. For an introduction to the topic see (e.g.) Puterman (1994), Hernandez-Lerma and Lasserre (1996), Bertsekas (2001, 2005), Feinberg and Shwartz (2002) or Baeuerle and Rieder (2011).

5. The modulo (mod) operation finds the remainder if dividing the number of the hour by 24. The plus 1 and minus 1 in the equation ensure that time runs from hour 1 to hour 24, and not from hour 1 to hour 0 (via hour 23).

6. Note that since we use 24 different transition matrices, our decision rule does depend on the hour of the day (which appears as one of its state variables), but it is stationary between days.

sion Process. Therefore, the simplifying approximation of a long but finite horizon by an infinite one is applied. The optimal strategy is a solution of the problem

$$\min_{k,\pi,x} c^{fix}k + \bar{T} \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[ \sum_{t=1}^T c^{var} x_t + VoLL \cdot (D_t + s_t - \sum x_t)^+ \right] \quad (2)$$

considering the state transitions (1) and the storage capacity  $\hat{S}$ .

In absence of any dynamic restrictions the optimization of the generation outputs  $x_t$  follows a merit order. The cheapest technology—in terms of variable cost—is used first; once its capacity is exhausted (we need  $x_t \leq k_t$ ), then the second cheapest technology is applied and so on. This gives rise to an aggregated variable cost step function  $C^{var}(L)$ , where  $L$  is the (residual) load plus storage charging. Load (net of storage discharge) that exceeds the installed generation capacity is treated as having a cost of  $VoLL$ . Inserting both into (2) gives

$$\min_{k,\pi} C(\pi, k) = c^{fix}k + \bar{T} \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[ \sum_{t=1}^T C^{var}(D_t + s_t | k) + VoLL \cdot (D_t + s_t - \sum k)^+ \right] \quad (3)$$

Details of the numerical solution to the stochastic optimization problem can be found in the appendix. In section 5.5 this solution will be compared to an equivalent perfect foresight problem. In this case a period of five years in hourly resolution is simulated for specific draws of the residual load  $D_t$  generated by the Markov process defined in Section 3. The resulting costs are annualised as in the stochastic model. The welfare-maximizing capacity-, output- and storage-decisions under the perfect foresight assumption are the solution of the following problem with  $T = 365 \times 24 \times 5$  and  $\bar{T} = 365 \times 24$ :

$$\min_{x_t, k_t, s_t} c^{fix}k + \frac{\bar{T}}{T} \sum_{t=1}^T (c^{var} x_t + VoLL \cdot y_t) \quad (4)$$

$$\text{s.t.: } y_t \geq D_t + s_t - x_t \quad (5)$$

$$\hat{S} \geq S_t = S_{t-1} + s_t \geq 0$$

$$\bar{s} \geq |s_t| \quad (6)$$

$$k \geq x_t \geq 0 \quad (7)$$

$$y_t \geq 0 \quad (8)$$

$y_t$  is the amount of energy potentially not covered by generation capacities  $k_t \in \{0, 5, 10, 15, \dots\}$ .  $x_t$  is the generation level,  $s_t$  the change in the state of charge  $S_t$ .  $\bar{s}$  refers to the maximum charging and discharging capacity of storage.

### 3. RESIDUAL LOAD

This section asks how valid it is to model the hourly residual load in Germany as a 24-hour or diurnal Markov Chain with 5 GW steps.<sup>7</sup> Our data period is from 2011 to 2015. We start with the total hourly load on the power system as given by ENTSO-E, which excludes self-generation by industry. We subtract renewable generation—in the case of Germany, mainly wind and solar power—

7. See Shamshad et al. (2003) and Brokish and Kirtley (2009) for a discussion of Markov Load modelling.

**Table 1: Benchmarks of Simulated Residual Load in Germany (2011–2015)**

	Load	Wind	Solar	Residual Load
Average annual total (TWh)	480	63	34	382
Day Mean (GW)	1316	174	94	1047
Hourly Mean (GW)	55	7	4	44
Max (GW)	79	33	26	78
Min (GW)	29	0	0	4
Renewable share of Load (%)		13	7	

to get our residual load (Table 1).<sup>8</sup> To eliminate the impact of an increasing share of solar and wind generation, we use the actual hourly load factors scaled by the maximum generation in 2015.

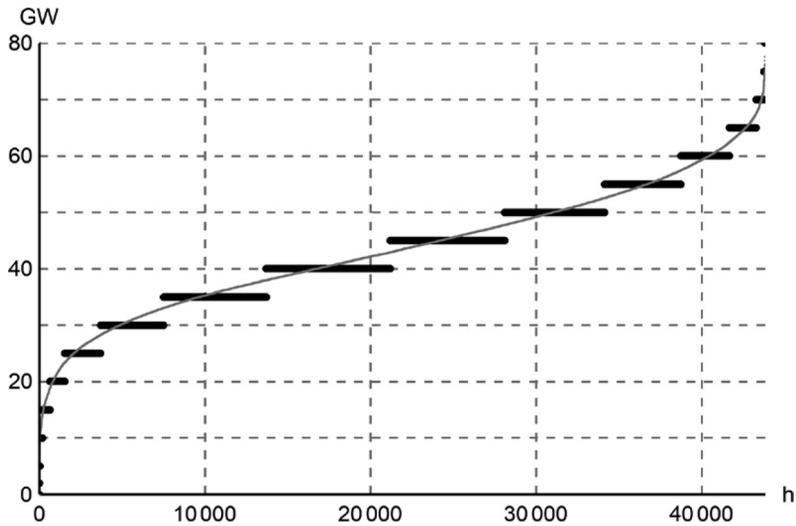
The annual load in Germany amounts to 480 TWh. Wind stations generated 13% of this load and 7% came from solar PV panels. We model the residual load, which is thus reduced to 80% of the gross demand. The peak (gross) load is 79 GW. At their peaks, 33 GW of wind power and 26 GW of solar electricity are produced. Nevertheless, the residual load for thermal generators peaks at 78 GW, since high demand can coincide with very low levels of intermittent renewable generation. The gross load will never fall below 29 GW; the minimum of the residual load, however, is only one seventh of this lower bound (4 GW). It is obvious that renewables a) cover a substantial share of the load (20%), b) lower the minimum load to almost zero but c) hardly decrease peak load. The residual load duration curve is shown as the continuous line in Figure 1.

For this residual load a Markov process of transition matrices  $P_h$  with  $h = 1, \dots, 24$  rounded to 5 GW steps is estimated. The (consistent) maximum likelihood estimators of the transition matrix of a homogeneous Markov chain are the numbers of state transitions normalized per line. The  $16 \times 16$  matrix  $P_1$  gives the probability of reaching each load step in period 2, given the load in period 1, and so on. The probability of reaching each load step in period 3, relative to the load in period 1, is thus  $P_1 \cdot P_2$ . The steady state distribution of load values at  $t=1, p_i^*$ , is derived by solving the equation  $p_i^* = p_i^* \cdot P_1 \cdot \dots \cdot P_{24}$ . The stationary distributions for each hour are presented in Table 2A and visualized in Figure 2. It is clear that the residual load variance during the night is lower than during the day. There are 2 peaks at around 8 am and 5pm. Peak load (which is represented by a step at 80 GW) is reached at 5 pm (Table 2A).

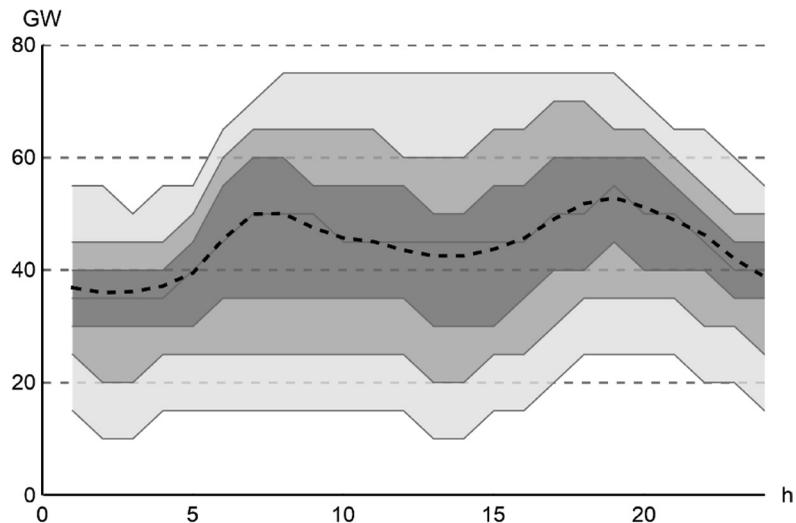
The stationary probabilities are weighted by  $1/24$  and aggregated over the day to give the load duration (Table 2B). The latter matches the average residual load of Table 1 to within 0.01 GW. In terms of the load-duration curve (Figure 1), the step function of our discretised load is a reasonable approximation. When modelling energy storage, the pattern of load over time is critical, and an anonymous referee has wondered whether the Markov Chain contains enough serial correlation to properly model the risk of a low-wind weather regime (Staffell and Pfenninger, 2018) lasting several days that could deplete storage during a long period of high demands. In Section 5.5 below, we run simulations over 100 draws from the Markov process, and observe a number of these events, which lead to lost load even when storage is optimised with perfect foresight. This implies that the Markov approach is able to capture this risk. Nonetheless, annual transition matrices may allow the system

8. Load data are taken from ENTSO-E. These values do not cover all the load in Germany, because industry's production for own consumption and German railways are not included, since hourly values are not available. Annual data from ENTSO-E or BMWI Energiedaten which include these loads are in a range of 530 to more than 600 TWh. For the purpose of the paper, this is not of major importance. Wind and solar generation are the sum of TSO estimates from Amprion, 50Hertz, TenneT and Transnet BW. All data have been accessed via [http://www.pfbach.dk/firma\\_pfb/time\\_series/ts.php](http://www.pfbach.dk/firma_pfb/time_series/ts.php).

**Figure 1: Load Duration: German Data 2011–2015 (continuous line), Stationary Distribution of the 24h-Markov Process in 5 GW Steps (bars)**



**Figure 2: Interpolated Discrete Densities of the Stationary Distributions of the 24-hour Residual Load Markov Process, with 0.5%, 5%, 20%, 50%, 80%, 95% and 99.9% Quantiles.**



to move from high through medium to low residual demand too easily,<sup>9</sup> and using seasonal matrices would keep the winter system mainly in a state of medium or high demand with more serial correlation. This is a matter for future research; in this paper, we next explain our optimisation approach with a simple example.

9. The issue is that the same absolute demand level might represent either a summer day of above-average demands or a winter day of below-average demands. In summer, it is very unlikely that the demand would get higher, and probable that it will fall; the opposite is the case in winter. Annual transition matrices allow the demand to either rise or fall. Separate seasonal transition matrices would produce longer sequences of low-medium (summer) and high-medium (winter) demands.

**Table 2A: Stationary Distributions [%] of Residual Load Levels by Time of the Day**

Residual Load GW	Residual load distribution by hour of the day [% per hour]																								
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
80																	0.1								
75								0	0	0	0	0	0	0	0	0	1	1	0						
70							1	2	2	1	1	1	1	1	1	2	5	5	4	0					
65						0	7	8	6	5	5	4	4	4	5	7	7	9	10	5	1	0			
60						5	14	13	10	9	8	7	7	7	7	7	9	13	16	16	9	3	1	0	
55	0	0	0	0	2	16	18	18	15	12	11	10	9	8	8	10	15	20	22	23	22	14	4	1	
50	4	3	3	5	10	20	17	17	17	16	16	14	13	13	15	17	19	18	17	20	25	25	15	7	
45	14	11	12	15	22	17	10	11	16	18	18	17	17	17	17	17	15	13	13	16	19	26	27	19	
40	28	26	27	28	25	12	10	11	12	13	15	16	15	15	15	13	11	10	10	11	13	18	28	31	
35	29	30	29	25	20	12	9	9	9	10	11	12	13	12	11	11	9	8	5	5	7	8	14	23	
30	14	17	17	15	11	9	7	6	7	7	8	9	10	10	9	9	6	3	2	2	3	4	8	11	
25	7	8	8	7	6	4	4	3	4	5	5	6	7	8	7	5	2	1	0	1	2	2	3	5	
20	3	3	3	3	2	2	2	2	2	2	2	2	3	3	3	2	1	0	0	0	0	0	1	2	
15	1	1	1	2	1	1	1	1	0	1	1	1	2	2	1	1	0			0	0	0	0	1	
10	0	1	1	0	0	0	0	0	0	0	0	0	1	1	0	0				0	0	0	0	1	
5	0.1	0.1	0.2	0.1	0.1	0.1								0.1	0.1									0	0
Mean	36	35	35	36	39	43	47	48	47	45	45	43	42	42	43	45	49	51	52	50	48	45	41	38	

Zeros represent values of less than 0.5%. Each column sums up to 100.

**Table 2B: Expected Load and Corresponding Generation**

Residual Load GW	Expected Residual Load		Generation without storage		
	Frequencies	Duration	Technologies	Capacities	Break even
	%	%		GW	%
80	0.002	0.002	Lost Load	5	0.03
75	0.13	0.13	C Turbine	5	
70	1.08	1.21	C Turbine	5	
65	3.60	4.81	C Turbine	5	
60	6.67	11.48	C Turbine	5	15.5
55	10.80	22.29	C Cycle	5	
50	14.45	36.74	C Cycle	5	42.5
45	16.46	53.20	Coal	5	
40	17.18	70.38	Coal	5	73.8
35	13.72	84.10	Lignite	5	
30	8.45	92.55	Lignite	5	92.9
25	4.51	97.06	Nuclear	5	
20	1.87	98.93	Nuclear	5	
15	0.81	99.74	Nuclear	5	
10	0.23	99.97	Nuclear	5	
5	0.03	100.00	Nuclear	5	
Sum	100		Total	80	

The “expected frequencies” column gives the mean value across each row (and also sums to 100). The “expected duration” column gives the cumulative percentage of periods in which residual load is at or above the level given in that row. The next columns give the type of generator that would be selected to serve each load block in a model without storage, on the basis of costs presented in Table 6 below, which lead to the break-even load factors (at which adjacent technologies have equal costs) shown in the final column.

#### 4. CASE 1: “MINIMAL” ENERGY SYSTEM MODEL

We start with a small-scale model to give intuition for the use of storage, since we are not aware of any analytical solution of the optimization over Markov chains. In our model electricity is generated by a base load and a peak load technology. The peak load technology has double the vari-

**Table 3: Technology Parameters and Break-even Utilization**

Technologies	Lost Load	Peak	Base
Fixed cost	0	18	93
Variable cost	8	2	1
Breakeven load factors		3%	75%

**Table 4: Residual Load Transition Matrix**

$$\text{Prob}(D_i | D_j) = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 & 0 \\ 1/6 & 1/2 & 2/6 & 0 & 0 \\ 0 & 1/8 & 3/4 & 1/8 & 0 \\ 0 & 0 & 2/6 & 1/2 & 1/6 \\ 0 & 0 & 0 & 3/4 & 1/4 \end{pmatrix}$$

(load level 1,1 is in the upper left cell)

able cost of the base load technology, but much lower fixed costs (Table 3). Losing load involves no fixed costs, but four times the variable costs of peak load generation. Total costs are the sum of the expected per-period variable cost multiplied by 100 periods and fixed cost. The bottom row gives the break-even load factors at which the two adjoining technologies would have equal total costs.<sup>10</sup> For these parameters, it is efficient to shed load instead of building peak capacity that would have a load factor of less than 3%, and base load capacity should run with a load factor of 75% or more.

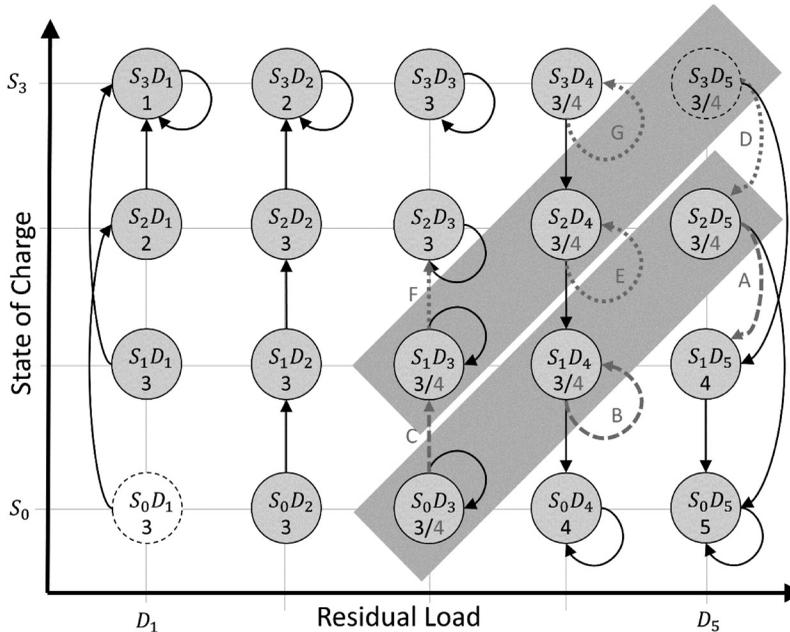
It is assumed that residual load can take discrete values from 1 to 5 and follows a mean-reverting and persistent Markov process like the one estimated in the previous section (Table 4). Without storage generation equals residual load and it is straightforward to determine the expected load duration as stationary probabilities of the process: (100, 93.6, 74.5, 23.4, 4.3) [in % starting with the lowest level]. The screening-curve approach then gives optimal capacities as 2 base load units and 3 of peaking, with no load lost, which we write as the vector (2,3,0). Total system costs are 637.

We now introduce storage with (energy and power) capacity of 3 units. The state of charge can take values from 0 to 3. Each state is now described by a residual load value and the state of charge. The state transition including the strategy defines a Markov Chain whose stationary distribution can be determined by solving a linear equation system. We represent this strategy in Figure 3. The level of demand is on the horizontal axis, and the state of charge at the start of the period on the vertical. The upper row of numbers in each circle give the state of charge at the start of the period and demand level—the top left circle (S3/D1) has storage of 3 units and demand of 1. The lower number gives the chosen level of generation in a strategy—some cells contain two numbers as we consider alternative strategies. Generation is equal to the level of demand, plus the amount of storage charging, less discharging. The arrows show the amount of charging or discharging and hence the state of charge in the next period, but do not represent the possibility that the level of demand will change. The two cells with dotted outlines are never reached, given the strategy chosen.<sup>11</sup>

10. The break-even load factor (expressed as a fraction) equals  $(\text{Fixed Costs}_A - \text{Fixed Costs}_B) \div (\text{Variable Costs}_B - \text{Variable Costs}_A)$  where Fixed Costs are annualised and the variable costs are the total amount for a whole year’s operation at full capacity.

11. For example, the left-hand column, with a demand of 1, can only be reached after a period with demand of 1 or 2, but the strategy for periods with those demands gives a minimum storage level of 1 in the following period (from cell S0/D2), and so we would never get to cell S0/D1.

**Figure 3: Strategy as a Function of Residual Load and State of Charge**



Graph visualizing a strategy = the change in state of charge defined for any state as arrows. A, B and C (etc.) indicate local changes in the strategy. The graph is a pre-step to the complete graph of the Markov Chain that also includes the residual load transitions. The lower numbers give the state-specific generation after charging/discharging. Dotted states are not reached.

A purely arbitrage-driven (“Full Arbitrage”) strategy without precautionary motives would attempt to equilibrate generation levels perfectly—in this case to 3 units of generation. If demand equals 2 units (the second column), 3 units can be generated and 1 placed into storage, shown by the vertical arrows pointing up. This is not possible if the storage is already full (in the top row) and so generation in that cell falls to 2 units. Similarly, with demand of 4 units (the fourth column), 1 unit can be taken from storage (the vertical arrows pointing down), except in the bottom row, when storage is empty and a fourth unit must be generated. Only 1 unit can be generated in the top left cell, and 5 units of generation (possibly including “generation” by declaring a loss of load event) are required in the bottom right cell. This is the strategy represented by the black numbers and solid lines in Figure 3. It is straightforward to determine the stationary probability of each state, the expected level of output from each generator and hence optimal capacities, all shown in Table 5. As there is less variation in the resulting load compared to the model without storage, a unit of peak load generation is substituted by base load, giving capacities (3, 2, 0). Total costs are reduced by 1%. The right-hand column of the probability matrix shows that the storage is fully charged for 40.8% of the time, and fully discharged for 38.6%.

We now consider deviations from the “full arbitrage” charging strategy that reduce the risk of high generation levels and further reduce costs. With change A (Figure 3), generation in cell  $S_2/D_5$  is increased to 4 (the second number, in grey) so that discharging is reduced by 1 unit (the grey dashed line). Thus, if in the next period residual load would stay at 5 then generation level 5 could be avoided. This changes the stationary probabilities and the load duration: the probability to reach generation level 5 is reduced by a third of the level with full arbitrage. In the same way, it is possible to modify the full arbitrage strategy by changes A, B and C (Figure 3) and determine the resulting load duration and costs (Table 5).

**Table 5: Results from the Simple Model, given Storage Strategies:**

“Full Arbitrage”	“Full Arbitrage” + A-C	“Full Arbitrage” + A-G (optimal)
Stationary probabilities of the states and row total [%]		
$\begin{pmatrix} 5.6 & 13.0 & 19.6 & 2.5 & 0 &   & 40.8 \\ 0.3 & 1.7 & 5.9 & 2.0 & 0.4 &   & 10.3 \\ 0.4 & 2.0 & 5.9 & 1.7 & 0.3 &   & 10.3 \\ 0 & 2.5 & 19.6 & 13.0 & 3.5 &   & 38.6 \end{pmatrix}$	$\begin{pmatrix} 5.9 & 14.1 & 21.9 & 2.7 & 0 &   & 44.7 \\ 0.5 & 2.3 & 7.3 & 2.3 & 0.5 &   & 12.8 \\ 0 & 2.7 & 2.0 & 8.4 & 1.9 &   & 33.1 \\ 0 & 0 & 1.9 & 5.7 & 1.9 &   & 9.5 \end{pmatrix}$	$\begin{pmatrix} 6.4 & 17.5 & 38.1 & 9.5 & 1.6 &   & 73.1 \\ 0 & 1.6 & 11.0 & 5.4 & 1.3 &   & 19.4 \\ 0 & 0.1 & 1.2 & 2.1 & 0.7 &   & 4.1 \\ 0 & 0 & 0.7 & 2.0 & 0.7 &   & 3.4 \end{pmatrix}$
Expected Load duration for generators, taking account of storage (dis)charging, cumulated		
(100, 94.4, 81.0, 16.9, 3.5)	(100, 94.1, 79.5, 20.3, 1.9)	(100, 93.6, 76.1, 25.3, 0.7)
Optimal generation capacities (Base, Peak, Lost Load)		
(3, 2, 0)	(3, 1, 1)	(3, 1, 1)
Total cost [%] relative to those without storage		
-1%	-1.8%	-2.4%

With these changes, the probability of an empty storage falls from 38.6% to 9.5%—particularly, the probability of reaching load level 5 with an empty storage halves from 3.5% to 1.9%. Thus, losing load appears sufficiently rare to be acceptable and generation capacity is reduced to four units. This comes at the cost that generation in the states S0/D3, S1/D4 and S2/D5 increases from 3 to 4. Additional changes to the strategy (D, E, F and G; Figure 3) further halve the probability of reaching load level 5 with an empty storage. Total costs are 1.4% lower than with the “full arbitrage” strategy. This strategy (found with the algorithm described in Appendix I, checked by brute force) proves to be optimal. Further changes to the strategy could reduce the loss of load probability below 0.7% but the storage would be discharged less. This means it less frequently substitutes for peak generation, reducing the benefits of having the storage. This means that making the strategy more precautionary is not the best option.

The attractiveness of precautionary strategies depends on VoLL. Deviations D, E, F and G reduce the loss of load probability by 1.2%, but they also reduce the 100-period output of the baseload generators by 3.9 units (= (79.5%–76.1% + 94.1%–93.6%) x 100). The precautionary strategy gives a greater probability of entering a low-demand state with high levels of storage, when the baseload generators cannot all run. Since peaking generators with a variable cost of 2 are used to charge that storage, instead of baseload generators with a variable cost of 1, the change adds 3.9 to generation costs. Reducing lost load by 1.2 units gives a net saving of 7.2 (VoLL of 8 per unit, less the variable cost of the peaking generator, which is 2 per unit). If VoLL were lower, however, equal to 5 or less, those four deviations (taken together) would not be worthwhile, since the net saving from avoiding lost load has fallen to 3 per unit. Over 100 periods, the three deviations would save 3.6 from avoiding lost loads, less than the 3.9 cost of substituting peaking for baseload generation.

If VoLL were higher, on the other hand, precautionary strategies might also be sub-optimal. Adopting deviations A-G gives a loss of load probability of 0.7%, in return for saving the 18-unit cost of a peaking generator. If VoLL were 32, the net cost of a unit of lost load would be 30, and the expected cost of unserved energy over 100 periods would be 21. That makes the precautionary strategy presented here worse than the full arbitrage strategy with enough generating capacity to always serve the load.

Hence, there is a window of VoLL such that precautionary storage strategies together with the substitution of generation capacity are optimal. But can we expect this window to include realistic levels of VoLL and persist under more realistic conditions? This question is addressed in the next section.

## 5. CASE 2: “REAL WORLD” ENERGY SYSTEM MODEL

We now apply our model to the German load data from Section 3 with generator cost data based on Schröder et al. (2013). In our central case, we analyse the impact of adding storage with a charging/discharging capacity of 30 GW and an energy capacity of 300 GWh (equivalent to 6 hours of average load in Germany; for an overview of storage potential in Europe see Ess et al. (2012)).

Table 6 presents the costs for five types of plant, with investment costs converted to annuities at an interest rate of 9%. The bottom row gives the load factors at which two adjacent plant types have equal costs over the year. Nuclear plants (with low variable costs but the highest fixed costs) are the cheapest option if they can run at a load factor of more than 93%; lignite for load factors between 74% and 93%, and so on. We present our results in terms of costs over one year, scaling the representative 24 hours by a factor of 365. The grid size for the levels of residual demand is set to 5 GW.

**Table 6: Cost Data and Break-even Load Factors**

Technology	Nuclear	Lignite	Coal	Combined Cycle	Combust Turbine	Lost Load
Variable Cost <sup>a</sup> €/MWh	14	38	42	59	94	18,000
Fixed Cost <sup>b</sup> €/kW-year	387	189	163	103	55	0
Breakeven load factor %	93	74	43	16	0.03=3h/a	

Source: derived from Schröder et al. (2013).

<sup>a</sup> (Fuel + variable O&M)

<sup>b</sup> (Capital cost and fixed O&M)

Following e.g. Qiu and Girgis (1993), the social planner can decide upon capacity and production in a way that some demand is not met, incurring a cost of  $VoLL$  instead, if this reduces system costs. Therefore, losing load was modelled as the application of a virtual technology with variable costs of  $VoLL$  and without fixed costs. We calibrated it to a reliability target of 1 day per ten years or 3.5 hours per year (Enernex, 2006)<sup>12</sup> by assuming that losing load ‘breaks even’ with the most expensive technology—a combustion turbine—for the highest 3.5 hours of load per year. The implicit  $VoLL$  exceeds the variable cost of the combustion turbine by 200 times.

We run several versions of the model, with and without storage; with perfect foresight and with no ability to predict the future beyond the probability distributions of the Markov process. First (Section 5.1), we validate our discrete Markov process for demand by comparing the outcomes—with perfect foresight in all cases—for the actual (continuous) loads and for the Markov process, with and without storage. We took 100 draws of the Markov process, ran the model (with perfect foresight) for each of these, and calculated the means of the results. We find a very close agreement in optimal capacities and in costs for the cases with no storage. The two cases with storage are slightly further apart, but still acceptable.

We then (Section 5.2) present three scenarios for the Markov process, our core results. First, we calculate the optimum strategy for a scenario with no storage. Second, we add 300 GWh of storage energy capacity, with 30 GW of power capacity, and find the optimal operating and investment strategy for storage and conventional generation. Conventional generating capacity is significantly less than the maximum demand of 80 GW, and so the optimal storage strategy involves some precautionary storage, forgoing short-run arbitrage opportunities in order to retain energy in case of

12. As a rule of thumb for a system-wide loss of load, commercial users estimate it at several thousand €/MWh. For a literature review see London Economics (2013a,b) or Growitsch et al. (2013).

a prolonged period of high demand. We discuss this scenario in some detail, as it contains our main results on how precautionary storage would work in practice. We also compare it with a scenario in which we optimise storage, holding conventional generation capacities constant, and find that just over half the value of storage comes from the ability to adjust the capacity mix in response to the new output-duration curve.

We follow these with four additional scenarios (Section 5.3) for the Markov process, varying first the energy and then the power capacity of storage. This allows us to calculate how the value of storage varies with each dimension of capacity. For the scenarios that we tested, the value of storage is more sensitive to its energy capacity than to its power capacity.

In Section 5.4, we consider two scenarios in which it is considered socially unacceptable for the system operator to shed load, and so total generating capacity is held at the level of maximum demand (80 GW). One scenario has no storage capacity, while the other adds our base-case amount of storage (300 GWh and 30 GW). This storage is used for optimal arbitrage and not for precautionary reasons, since there is always enough generation capacity (the model assumes there are no outages). While there are still benefits from adjusting the mix of capacity types, the overall gain from storage is roughly halved compared to scenarios where occasionally losing load is optimal.

Finally, in Section 5.5 we compare our results under uncertainty with the perfect foresight findings based on Markov simulations of residual load as validated in Section 5.1. We find that the value of storage under uncertainty is 27% lower than in the model with perfect foresight.

The results are presented in Tables 7A and 7B. Each scenario is numbered from 1 to 15 according to its order in those tables; we use the abbreviation PF for scenarios with perfect foresight, and MPS for the Markov Process Strategies. The PF scenarios are further distinguished by the use of the original continuous load (PF-C), the actual load rounded to 5-GW steps (PF-RC) and by running the Markov process to generate (perfectly foreseen) time series (PF-M). We show the energy and the power capacity of the storage with two numbers—scenarios with 0/0 have no storage, whereas our base case, 300/30, has 300 GWh of energy capacity and a maximum charge/discharge rate of 30 GW. Two suffixes show the cases of optimal storage operation for fixed generation capacities (KF) and when it is not acceptable to lose load (NLL).

## **5.1 Validation of Residual Load Generated by a Markov Process**

We first validate our Markov process model, comparing results for generation capacities and costs against those based on the original load data, assuming perfect foresight. There are two steps which might distort our results. First, the data is discretised in 5 GW steps; second, we replace the actual sequence of discretised demands with the Markov transition matrices. We need to ensure that neither step leads to significant changes. The first step is tested by optimising the model of equations 4–8 against the original and the discretised loads, first without and then with storage. To test the second step, we use the Markov process to generate 100 5-year runs of residual loads, optimise the model against each run, and take the average. We solve the deterministic model using CPLEX in GAMS; a typical deterministic run takes 2 minutes on a standard laptop. As well as validating the model against a more traditional approach, these perfect foresight runs can be compared with our later results to show the effect of ignoring uncertainty.

The results without storage (1. PF-C 0/0, 2. PF-RC 0/0 and 3. PF-M 0/0) in the top three rows of Tables 7A and 7B are very similar. Optimal capacities differ by less than 2 GW and total costs by less than 0.1%. This could be expected as the optimal system defined by equations 4–8 depends on the residual load duration only—and the stationary Markov process distribution fits the

**Table 7A: Scenarios of Storage Management and Capacity Adjustment**

Scenario Code	Storage		Generation [GW]					Total	Lost Load
	GWh	GW	Nuc.	Lig.	Coal	C.C.	C.T.		
Perfect Foresight results—validation									
1. PF-C 0/0	0	0	27	9	9	10	19	75	3
2. PF-RC 0/0	0	0	25	10	10	10	20	75	5
3. PF-M 0/0	0	0	26	9	10	10	20	75	3
4. PF-C 300/30	300	30	34	6	6	5	13	63	0
5. PF-RC 300/30	300	30	35	5	5	5	15	65	0
6. PF-M 300/30	300	30	34	6	5	5	10	60	0*
Core results—optimization with Stochastic Markov Process for residual load									
7. MPS 0/0	0	0	25	10	10	10	20	75	5
8. MPS 300/30 KF	300	30	as above						
9. MPS 300/30	300	30	30	5	15	0	10	60	20
Sensitivity:									
Varying storage energy capacity (holding storage power constant) with stochastic residual load									
10. MPS 150/30	150	30	30	10	5	5	15	65	15
9. MPS 300/30	300	30	30	5	15	0	10	60	20
11. MPS 450/30	450	30	30	5	15	0	10	60	20
Sensitivity:									
Varying storage power capacity (holding storage energy constant) with stochastic residual load									
12. MPS 300/10	300	10	30	10	5	5	15	65	15
13. MPS 300/15	300	15	30	5	15	0	10	60	20
9. MPS 300/30	300	30	30	5	15	0	10	60	20
Sensitivity:									
No need for precautionary storage as conventional capacity kept equal to maximum demand									
14. MPS 0/0 NLL	0	0	25	10	10	10	25	80	0
15. MPS 300/30 NLL	300	30	30	10	5	5	30	80	0

\* The mean amount of lost load is very close to zero, but there are a few runs where 5 GW of load is occasionally shed

empirical load duration very well, as already discussed in Section 3. Without storage, the optimal generation capacity is 75 GW. The occasional loss of 5 GW of load is accepted; the probability of an 80 GW residual load is 0.0023%, below the level of 0.035% at which load-shedding breaks even with peaking capacity.

With storage and perfect foresight, the differences between 4.PF-C 300/30 and 5.PF-RC 300/30 caused by rounding are equally small. However, the runs of the Markov process reported in 6.PF-M 300/30 do not contain any seasonal patterns, and this reduces generation capacity compared to PF-RC 300/30 by 5 GW of combustion turbines, from 65 GW to 60 GW. This helps reduce the total cost of generation by 1.4% compared to 5.PF-RC 300/30. The cost estimates across the scenarios with storage therefore remain close, but the value of storage depends on the difference between these estimates and those from scenarios without storage. The estimated value of storage under perfect foresight (in the final column of Table 7B) is not much affected by rounding, but it does increase by 25% (from 4.4 to 5.6 Euro/kWh of storage capacity) when we use the non-seasonal Markov process to generate our load series. Intuitively, the Markov process removes seasonal autocorrelation and gives more short-term variability that can be exploited by the storage and results in an overestimation of its value. The estimated costs and capacities are close enough, however, to support our view that the Markov residual load process is a sensible way to introduce uncertainty into an electricity model.

**Table 7B: Scenarios of Storage Management and Capacity Adjustment**

Scenario Code	Comment	Variable Cost €/MWh	Fixed Cost Mio €	Total Cost Mio €	Storage Value €/kWh	
Perfect Foresight results—validation						
1. PF-C 0/0	Continuous residual load 2011–15	26.41	15,829	25,923		
2. PF-RC 0/0	As above rounded to 5 GW steps	27.69	15,317	25,900		
3. PF-M 0/0	Av. Markov residual load process	27.10	15,544	25,896		
4. PF-C 300/30	Continuous residual load 2011–15	21.43	16,309	24,500	4.7	
5. PF-RC 300/30	As above rounded to 5 GW steps	20.82	16,636	24,592	4.4	
6. PF-M 300/30	Av. Markov residual load process	21.34	16,093	24,244	5.6	
Core results—optimization with Stochastic Markov Process for residual load						
7. MPS 0/0	Optimal cap. with no storage	27.69	15,317	25,899		
8. MPS 300/30 KF	Opt. storage management	Capacities fixed	26.30	15,317	25,368	1.8
9. MPS 300/30		Capacities adjusted	23.89	15,547	24,676	4.1
Sensitivity:						
Varying storage energy capacity (holding storage power constant) with stochastic residual load						
10. MPS 150/30	Power constraint 30 GW	24.35	15,649	24,954	6.3	
9. MPS 300/30	Power constraint 30 GW	23.89	15,547	24,676	4.1	
11. MPS 450/30	Power constraint 30 GW	23.36	15,547	24,475	3.2	
Sensitivity:						
Varying storage power capacity (holding storage energy constant) with stochastic residual load						
12. MPS 300/10	Power constraint 10GW	23.93	15,649	24,793	3.7	
13. MPS 300/15	Power constraint 15GW	23.93	15,547	24,692	4.0	
9. MPS 300/30	Power constraint 30 GW	23.89	15,547	24,676	4.1	
Sensitivity:						
No need for precautionary storage as conventional capacity kept equal to maximum demand						
14. MPS 0/0 NLL	Not acceptable to lose load	27.64	15,590	26,155		
15. MPS 300/30 NLL		23.32	16,599	25,511	2.1	

Storage is valued by the system cost savings (including investment) relative to the storage capacity.

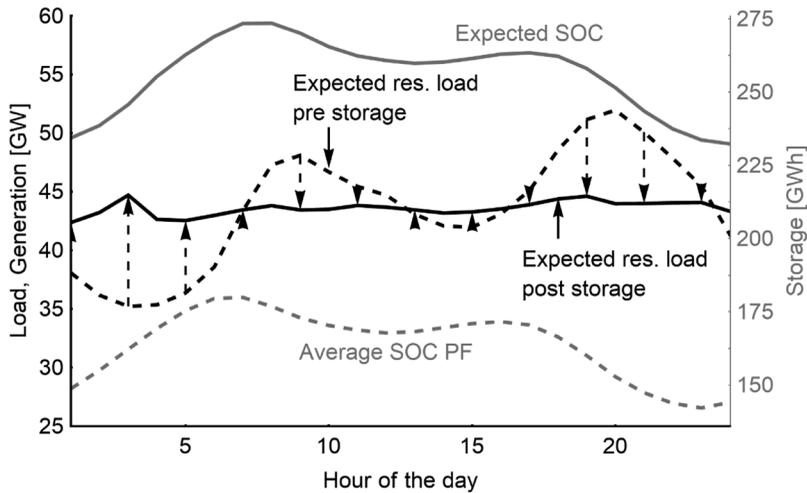
## 5.2 Optimisation of the Stochastic Base Case

We now turn to the case of uncertain residual load, based on the diurnal transition probabilities derived in Section 3, with 300 GWh storage and 30 GW un-/loading capacity (9.MPS 300/30). The solution algorithm is presented in the appendix, but consists of optimising the operation of generation and storage as a cost-minimising Markov Decision Problem for fixed capacities, obtaining the expected duration curve for generation, and adjusting generation capacities to be optimal for this duration curve. The process then iterates until an optimal solution is found (and checked by direct comparisons against neighbouring capacity mixes). We implemented our algorithm in Mathematica; our central case with storage took 45 minutes on a standard laptop, including some pre-processing time. We compare the results to the optimal solution of the stochastic model without storage (7.MPS 0/0), which depends only on the load-duration curve and is identical to that for the rounded loads (2.PF-RC 0/0) presented in the previous section.

The optimal strategy with storage reduces non-intermittent generation capacity by 15 GW from 75 GW to 60 GW (Table 7A). This means that 20 GW of discharge from storage would be required to meet the peak demand of 80 GW—well within the 30 GW that we have assumed.<sup>13</sup> There

13. We continue to use the higher power capacity in most of our scenarios to ensure that it does not suddenly become a binding constraint affecting the solutions we obtain.

**Figure 4: Expected Loads and State of Charge (SOC) in the Markov model and SOC in one Draw of the Perfect Foresight model based on the Markov Process.**



is a positive, but very small, probability of a post-storage load of more than 60 GW, which would occur if a low level of stored energy coincided with a high demand. This probability is low enough that accepting the risk of losing load is better than building additional generating capacity.

Moving from the load-duration to the pattern of load over time, Figure 4 confirms that the mean pre-storage residual load has the same structure as shown in Figure 2. There is an overnight trough, with a morning and a (higher) evening peak separated by a period of intermediate load. Figure 2 confirms that there is significant variability around all of these means—in particular, the highest possible value of 80 GW was seen at 5 pm, well before the highest mean value at 8 pm.

On average, storage operation conforms to intuition: loading during the low residual load time from midnight until 7 am raises the mean SOC (the top line of Figure 4) from 230 GWh to 270 GWh. Shaving the 9 am peak reduces the mean SOC to 260 GWh, offset by small amounts of charging (in expectation) during the afternoon low load around 3 pm. The evening peak gives another period of unloading from 6 pm, with the expected SOC falling back to 230 GWh at midnight. The mean post-storage residual load varies far less than the residual load before storage. Thus, it confirms the stabilization reached by the storage operation during the day.

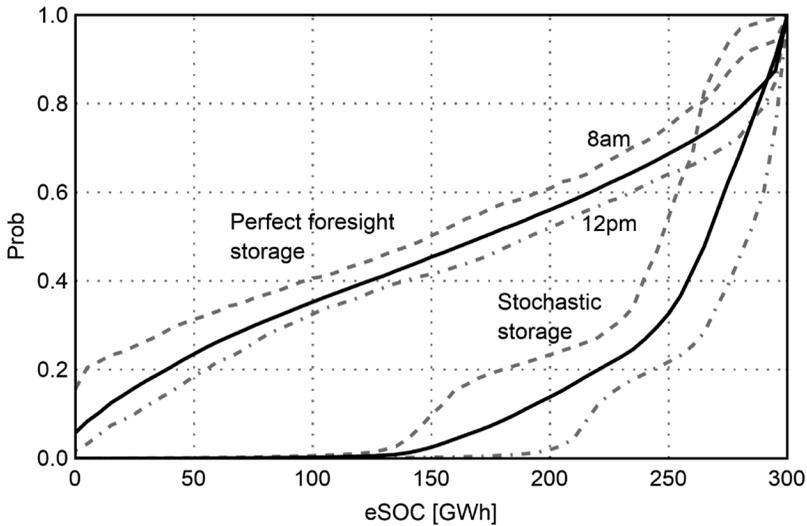
Besides these elements of the storage strategy to exploit the expected diurnal scarcity pattern, there are significant strategy elements built in to further deal with the uncertain development of the residual load. These elements can be identified by taking a closer look at the aggregate distribution of storage levels (shown in Figure 5) and the optimal storage strategies.

The SOC is distributed unequally with a median at 85% of the storage capacity (the right-hand set of lines in Figure 5); close to its upper capacity limit and far away from its lower capacity limit.<sup>14</sup> This is achieved by establishing a (time-varying) lower threshold SOC that is almost never undercut. Figure 6 shows that this threshold is highest (190 GWh) at 8 am, ready for the start of the higher-load daytime hours, and lowest at 10 pm (with a value of 110 GWh) when several hours of night-time recharging can be expected to follow.

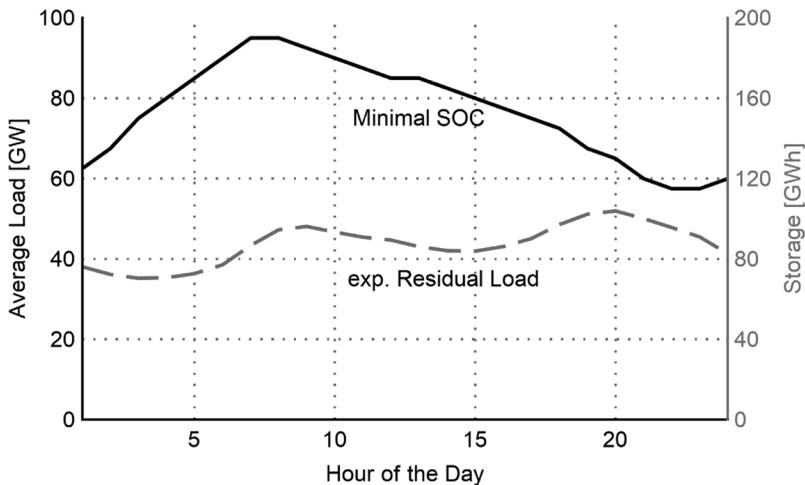
The genesis of this threshold can be derived by examining the optimal charging strategy for the high-demand hour 18, shown in Figure 7. In the top right part of the Figure, the storage is

14. The differences to the optimal storage operation under perfect foresight will be discussed in Section 5.5.

**Figure 5: Aggregate Distribution of the Storage Level at 8 AM (dashed) and 12 PM (dotted), and Diurnal Average (solid)**



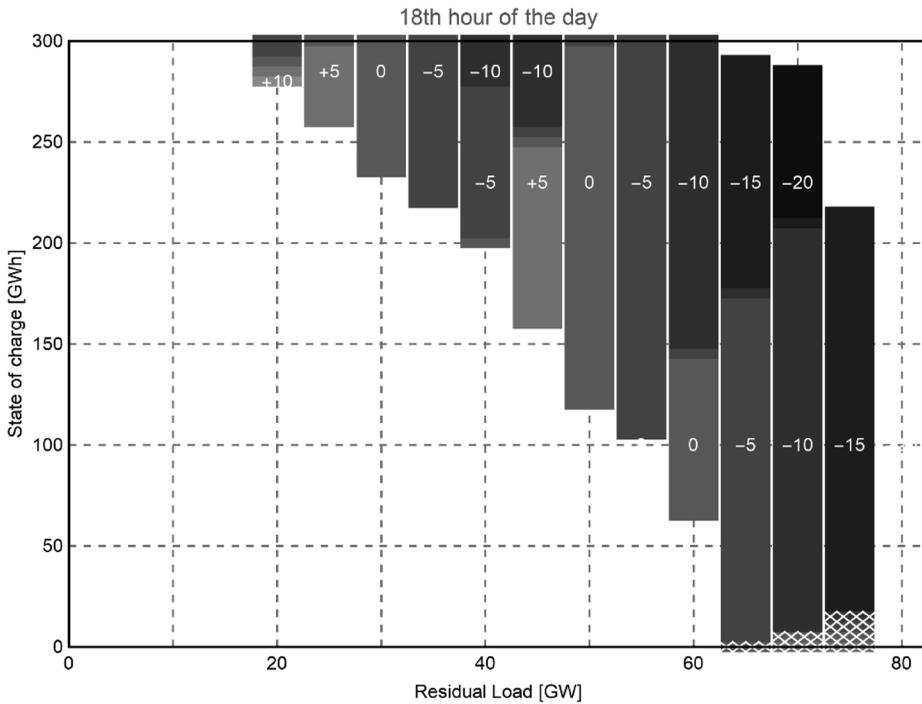
**Figure 6: Lower Level of the State of Charge (solid curve) and Average Load (dashed curve).**



nearly full and is effectively following a “full arbitrage” strategy. The row of numbers level with the 230 GWh point on the vertical (state of charge) axis shows that as load rises from 45 GW to 70 GW, the storage switches from 5 GW of charging to 20 GW of discharging. The amount of conventional generation needed remains at 50 GW, shown in the corresponding area of Figure 8. In both figures, states with a very low probability of occurring, given the optimal strategy, are blank.

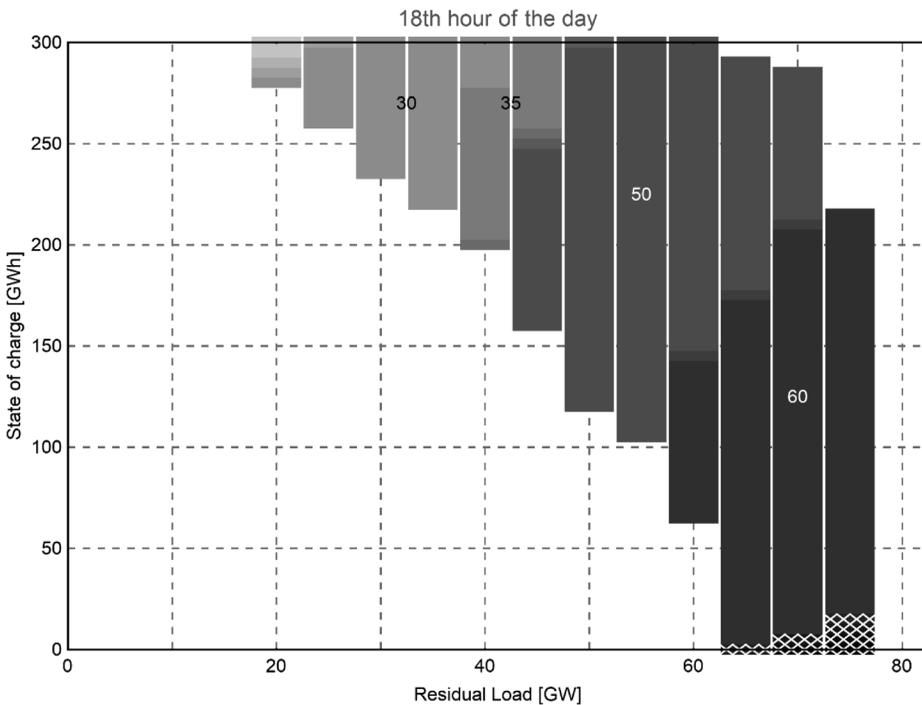
Moving down Figure 7, if the store holds less energy, the optimal strategy involves less discharging. For example, with a pre-storage load of 60 GW, the optimal strategy with a state of charge (an SOC) of 150 GWh or more is to discharge 10 GW and generate 50 GW. If the SOC is 140 GWh or less, it is best to stop discharging and meet the entire load from generation. This reduces the probability of reaching a future load of 65 GW or more (which is greater than the generation capacity with this strategy) with insufficient stored energy to avoid lost load.

**Figure 7: Optimal Charging Strategy in Hour 18**



States with probability  $\leq 1 \times 10^{-6}$  are blank; lost load marked by white hatching.

**Figure 8: Non-intermittent Generation Resulting from the Optimal Charging Strategy**



Lost load marked by white hatching

Similarly, for a pre-storage load of 65 GW, the optimal strategy switches from discharging 15 GW with an SOC of 180 GWh or more to discharging just 5 GW with an SOC of 170 GWh or less. For a load of 55 GW, the strategy even switches to charging (by 5 GW, the maximum possible) if the SOC falls to 95 GWh or less. Note that these changes occur long before the SOC has fallen so low that higher discharges are physically impossible; we are seeing precautionary storage. Figure 8 shows that this results in many states with generation of 60 GW, even though the storage could physically have been discharged by enough to reduce generation to 50 GW.

The bottom right corner of Figure 7 shows a few states with cross-hatching, in which some load is lost because the sum of generation and storage discharge is less than demand. All of these states are rare; the overall probability of losing load is less than 0.004%. In most of these states, the amount of stored energy is so low that it is physically impossible to discharge enough electricity to fill the gap between demand and generation—consider the case of an empty store and load of 65 GW when generation capacity is only 60 GW. The diagram does imply that load would be lost in a period with demand of 75 GW and an SOC of 15 GWh, which could be physically met by draining the store. However, as a 75GW load in hour 18 is certain to be followed by a load of either 70 GW or 75 GW in hour 19, then load must be lost in one of those two periods and it does not matter which.

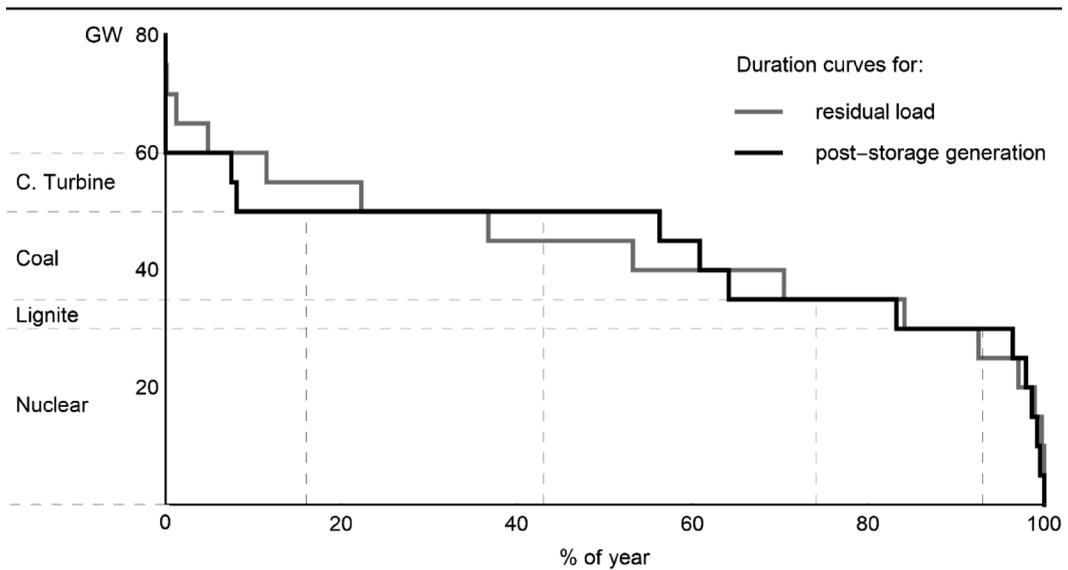
One other feature of the operating strategy deserves comment. In the same way that storage is operated to give generation of 50 GW if the residual load is high and the SOC good enough, the area at the centre of the top of Figure 8 shows a range of demands for which generation is kept at 30 GW. Figure 7 reveals that while the very lowest loads lead to charging (unless the store is already full), discharging is seen for a load of 35 GW or 40 GW. This would reduce the output from the 7<sup>th</sup> generation unit (which is a lignite unit with a variable cost of €38/MWh) and must imply a high probability that the store can be refilled later at a time when a nuclear unit with a variable cost of €14/MWh is the marginal generator. For the lowest SOC likely to be seen with a load of 40 GW, the store does switch to charging, giving generation of 45 GW, but Figure 8 shows that there are very few states with generation at this level at this time.

While Figures 7 and 8 do not show the probability of reaching each state, or the strategy for any other hours, Figure 9 shows the overall load-duration curves, with and without storage. The post-storage load-duration has long horizontal segments at 30 GW, 35 GW and (in particular) 50 GW—there is a 48% probability that generation will equal 50 GW. There is also a 7% probability that the full 60 GW of generating capacity will be in use, whereas without storage, the probability of needing the last 10 GW of (optimal) capacity was under 1.5%.

The generation duration curve allows us to identify the optimal capacity mix, although it should be remembered that the capacity mix and storage strategy are jointly determined. The dashed vertical lines show the break-even points between different generation types, from Table 6. The right-hand vertical line and lowest horizontal line show that 30 GW of nuclear capacity is optimal, since generation is 30GW or more for more than 93% of the time. The second pair of lines show that the 74<sup>th</sup> percentile of generation is 35 GW, giving an additional 5 GW of Lignite capacity. Note that the 16<sup>th</sup> and 43<sup>rd</sup> percentiles of generation are identical, at 50 GW, and so there is no need for CCGT capacity in this equilibrium. The tenth tranche of capacity runs enough for coal to be the cheapest option; the eleventh for so little time that a combustion turbine is cheapest. This particular result is brought about by the specific numbers in our case study; we do not claim that it is general. Finally, there is a very small probability that “generation” (including lost load) exceeds 60 GW at times of high demand and depleted storage, but this is sufficiently rare that it is optimal to keep generation capacity at 60 GW.

Overall, optimal storage operation and capacity adjustment reduced total cost by 4.7%. Putting it another way, the value of storage is 4.1 € /kWh of storage capacity (comparing scenarios 7.MPS 0/0 and 9.MPS 300/30 in Table 7B). We also optimised the operation of storage while hold-

**Figure 9: Load- and Generation-Duration Curves, and the Optimal Generation Capacity Mix.**



ing generation capacity constant (scenario 8.MPS 300/30 KF), finding that this gave 44% of the cost savings, implying 56% are enabled by generation capacity adjustments.

### 5.3 Sensitivity of the Results with respect to Storage Capacity

How does changing the amount of storage affect its value? An extension of the energy capacity from 300 GWh to 450 GWh (11.MPS 450/30) holding the un-/loading capacity constant at 30 GW, did not change optimal generation capacities. However, the additional storage capacity allowed more arbitrage, decreasing variable costs from 23.89 Euro/MWh to 23.36 Euro/MWh. This induced a total cost decline of 0.8%. The average value of storage declined from 4.1 Euro/kWh to 3.2 Euro/kWh, or by 22%.

Halving the storage capacity to 150 GWh (9.MPS 150/30) made it impossible to sustain the same precautionary operation as with 300 GWh. Therefore 5 GW less of the generation capacity could be substituted with the storage. Also, the levelisation of the load was less complete, so that 5 GW of combined cycle capacity remained operational instead of being substituted by coal and lignite. Therefore, fixed costs could not decrease as much as in the 300 GWh case and the variable cost of 24.35 Euro per MWh exceeded the 300 GWh storage case (23.89 Euro/MWh). Total cost savings fell by 22%, but the per-kWh average value of storage capacity increased by 54% from 4.1 Euro/kWh to 6.3 Euro/kWh. The marginal value of storage energy capacity declines rapidly from 6.3 Euro/kWh (for the first 150 GWh) over 1.9 Euro/kWh (the next 150 GWh) to 1.4 Euro/kWh (the third 150 GWh).

So far, the simulations have been run with constant storage power capacity of 30 GW. This exceeds requirements as the distribution of charging and discharging amounts (for scenario 9.MPS 300/30) in Table 8 reveals. While in 31.9% of all cases the storage is neither charged nor discharged, the probability of an un-/loading speed greater than  $\pm 15$  GW is less than 0.03%. Therefore, a reduction of the un-/loading speed limit to  $\pm 15$  GW (13.MPS 300/15) does not induce capacity adjustments, even though peak load of 80 GW can no longer be served by the combination of maximum

**Table 8: Probability distribution of un-/loading speeds.**

Probabilities %	Un-/loading speed [GW]												
	-30	-25	-20	-15	-10	-5	0	+5	+10	+15	+20	+25	+30
MPS 300/30	0	0	0.02	1.6	10.7	23.3	31.9	19.4	9.7	3.2	0.07	0	0
PF-M 300/30	0	0.2	1.3	4.2	9.9	17.2	32.2	19.7	10.9	3.4	0.9	0.1	0.03

unloading and a generation capacity of 60 GW. However, these rare cases increase variable costs from 23.89 Euro/MWh to 23.93 Euro/MWh and the average value of storage decreases slightly, by 0.1 Euro/kWh to 4.0 Euro/kWh.

A significant effect only occurs if the un-/loading speed is limited to  $\pm 10$  GW (12.MPS 300/10). In this case only 15 GW of generation capacity can be substituted by the storage instead of 20 GW (MPS 300/30). This reduces the average value of storage from 4.1 Euro/kWh to 3.7 Euro/kWh.

#### 5.4. Sensitivity of the Results with respect to the Lost Load Option

The discussion at the end of Section 4 showed that load shedding becomes uneconomic if the VoLL is either close to peak generation costs or if it is extremely high. In both cases the precautionary motive disappears and the storage is used for arbitrage only. We simulated the case of a very high VoLL by omitting the option to lose load.

Without storage, (14.MPS 0/0 NLL), this increases total costs by 1% (compared to 7.MPS 0/0) because another 5GW of combustion turbines have to be installed, even though very rarely used. Storage (15.MPS 300/30) is used for arbitrage only with no need for precautionary storage, which reduces the threshold storage level and extends the potential to change the post-storage residual load. Compared to scenario 9.MPS 300/30, 10 GW of the coal generation capacity is substituted with 5GW of lignite (more intensively used) and 5 GW of Combined Cycle capacity (with a lower load factor).

Total costs decrease by 2.5% with variable cost of 23.32 Euro/MWh—the lowest among all stochastic scenarios—but higher fixed costs. Without capacity substitution however, the value of storage halves to 2.1 Euro/kWh compared to the base case. This reinforces the observation that arbitrage alone is not an efficient usage of storage and that around half of its value stems from its precautionary operation.

#### 5.5. Comparison to Perfect Foresight Optimization Results

Finally, we compare the value of storage under perfect foresight of residual load (from 3.PF-M 0/0 and 6.PF-M 300/30) with its value under uncertainty (from 7.MPS 0/0 and 9.MPS 300/30). This allows us to estimate the impact of the stepwise information processing on the optimal storage strategy and capacity decisions.

Without storage the average of the perfect foresight results (3.PF-M 0/0) closely fit the stochastic optimization results (7.MPS 0/0). Installed capacity differs by less than 1 GW—aside from lost load with 2 GW.<sup>15</sup> Total costs differ by 0.1% only.

With storage (6.PF-M 300/30) generation capacity could be reduced to 60 GW—the same value achieved with limited foresight (9.MPS 300/30)—but furthermore lost load could be avoided

15. The results for 3.PF-M 0/0 give an average lost load of 3GW, suggesting that 5 GW of load is lost in 60% of our runs and none in the others. This is inevitably different from the 5 GW steps of the Markov process.

completely. This is achieved by a more ‘case specific’ storage operation that only reserves storage at times when high residual loads actually follows. This gives a more equal distribution of SOCs (shown in Figure 5). Table 8 shows that there is also a small probability of using more than 20 GW of un-/loading capacity to better match storage to the pattern of load. Storage allows the industry to take further advantage of lower-cost generators, with more nuclear capacity (and Combined Cycle) in 6.PF-M 300/30 than in 9.MPS 300/30, and less coal capacity. This improved storage handling and the capacity adjustments increase the storage value from 4.1 Euro/kWh under uncertainty to 5.6 Euro/kWh with perfect foresight. Models with too much foresight are likely to significantly over-estimate the value of electricity storage.

## 6. CONCLUSION

We analyse the optimal operation of storage and generation, and generation capacity choices, when the load net of wind and solar output is uncertain. This uncertainty is modelled as a Markov process, and the operating decisions are the solution to a Markov Decision Problem. Our algorithm chooses the least-cost generation capacity mix for the resulting duration curve of generator outputs, and iterates to a revised Markov Decision Problem until the solution converges.

This model is quantified with residual load (net of wind and solar generation) modelled as a diurnal Markov process from the residual load transitions in Germany in 2011 to 2015, with contemporary cost data for non-intermittent generation technologies and 300 GWh of storage (equivalent to six hours of average residual load). In this “real world” case, storage substituted for 15 GW of peak generation capacity. Uncertainty over the future evolution of demand made it optimal to keep the storage at a high charge level for most of the time to minimise the risk of a series of high-demand periods leading to the loss of load—we call this precautionary storage. Models with perfect foresight (or perfect foresight within a scenario) allow for more arbitrage and exaggerate the value of storage—our central estimate for the Markov model gives a value of €4.1/kWh of storage capacity, 27% below the perfect foresight estimate of €5.6/kWh. It is easier to forecast loads than we assume, which would tend to reduce the difference, but we also make no allowance for generation outages, which might increase it. Over the range of values that we simulated, we found that the value of storage is very sensitive to its power (GWh) capacity, but relatively insensitive to its energy (GW) capacity for charging and discharging.

Will current electricity market designs give the right incentives for storage? The “European” paradigm of an energy-only market (Green, 2008) relies on the prospect of high peak prices to give incentives for investment in generation capacity. As noted by an anonymous referee, “precautionary” storage can be seen as arbitrage against the occasional risk of extreme prices when load might be lost. Under the energy-only paradigm, storage operators would calculate the expected profits to be made from precautionary and from normal arbitrage, and maintain a suitably high level of charge.

Several of the electricity markets in the US include a capacity market to provide stronger incentives. The fear is that an energy-only market would lead to “missing money” as peak prices rarely rise to the levels necessary to recover generators’ full costs. Various designs of capacity market provide another revenue stream, and several European countries (amongst others) are also providing extra support. Storage operators are already allowed to participate in the British capacity market, and their power capacity is de-rated to reflect the risk that they might run out of energy. While the capacity market rules include penalties for not delivering electricity when it is needed, would these be strong enough to guarantee the right amount of precautionary storage if short-term

arbitrage opportunities are available? One possibility comes from the “strategic reserve” model in which a few generators are held out of the market by the system operator until the loss of load looks likely; an equivalent for storage could require the state of charge to be kept at or above our “threshold” level unless specific permission was given.

The modelling in this paper could be extended in various ways. It would be straightforward, but would considerably extend solution times and storage requirements, to use a finer resolution when discretising the load data (e.g. 2.5 GW rather than 5 GW steps). Second, it would be desirable to introduce seasonal patterns of demand alongside the daily patterns of time inhomogeneity presently included in the model. One risk facing countries with a high proportion of wind generation is that a long-lived “cold and calm” weather pattern of low wind speeds would deplete energy storage and lead to power cuts. Our Markov chain approach can create this kind of (residual) demand pattern, but allowing the transition matrices to vary with the season might represent this risk more accurately.

Third, the Markov process based on observed residual load transitions actually understates the knowledge of the decision makers. Forecasts of wind speed, in particular, have reached a significant level of predictability. There is thus more—but not infinite—foresight in the real world than we have considered here. In follow-up work (Geske and Green, 2018), we propose a way to deal with this problem in an accurate manner based on the modelling approach in this paper.

## ACKNOWLEDGMENTS

Funding from the Engineering and Physical Science Research Council via project EP/L014386/1, Business, Economics, Planning and Policy for Energy Storage in Low-Carbon Futures, is gratefully acknowledged. We would like to thank the editor and three anonymous referees for helpful comments.

## REFERENCES

- Bauerle, N. and U. Rieder (2011). *Markov Decision Processes with Applications to Finance*. Berlin: Springer. <https://doi.org/10.1007/978-3-642-18324-9>.
- Bertsekas, D.P. (2001). *Dynamic Programming and Optimal Control, Vol. I*. Belmont, MA.: Athena Scientific.
- Bertsekas, D.P. (2005). *Dynamic Programming and Optimal Control, Vol. II*. Belmont, MA.: Athena Scientific.
- Brokish, K. and J. Kirtley (2009). “Pitfalls of Modeling Wind Power Using Markov Chains.” Power Systems Conference and Exposition, 2009. PSCE’09. <https://doi.org/10.1109/PSCE.2009.4840265>.
- Cai, Y., K.L. Judd, T.S. Lontzek, V. Michelangeli, and C.-L. Su (2017). “A Nonlinear Programming Method for Dynamic Programming.” *Macroeconomic Dynamics* 21: 336–361. <https://doi.org/10.1017/S1365100515000528>.
- Cruise, J. and S. Zachary (2015). “The Optimal Control of Storage for Arbitrage and Buffering, with Energy Applications.” arXiv:1509.05788. <https://doi.org/10.1109/CISS.2014.6814090>.
- Durmaz, T. (2016). “Precautionary Storage in Electricity Markets.” NHH Dept. of Business and Management Science Discussion Paper No. 2016/5. <https://doi.org/10.2139/ssrn.2738004>.
- EnerNex Corporation (2006). “2006 Minnesota Wind Integration Study.” Final Report, Technical report, <https://www.leg.state.mn.us/docs/2007/mandated/070001/volumeI.pdf>.
- Ess, F., L. Haefke, J. Hobohm, F. Peter and M. Wunsch (2012). “Bedeutung der internationalen Wasserkraft-Speicherung für die Energiewende.” Prognos AG, Studie für den Weltenergieat—Deutschland e.V.
- Feinberg, E.A. and A. Shwartz (2002). *Handbook of Markov Decision Processes*. Boston, MA: Kluwer Academic Publishers. <https://doi.org/10.1007/978-1-4615-0805-2>.
- Geske, J. and R.J. Green (2018). “Optimal storage in a renewable system—Ignoring renewable forecast is not a good idea!” Mimeo, Imperial College Business School.
- Green, R.J. (2008). “Electricity Wholesale Markets: Designs Now and in a Low-carbon Future.” *The Energy Journal* 29(Special Issue): 95-124. <https://doi.org/10.5547/ISSN0195-6574-EJ-Vol29-NoSI2-6>.

- Growthsch, C., R. Malischek, S. Nick and H. Wetzel (2013). “The Costs of Power Interruptions in Germany—an assessment in the light of the Energiewende.” EWI Working Paper No 13/07.
- Hernandez-Lerma, O. and J.B. Lasserre (1996). *Discrete-time Markov Control Processes*. New York: Springer-Verlag. <https://doi.org/10.1007/978-1-4612-0729-0>.
- Kallrath, J., P.M. Pardalos, and M. Scheidt (Eds.) (2009). *Optimization in the Energy Industry*. Berlin: Springer. <https://doi.org/10.1007/978-3-540-88965-6>.
- Lauschke, A. (2011). “Affine-Scaling Interior Point Method.” Wolfram Demonstrations Project. <http://demonstrations.wolfram.com/AffineScalingInteriorPointMethod/> (downloaded 8 March 2019).
- London Economics International LLC (2013). “The Value of Lost Load (VoLL) for Electricity in Great Britain.” Final Report for OFGEM and DECC.
- Möst, D. and D. Keles (2010). “A survey of stochastic modelling approaches for liberalised electricity markets.” *European Journal of Operational Research* 207(2): 543–556. <https://doi.org/10.1016/j.ejor.2009.11.007>.
- Niesen, U., D. Shah, and G.W. Wornell (2009). “Adaptive Alternating Minimization Algorithms.” *IEEE Transactions on Information Theory* 55(3): 1423–9. <https://doi.org/10.1109/TIT.2008.2011442>.
- O’Sullivan J.A., (1998). “Alternating Minimization Algorithms: From Blahut-Arimoto to Expectation-Maximization.” In: Vardy A. (ed.) *Codes, Curves, and Signals*. The Springer International Series in Engineering and Computer Science (Communications and Information Theory), vol. 485. Boston, MA: Springer. [https://doi.org/10.1007/978-1-4615-5121-8\\_13](https://doi.org/10.1007/978-1-4615-5121-8_13).
- Powell, W.B., A. George, H. Simão, W. Scott, A. Lamont and J. Stewart (2012). “SMART: A Stochastic Multiscale Model for the Analysis of Energy Resources, Technology, and Policy.” *INFORMS Journal on Computing* 24(4): 665–682. <https://doi.org/10.1287/ijoc.1110.0470>.
- Puterman, M.L. (1994). *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. New York: John Wiley & Sons. <https://doi.org/10.1002/9780470316887>.
- Qiu, J. and A.A. Girgis (1993). “Optimization of power-system reliability level by stochastic-programming.” *Electric Power System Research* 26(2): 87–95. [https://doi.org/10.1016/0378-7796\(93\)90021-6](https://doi.org/10.1016/0378-7796(93)90021-6).
- Rebennack, S., and J. Kallrath (2017). *Stochastic Optimization in Energy Systems—With Practical Implementations in GAMS*. Berlin: Springer.
- Ross, S.M. (1983). *Introduction to Stochastic Dynamic Programming*. New York, NY: Academic Press. <https://doi.org/10.1016/B978-0-12-598420-1.50006-4>.
- Schröder, A., F. Kunz, J. Meiss, R. Mendelevitch, and C. von Hirschhausen (2013). “Current and Prospective Costs of Electricity Generation until 2050.” Data Documentation 68, Berlin: Deutsches Institut für Wirtschaftsforschung.
- Shamshad, A., M.A. Bawadi, W.M.A., Wan Hussin, T.A. Majid, and S.A.M. Sanusi (2003). “First and Second Order Markov Chain Models for Synthetic Generation of Wind Speed Time Series.” *Energy* 30: 693–708. <https://doi.org/10.1016/j.energy.2004.05.026>.
- Simonsen, R., B. Weron, and B. Mo (2004). “Structure and Stylized Facts of a Deregulated Power Market.” MPRA paper 1443, University of Munich, Germany.
- Sioshansi R, P. Denholm, T. Jenkin, and J. Weiss (2009). “Estimating the Value of Electricity Storage in PJM: Arbitrage and Some Welfare Effects.” *Energy Economics* 31(2):269–277. <https://doi.org/10.1016/j.eneco.2008.10.005>.
- Staffell I., and S. Pfenninger (2018). “The Increasing Impact of Weather on Electricity Supply and Demand.” *Energy* 145: 65–78. <https://doi.org/10.1016/j.energy.2017.12.051>.
- Teng, F., J. Miles, A. Thomson, G. Strbac, N. Brandon, and D. Pudjianto (2015). “Potential Value of Energy Storage in the UK Electricity System.” *Proceedings of the ICE—Energy* 168(2): 107–117. <https://doi.org/10.1680/ener.14.00033>.
- Tidball, R., J. Bluestein, N. Rodriguez, and S. Knoke (2010). “Cost and Performance Assumptions for Modeling Electricity Generation Technologies.” Golden, Colorado: NREL. <https://doi.org/10.2172/1219277>.
- Tijms, H.C. (2003). *A First Course in Stochastic Models*. Chichester: John Wiley & Sons. <https://doi.org/10.1002/047001363X>.
- Trick, M.A. and S.E. Zin (1993). “A Linear Programming Approach to Stochastic Dynamic Programming.” working paper, Carnegie Mellon University.
- Trick, M.A. and S.E. Zin (1997). “Spline Approximations to Value Functions—Linear Programming Approach.” *Macroeconomic Dynamics* 1: 255–277. <https://doi.org/10.1017/S1365100597002095>.
- Wallace, S.W. and S.-E. Fleten (2003). “Stochastic Programming Models in Energy.” In A. Ruszczyński and A. Shapiro, eds., *Handbooks in OR & MS*, 10: 637–77. [https://doi.org/10.1016/S0927-0507\(03\)10010-2](https://doi.org/10.1016/S0927-0507(03)10010-2).
- Weisser, H. (2007). “The Security of Gas Supply—a Critical Issue for Europe?” *Energy Policy* 35: 1–5. <https://doi.org/10.1016/j.enpol.2005.10.002>.
- Zucker, A., T. Hincliffe, and A. Spisto (2013). “Assessing Storage Value in Electricity Markets—A literature review.” Brussels: European Commission.

**APPENDIX: DETAILS OF THE NUMERICAL SOLUTION**

The nonlinear optimization problem (3) has a 2-block structure in capacities and the storage strategy. In these cases, “the alternating minimization or projection algorithm has been extensively used ...”.<sup>16</sup> It consists of the division of an optimization problem into a sequence of two simpler ones. The limit of the optimal strategies are stationary points of the sequence. Sufficient uniqueness conditions for the optimality of the limit of the sequence can be provided. With capacities  $k$  and the storage strategy  $\pi$  the algorithm that can be applied to solve (3) is:

1. Choose initial capacities  $k_0 \in K$ .
2. Determine the optimal storage strategy  $\pi_n$  given capacities  $k_{n-1}$  and the optimal capacity response  $k_n$  to storage strategy  $\pi_{n-1}$

$$\pi_n \in \arg \min_{\pi \in \Pi} C(\pi, k_{n-1}) \tag{10}$$

$$k_n \in \arg \min_{k \in K} C(\pi_{n-1}, k) \tag{11}$$

3. If the series  $(\pi_n, k_n)_{n \geq 1}$  converges the limit is a solution candidate to the (3).

Problem (10) is a Markov decision process and (11) a long-term capacity optimization problem. For both conditions of existence and solutions are available (e.g. for the MDP: Ross, 1983). In detail, the appeal of the decomposition to the nonlinear joint strategy and capacity optimization problem<sup>17</sup> (3) is that (10) can be discretized and solved as a linear problem and (11) can be solved directly via the screening curve approach. That means given the optimal storage strategy  $\pi_{n-1}$  an expected residual load duration  $RLD[\pi_{n-1}]$  can be derived and from this residual load duration, the optimal capacities  $k_n = \text{Screening}(RLD[\pi_{n-1}])$ .

For the solution of the MDP (10) several numerical options are available. We favour the seldom-applied linear programming approach<sup>18</sup> as the solution of (11) via the screening curve approach requires the post-storage expected load-duration, that is the stationary probabilities of the state of charge and residual load. The latter can be derived directly from the solution of the LP. So, the iterated cumbersome discretization of the optimal policy with respect to the state space and the (probably) more difficult and time-consuming computation of stationary probabilities of iteration schemes (e.g. policy iteration) can be avoided. These advantages (in addition to robustness with respect to step shape of the marginal costs and the simplicity to integrate storage capacity constraints) come at the cost of a problem size that grows exponentially as the discretization is refined. Let the optimal strategy derived by the LP be  $\pi_n = LP(k_n)$ . Details of this solution can be found in the next section.

To determine the optimal capacity we started with an equal share capacity equipment  $k_0$  whose elements total to peak load. With this capacity equipment we determined the optimal storage strategy  $\pi_0 = LP(k_0)$ . The next stage capacity equipment  $k_1 = \text{Screening}(RLD[\pi_0])$  converged, to a closed-by suboptimal steady state. To avoid this it proved sufficient to analyse total costs of all neighbouring capacities of  $k_1$ . To be specific the neighbourhood capacities  $N$  of a capacity  $k_1 = (k_1, \dots, k_i, k_{i+1}, \dots, k_J)$  are the elements of the set

16. Niesen, Shah and Wornell, 2009; overview: O’Sullivan, 1998

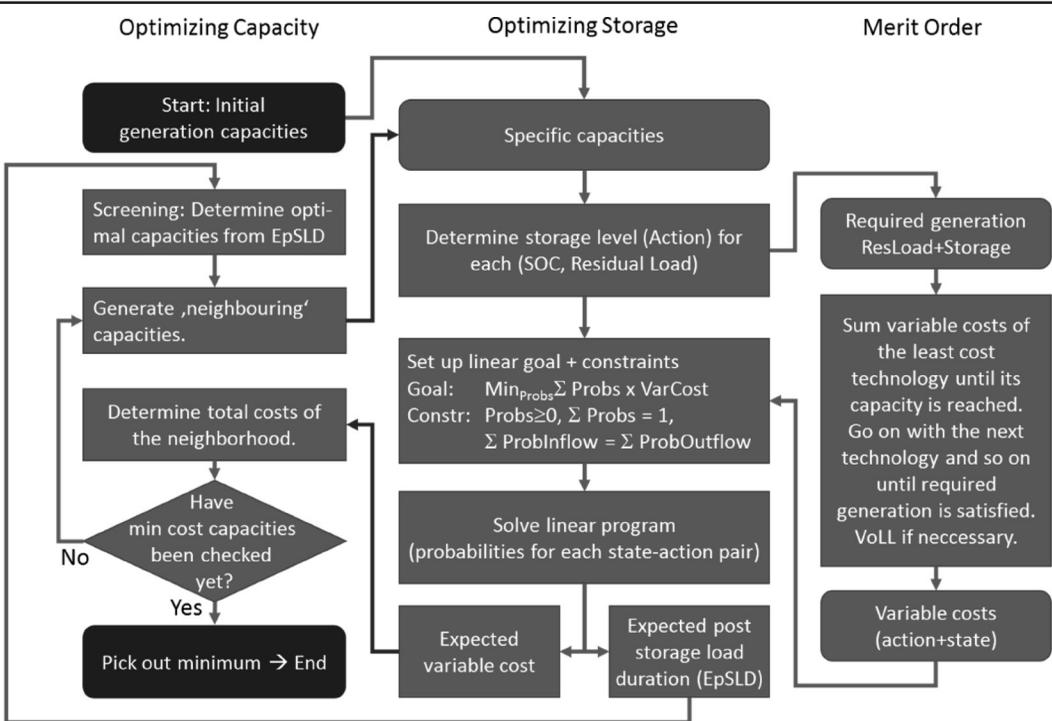
17. For a direct approach see Cai et al 2017.

18. One of the few applications of the LP approach to stochastic optimization problems can be found in Trick and Zinn 1993 and 1997.

$$N(k_i) = \{k : (\dots, k_i + \varepsilon, k_{i+1} - \varepsilon, \dots) \mid i = 1, \dots, I - 1, \varepsilon \in \{+\mu, 0, -\mu\}\}$$

with the smallest capacity unit  $\mu$  (which was 5 GW throughout). We then solved  $\min_{k \in N(k_i)} C(LP(k), k)$ .<sup>19</sup> To guarantee optimality, the global optimum was determined by brute force for a few early cases to understand the problem. The algorithm is presented as a flow diagram in Figure 10.

**Figure 10: Flow Diagram of the Algorithm to Solve the Stochastic Optimization Problem.**



**Linear Programming solution of stationary Markov decision processes:**

With the discrete states  $\{1, \dots, |\hat{S}|\}$  in 5 GWh steps and demand  $\{1, \dots, |D|\}$  in 5 GW steps the set of admissible actions  $A$  in a state  $(S, D)$  are

$$A((S, D) \mid k, \hat{S}) = \{s \in \mathbb{N} \mid 1 \leq s + S \leq \hat{S}, 1 \leq s + D \leq |D|\} \tag{13}$$

We need a decision rule to choose the probability of choosing each alternative action  $s \in A(S, D)$  in state  $(S, D)$ , which we denote as  $d((S, D), s)$ . Applied to the storage problem, (Ross, 1983) proves that from the solution of the linear program:

$$\max_{d((S, D), s)} \sum_{(S, D), s \in A(S, D)} d((S, D), s) C(D + s \mid k) \tag{14}$$

19. The non-convexity is most probably caused by the discretization of the problem, as perturbations of the storage strategy may only pay off if the capacity response is sufficient and capacities need to be fully utilized—not possible if capacities are adjusted by screening.

$$\sum_{(S,D),s \in A(S,D)} d((S,D),s) = 1 \tag{15}$$

$$d((S,D),s) \geq 0 \forall S,D \tag{16}$$

$$\sum_{s \in A(S,D)} d((S,D),s) = \sum_{S',D',s \in A(S',D')} d((S',D'),s) Prob((S',D'),(S,D),s) \forall S,D \tag{17}$$

$$Prob((S',D'),(S,D),s) = \begin{cases} P(D'|D) & S' = s + S \\ 0 & \text{otherwise} \end{cases} \tag{18}$$

the decision rule can be derived as follows:

$$Prob(s|(S,D)) = \frac{d((S,D),s)}{\sum_{s' \in A(S,D)} d((S,D),s')} \tag{19}$$

MDPs (discretized as we use them in the LP approach) can be interpreted as minimum-cost flow problems (MCFP). Their states correspond to the nodes of an MCFP, the discrete actions to pipelines directing a flow of probabilities respectively a liquid, that is conserved at each state (node). The actions are selected as pipes to minimize the costs of their usage. This interpretation enables the graphical representation of the MDP as pipeline system that is exploited in section 4.<sup>20</sup>

The LP framework has a further advantage. While it does not allow for a straightforward deviation of the optimal strategy, the determination of the load duration given a strategy requires only the computation of the stationary probabilities of a Markov Chain. It is thus straightforward to analyse variations of elements of the optimal strategy separately and quantify contributions to the optimal solution—for specific cases. This feature greatly deepens the understanding of the optimal strategy.<sup>21</sup>

The algorithm was implemented with Mathematica. The code is available as an online appendix.

20. We used Wolfram’s Affine-Scaling Interior Point method, which is “on par with several commercial network simplex algorithms for multi-commodity network problems” (Lauschke, 2011). A solution of the LP with a network simplex algorithm—known to be the fastest approach for this set of problems—was unnecessary for a problem of our difficulty.

21. It might be considered as an intuitive substitute to the analytical analysis of first order conditions in a Bellman equation based approach. The latter is generally not suited to give any insights into the structure of the stationary probabilities and hence the expected load duration.



The IAEE is pleased to announce that our leading publications exhibited strong performances in the latest 2018 Impact Factors as reported by Clarivate. The *Energy Journal* achieved an Impact Factor of 2.456 while *Economics of Energy & Environmental Policy* saw an increase to 2.034.

Both publications have earned SCIMago Journal Ratings in the top quartile for Economics and Econometrics publications.

IAEE wishes to congratulate and thank all those involved including authors, editors, peer-reviewers, the editorial boards of both publications, and to you, our readers and researchers, for your invaluable contributions in making 2018 a strong year. We count on your continued support and future submission of papers to these leading publications.