

# **Coaction of Input Parameters and Model Sensitivities in Numerical Power System Modeling**

Von der Fakultät für Mathematik und Naturwissenschaften der Carl  
von Ossietzky Universität Oldenburg zur Erlangung des Grades eines

Doctor rerum naturalium

angenommene Dissertation  
von Herrn Bruno Umberto Schyska  
geboren am 19.10.1986 in Berlin

Gutachter: Prof. Dr. Joachim Peinke

Weiterer Gutachter: Prof. Dr. Alexander Hartmann

Tag der Disputation: 17.02.2021

To Ante, Beeke, Tue and Ove

Ich habe einen Kopf  
und zwei Augen.  
Ich kann seh'n, sprechen, denken und glauben.  
Ich hab einen Namen und ein Gesicht.  
Ich sag "Hallo, Hurra, hier bin ich".

[aus dem Lied "Hier bin ich" von Gisbert zu Knyphausen und Nils Koppruch]

Oldenburg, January 5, 2022  
*Bruno Umberto Schyska*

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Basics of Power System Modeling</b>	<b>3</b>
2.1	Mathematical Formulation . . . . .	3
2.2	The Numerical Complexity of Power System Expansion Models . . . . .	5
2.3	Duality and Locational Marginal Pricing . . . . .	7
2.4	Data, Models and Software . . . . .	10
2.5	Coacting Parameters . . . . .	12
<b>3</b>	<b>Research Question 1</b>	<b>17</b>
3.1	Aims and Motivation . . . . .	17
3.2	Methodology and Data . . . . .	18
3.2.1	Reference Site Selection: Spatio-Temporal Clustering . . . . .	18
3.2.2	Upscaling and Evaluation . . . . .	20
3.2.3	Wind Power Generation Data . . . . .	21
3.2.4	Circulation Weather Types . . . . .	22
3.3	Results . . . . .	23
3.3.1	Cluster Centres and Reference Site Weights . . . . .	23
3.3.2	Upscaling Evaluation . . . . .	24
3.3.3	Benefit from Training for Weather Types . . . . .	26
3.4	Discussion and Conclusions . . . . .	27
<b>4</b>	<b>Research Question 2</b>	<b>29</b>
4.1	Aims and Motivation . . . . .	29
4.2	Data and Methods . . . . .	31
4.2.1	Power System Expansion Modelling . . . . .	31
4.2.2	Locational Marginal Prices for Electricity . . . . .	32
4.2.3	Cost of Capital Scenarios . . . . .	33
4.2.4	Generation and Load Data . . . . .	35
4.3	Results . . . . .	35
4.4	Discussion . . . . .	41
4.5	Conclusion . . . . .	43
<b>5</b>	<b>Research Question 3</b>	<b>45</b>
5.1	Aims and Motivation . . . . .	45
5.2	Defining a misallocation metric . . . . .	46
5.3	Results . . . . .	48
5.3.1	The Sensitivity to the Capital Cost of Generation Capacities . . . . .	48

5.3.2	The Sensitivity to the Capacity Factor Time Series . . . . .	49
5.3.3	The Sensitivity to Reduced Spatial and Temporal Resolution . . . . .	53
5.4	Discussion . . . . .	57
5.5	Conclusions . . . . .	58
5.6	Experimental Procedures . . . . .	59
5.6.1	Power System Model and Data . . . . .	59
5.6.2	Reducing the Spatial and Temporal Resolution of PSEM . . . . .	60
5.6.3	Computing the misallocation metric . . . . .	61
<b>6</b>	<b>Extending Research Questions and Outlook</b>	<b>63</b>
6.1	Correlation Lengths of Wind and Solar Power . . . . .	63
6.2	Tracing the Flow of Electricity . . . . .	65
6.3	Parametric Optimisation . . . . .	67
6.3.1	Introduction . . . . .	68
6.3.2	Methodology . . . . .	68
6.3.3	Results . . . . .	70
6.3.4	Conclusion . . . . .	72
6.4	Stochastic Optimisation . . . . .	72
6.5	Outlook . . . . .	76
<b>7</b>	<b>Summary and Conclusions</b>	<b>77</b>
	<b>References</b>	<b>79</b>
	<b>Appendix</b>	<b>89</b>

# Nomenclature

## Abbreviations

LCOE	Levelized Cost of Electricity	[€/ MWh]
Li-ion	Lithium-Ion (battery)	
LMPE	Locational Marginal Prices of Electricity /MWh	[€]
OCGT	Open Cycle Gas Turbine	
PSEM	power system expansion model(s)	
ROR	runoff river	

## Decision Variables

$\bar{f}_l$	nominal transmission capacity of transmission line $l$	[MWkm]
$\bar{g}_{n,s}$	nominal generation capacity of generation source $s$ at bus $n$	[MW]
$f_{l,t}$	electricity transmission via transmission line $l$ at time $t$	[MWh]
$g_{n,s,t}$	electricity generation of resource $s$ at bus $n$ and time $t$	[MW]
$s_{n,t}$	load shedding at bus $n$ and time $t$	[MWh]
$soC_{n,s,t}$	state of charge of storage unit of category $s$ at bus $n$	[MWh]

## Indices

$l$	transmission line $\in \mathcal{L}$
$n$	bus $\in \mathcal{N}$
$s$	resource $\in \mathcal{S}$
$t$	time step

## Parameters

$\bar{G}_{n,s}^{min/max}, \bar{F}_l^{min/max}$	upper/lower bounds for nominal transmission and generation capacities	[MW]
$\tilde{G}_{n,s,t}$	capacity factor of resource $s$ at bus $n$ and time $t$	[a.u.]
$\varepsilon_{n,s}, \eta_{n,s}$	efficiencies	[a.u.]

## Contents

---

$C_l$	annualised investment cost of transmission line $l$ at bus $n$	[€/ MWkm]
$C_{n,s}$	annualised investment cost of resource $s$ at bus $n$	[€/ MW]
$CAP_{CO_2}$	global CO <sub>2</sub> emission limit	[tons]
$CAP_{trans}$	global transmission capacity limit	[MWkm]
$D_{n,t}$	electricity demand at bus $n$ and time $t$	[MWh]
$E_{n,s}$	CO <sub>2</sub> emissions of resource $s$ at bus $n$	[tons / MWh]
$inflow_{n,s}$	natural inflow to storage unit $s$ at bus $n$	[MWh]
$K_{n,l}$	incidence matrix	[]
$L_l$	length of transmission line $l$	[km]
$O_{n,s}$	marginal cost of resource $s$ at bus $n$	[€/ MWh]
$r_{n,s}$	power-to-energy ratio of storage units	[hours]
$spill_{n,s}$	spillage of natural inflow to storage unit $s$ at bus $n$	[MWh]
$W_t$	snapshot weighting	[a.u.]

### Sets

$\mathcal{L}$	Transmission lines of the power system
$\mathcal{N}$	Buses of the power system
$\mathcal{S}$	Resources available for expansion

# Chapter 1

## Introduction

In order to address the issue of climate change and sustainability, countries around the world make great efforts to transform their power systems. In this context, large shares of renewable power sources need to be integrated into existing power systems [Pacala and Socolow, 2004]. This is a challenging task which requires a great number of political decisions and great financial and technical efforts. It is mainly the intermittency of the renewable resources, i.e. their *weather-dependence*, which causes difficulties for the operation of power systems and electricity markets [Pinson et al., 2007, Morales et al., 2013]. Basically, it makes it more difficult to ramp up electricity generation when the demand is high and to shut down plants when the demand is low. Renewable power plants are only partially committable. The need to account for the turbulent character of wind energy has for instance been stressed by Milan et al. [2013]. Haehne et al. [2019] showed that non-Gaussian time-correlated fluctuations in the renewable feed-in perturb electricity grids on a sub-second time scale and Pineda et al. [2016] showed that errors in wind power forecasts even affect the optimal expansion of a power system.

Solutions to cope with these issues proposed by research are manifold [Lund, 2007, Connolly et al., 2010]. Some of them are relatively straight-forward as, for instance, the large-scale integration of storage technologies [Steinke et al., 2013, Weitemeyer et al., 2016], which can store electricity in times of over-generation and dispatch it during times of low resource availability, the extension of the transmission grid [Rodriguez et al., 2014, Kies et al., 2016e] in order to distribute generation over large geographical areas and to smooth out fluctuations (see also Section 2.5), or the over-installation of renewable capacities to reduce the need for backup energy [Heide et al., 2011]. Others focus on the meteorological aspects and propose to use meteorological information to find the optimal spatial deployment of generation capacities [Grams et al., 2017] or to optimise the mix of different renewable generation sources [Ming et al., 2017, Jurasz and Ciapała, 2017, Santos-Alamillos et al., 2015]. Focusing on the flexibility options on the demand-side, Kies et al. [2016f] found that demand-side management can balance generation-side fluctuations for a renewable share of up to 65% in Europe. Even more flexibilities could be used by coupling the different sectors of the energy system, i.e. electricity, heat and transportation. The synergies between these sectors have been investigated by Brown et al. [2018]. Lund and Kempton [2008] proved the usefulness of vehicle-to-grid technologies for the integration of renewables. Hirth and Müller [2016], Chattopadhyay et al. [2017] and Pfenninger et al. [2014] proposed to deploy *system-friendly* generation assets which are able to deliver ancillary services. Accordingly, this would ease the operation of power systems with high shares of renewables and, hence, might support the integration of these new resources. A new market scheme which explicitly

accounts for the uncertainty in renewable power forecasts has been proposed by Morales et al. [2014].

All of these proposed solutions come with their own specific characteristics, their specific advantages and weaknesses and at their own specific cost. Making a decision about which of them is suited best and which should consequently be envisaged with priority is challenging. Commonly, power system models are the method of choice to support this process of decision-making. As power system models often are complex linear or mixed-integer programs, the (optimal) solution of these programs depends on a great number of parameters from different fields: economics, electrical engineering, meteorology and geography. The interplay of these different aspects, the way they *coact* towards the optimisation function, is complex, the sensitivity of power system models to this interplay and to different choices of the model parameters not yet fully explored.

The aim of this thesis is the description of the complex coaction of the input parameters in numerical power system models and the quantification of the power system models' sensitivities on different choices of model parameters and model designs. In particular, I will address the following research questions: (i) How is the relation between weather classes and wind power production?, (ii) How do regional differences in cost of capital influence the optimal design of power systems? and (iii) How large is the sensitivity of power system expansion models? I will show that (i) the availability of wind power can clearly be linked to the prevailing weather situation, that (ii) there is a need to consider regional differences in the cost of capital in defining the optimal design of a highly renewable European power system and that (iii) power system models are especially sensitive to the choice of the underlying weather data and the temporal resolution. In Section 6, more examples for the concept of parameter coaction will be given by investigating the correlation length in wind and solar power, the potential use of a flow tracing algorithm to allocate cost for grid expansion, the possible evaluation of forecast skill via a stochastic dispatch model and the use of parametric optimisation for reducing the model complexity.

# Chapter 2

## Basics of Power System Modeling

### 2.1 Mathematical Formulation

Throughout this thesis, a common formulation of a power system expansion model (PSEM) will be used. It has the aim to find the least expensive design of a power system given the constraint that the CO<sub>2</sub> emissions from power plants may not exceed an upper limit. In the formulation used here, it consists of two parts (compare Section 2.2): While the first part minimises the investment cost in generation, storage and transmission capacities keeping all capacities within given bounds, the second part minimises the operational cost of the power system ensuring the security of supply and keeping the generation lower or equal the capacities derived from the upper level. For volatile renewable resources – such as wind and solar PV – the available dispatchable capacity is additionally limited by prevailing meteorological conditions, i.e. the capacity factor time series. Furthermore, the consistency of the state of charge of storage units must be ensured.

Mathematically, the expansion problem can be formulated as a linear problem: Let us consider a power system described by a set of buses  $\mathcal{N}$ . Assume that these buses  $n \in \mathcal{N}$  are connected by lines denoted with the index  $l \in \mathcal{L}$ . Each bus is assigned with an inelastic electricity demand (or load) time series  $D_{n,t}$ . This load can be served by electricity generation from a set of generators and storage units installed at the respective bus  $\mathcal{S}_n$ . The generation itself may be connected with costs, the operational (or marginal) costs of the generator or storage  $O_{n,s}$  as well as CO<sub>2</sub> emissions  $E_{n,s}$ . Note that we assume time-independent marginal costs here. Furthermore, the generation is limited by the installed nominal power of the generator or storage  $\bar{g}_{n,s}$  and by the availability of the power source, which may vary in time  $\tilde{G}_{n,s,t}$ . For variable renewable sources  $\tilde{G}_{n,s,t}$  given in units of the nominal power  $\bar{g}_{n,s}$  is determined by the meteorological conditions, while for conventional generation technologies and storage units  $\tilde{G}_{n,s,t}$  is either constant or reflects downtimes due to maintenance and/or fuel shortages. Furthermore, we assume that the nominal power can be expanded. The related investment costs are denoted with  $C_{n,s}$ .

In order to be able to serve the electricity demand, the electricity must be distributed between the buses, where it is generated, and the buses, where it is consumed. The task of distributing electricity is accomplished by transmission lines. Similar to the generators, the lines are assigned with a nominal capacity  $\bar{f}_l$  and investment costs  $C_l$ . Let us assume that transmitting electricity is not related to any marginal costs. Denote the flow along line  $l$  with  $f_{l,t}$ .

Beyond the supply of electricity, storage units additionally fulfill the task of balancing temporal fluctuations in the availability of power by providing storage capacity expressed as

multiples of their nominal power. The factor has the unit of hours and can be interpreted as the number of hours the storage can supply electricity at full power. It is denoted with  $R_{n,s}$ . In contrast to generators, storage units cannot only supply electricity but also take energy up.  $g_{n,s,t}$  can also be negative. Each of these processes, i.e. storage dispatch  $d_{n,s,t}$  and storage uptake  $u_{n,s,t}$ , occur with a certain efficiency  $\eta_{n,s}^{dispatch}$  and  $\eta_{n,s}^{store}$  respectively. Additionally, there may be storage units exhibiting a natural inflow, e.g. hydro reservoir storage,  $inflow_{n,s,t}$ . Part of this inflow may also be spilled, which is described by the variable  $spill_{n,s,t}$ . All these variables are related by the state of charge, i.e. the amount of energy stored,  $soc_{n,s,t}$ .

Following these assumptions, the problem of minimising total system costs can mathematically be formulated as a linear optimisation problem of the following form

$$\min_{\bar{g}_{n,s}, g_{n,s,t}, \bar{f}_l, z_a} \left[ \sum_n \sum_s \left( C_{n,s} \bar{g}_{n,s} + \sum_t W_t O_{n,s} g_{n,s,t} \right) + \sum_l C_l \bar{f}_l \right] \quad (2.1)$$

Equation 2.1 reflects a rather simplified version of a power system expansion problem. It does not include binary variables originating from the consideration of unit commitment or of minimum/maximum down and up times for power generators (for instance). In this setup, the nominal power of generators and transmission lines  $\bar{g}_{n,s}$ ,  $\bar{f}_l$  and the electricity dispatch  $g_{n,s,t}$  are decision variables.  $z_a$  are auxiliary variables used to define the load flow. The network topology described by the incidence matrix  $K$ , electricity demand  $D_{n,t}$  and availability  $\tilde{G}_{n,s,t}$  enter as parameters (or *exogenous* data).  $W_t$  are optional snapshot weightings. Throughout this thesis only power system expansion problems of the same form as 2.1 will be considered.

The objective is subject to the following constraints:

1. At each bus  $n$  and at each time step  $t$  generation and exports must match the demand (**nodal balancing**):

$$\sum_s g_{n,s,t} - \sum_l K_{nl} f_{l,t} = D_{n,t}, \forall n, t \quad (2.2)$$

2. Generation may not exceed the available nominal power

$$g_{n,s,t} - \tilde{G}_{n,s,t} \cdot \bar{g}_{n,s} \leq 0 \quad (2.3)$$

3. The state of charge of the network's storage units is given by

$$soc_{n,s,t} = (1 - \eta_{n,s}^{stand})^{W_t} soc_{n,s,t-1} - W_t (1/\eta_{n,s}^{dispatch} d_{n,s,t} - \eta_{n,s}^{store} u_{n,s,t} - inflow_{n,s,t} + spill_{n,s,t}) \quad (2.4)$$

Here,  $\eta_{n,s}^{stand}$  describes the standing losses. In many cases an additional periodicity constraint is formulated, which fixes the state of charge at the first time step to the state of charge at the last time step, i.e.

$$soc_{n,s,t=0} = soc_{n,s,t=|\mathcal{T}|} \quad (2.5)$$

where  $\mathcal{T}$  is the set of time steps.

4. Global constraints:

- limit overall CO<sub>2</sub> emissions

$$\sum_{n,s,t} \frac{1}{\varepsilon_{n,s}} E_{n,s} g_{n,s,t} \leq CAP_{CO_2} \quad (2.6)$$

- limit overall transmission capacity

$$\sum_l L_l \bar{f}_l \leq CAP_{trans} \quad (2.7)$$

where  $L_l$  denotes the line length.

5. variable bounds

$$0 \leq g_{n,s,t} \quad (2.8)$$

$$\bar{G}_{n,s}^{min} \leq \bar{g}_{n,s} \leq \bar{G}_{n,s}^{max} \quad (2.9)$$

$$\bar{F}_l^{min} \leq \bar{f}_l \leq \bar{F}_l^{max} \quad (2.10)$$

$$0 \leq d_{n,s,t} \leq \bar{g}_{n,s} \quad (2.11)$$

$$0 \leq u_{n,s,t} \leq \bar{g}_{n,s} \quad (2.12)$$

$$0 \leq soc_{n,s,t} \leq r_{n,s} \bar{g}_{n,s} \quad (2.13)$$

$$0 \leq |f_{l,t}| \leq \bar{f}_l \quad (2.14)$$

6. The flow of electricity, i.e. active power, along any line can be expressed as a function of the net active power  $p_n$  and some auxiliary variables  $z_a$  [Hörsch et al., 2018b]:

$$f_l = f_l(p_n, z_a) \quad (2.15)$$

Note, that the choice of this function can significantly impact the computational cost for solving the optimisation problem. See the Appendix for a detailed description of load flow calculations.

This formulation is similar to the formulation of e.g. Brown et al. [2018], Schlachtberger et al. [2017], Schlott et al. [2018].

## 2.2 The Numerical Complexity of Power System Expansion Models

Beyond the objective function itself, the complexity of power system expansion problems and consequently the computational burden to solve them is mainly determined by the following factors: (i) the number of decision variables, (ii) the number of constraints (iii) the type of the decision variables, (iv) the number of parameter values to be stored in memory, (v) the sparsity or density of the constraint matrix and (vi) the shape of the solution space close to the optimum (e.g. Hörsch et al. [2018b]). Obviously (i), (ii) and (iv) are not independent.

When the number of decision variables increases, usually also the number of constraints and parameter values increases. (i), (ii) and (iii) set the requirements for the numerical solver algorithm applied to solve the expansion problem. Different algorithms need to be applied depending on whether the problem is purely continuous or whether it includes binary or integer variables for instance. The number of parameter values, however, mainly defines the storage requirements, which in many cases might be the limiting factor.

In our case, the linear program 2.1 (plus constraints) can be considered as a bilevel optimisation problem [Conejo et al., 2016]. The upper level is the investment problem of a fictitious system planner, who needs to decide about his investments in generation and transmission facilities. The system planner bases his decision upon the lower level problem. This lower level represents the market clearing aiming on maximizing social welfare. As such it determines the price at which electricity is sold and consequently the profit of the generation facilities. Mathematically, the upper level is constrained by (i) the upper level constraints 2.7, 2.9, 2.10 and a row of lower-level problems, one for each time-step  $t$ , each being represented by an optimisation problem with corresponding constraints itself. The upper level cannot be solved without having solved the lower level.

Additional complexity comes from the appearance of so called *complicating variables* [ibid.]. These complicating variables prevent the lower level problems to be solvable independently from the upper level. In our case, the complicating variables are the generation and transmission capacities  $\bar{g}_{n,s}$ ,  $\bar{f}_l$ , i.e. the decision variables of the upper level. Without knowing the capacities, the market cannot be cleared. This becomes clear when looking at the *complicating constraints* 2.3 through 2.15. These constraints link the lower level variables to the decision variables of the upper level. Note, that without the complicating variables and the corresponding complicated constraints the lower level problems could be solved independently from the investment decision, i.e. the upper level.

In general, two types of solution approaches exist for this kind of bilevel optimisation problems [Conejo et al., 2016]:

1. direct solution: all variables, constraints and parameters are constructed and solved at once by applying a solution algorithm for linear problems (e.g. the simplex algorithm or the interior-point algorithm [Boyd and Vandenberghe, 2004a]).
2. approaches based on decomposition techniques: decompose the upper and lower level in order to be able to solve the lower level independently from the upper level (e.g. Benders' decomposition [ibid.] )

While the second approach usually increases the computation time, because the global solution is calculated in an iterative manner, the direct solution usually requires more memory. For this thesis, I will only apply the direct solution approach. But since I am using a commercial solver, decomposition techniques may still be used under the hood. Most simulations were performed with the open-source software toolbox *Python for Power System Analysis* (PyPSA) [Brown et al., 2017] and the commercial solver *Gurobi* on the high performance cluster facilities of the University of Oldenburg and the DLR Institute of Networked Energy Systems.

## 2.3 Duality and Locational Marginal Pricing

The power system expansion problem described above can be considered as a linear programming problem of the following general form:

$$\text{Min}_x f(x) \quad (2.16)$$

$$\text{s.t. } h(x) = 0 \quad (2.17)$$

$$g(x) \leq 0 \quad (2.18)$$

with  $f(x) = c^T x$  being the linear cost function,  $h(x)$  representing the equality constraints for the nodal balancing and the state-of-charge and  $g(x)$  representing the inequality constraints including the complicating constraints (2.3). From (2.16) - (2.18) the corresponding *Lagrangian* can be derived as:

$$\mathcal{L}(x, \lambda, \mu) = f(x) + \lambda^T h(x) + \mu^T g(x) \quad (2.19)$$

As shown in Boyd and Vandenberghe [2004b], the infimum of (2.19)

$$d(\lambda, \mu) = \inf_x \mathcal{L}(x, \lambda, \mu) \quad (2.20)$$

defines a lower bound to the optimal solution of the optimisation problem (2.16) - (2.18) depending on the two parameters  $\lambda$  and  $\mu$ . Consequently, an optimisation problem can be defined aiming at finding the best – meaning maximum – lower bound of the original optimisation problem [ibid.]. This optimisation problem is called the *dual*,  $d(\mu, \lambda)$  the *dual function* and the original optimisation problem (2.16) - (2.18) the *primal*.

Here, I show that in the context of power system expansion problems – more precisely in the context of market clearing problems – the  $\lambda^*$  corresponding to the nodal balancing (Equation 2.2) can be interpreted as locational marginal prices of electricity and that linear power system expansion problems represent an efficient cost-allocation mechanism. Let us for now neglect any global constraints and any constraints effecting the state of charge of

storage devices in our power system expansion problem. Then, the Lagrangian reads:

$$\begin{aligned}
 \mathcal{L} = & \sum_n \sum_s \left( C_{n,s} \bar{g}_{n,s} + \sum_t W_t O_{n,s} g_{n,s,t} \right) + \sum_l C_l \bar{f}_l + \\
 & \lambda_{n,t} \left( \sum_s g_{n,s,t} - D_{n,t} - \sum_l K_{n,l} f_{l,t} \right) + \\
 & \mu_1 (g_{n,s,t} - \tilde{G}_{n,s,t} \bar{g}_{n,s}) - \\
 & \mu_2 g_{n,s,t} + \\
 & \mu_3 (\bar{g}_{n,s} - \bar{G}_{n,s}^{min}) + \\
 & \mu_4 (\bar{G}_{n,s}^{min} - \bar{g}_{n,s}) + \\
 & \mu_5 (\bar{f}_l - \bar{F}_l^{min}) + \\
 & \mu_6 (\bar{F}_l^{min} - \bar{f}_l) + \\
 & \mu_7 (f_{l,t} - \bar{f}_l) - \\
 & \mu_8 (f_{l,t} - \bar{f}_l)
 \end{aligned} \tag{2.21}$$

Applying the necessary conditions for optimality, namely *primal feasibility*

$$h(x^*) = D_{n,t} - \sum_s g_{n,s,t}^* - \sum_l K_{n,l} f_{l,t}^* = 0, \forall n, t \tag{2.22}$$

*complementary slackness*

$$\mu^T g(x^*) = 0 \tag{2.23}$$

and the *Karush-Kuhn-Tucker conditions* for stationarity

$$\nabla f(x^*) + \lambda \nabla h(x^*) + \mu \nabla g(x^*) = 0 \tag{2.24}$$

(see Boyd and Vandenberghe [2004b]) we finally find:

$$\begin{aligned}
 \sum_{n,s} \left( C_{n,s} \bar{g}_{n,s}^* + \sum_t W_t O_{n,s} g_{n,s,t}^* \right) + \sum_l C_l \bar{f}_l^* &= \sum_{n,s} \lambda_{n,t}^* g_{n,s,t}^* + \sum_{l,n,t} \lambda_{n,t}^* K_{n,l} f_{l,t}^* \\
 &= \sum_{n,t} \lambda_{n,t}^* D_{n,t}
 \end{aligned} \tag{2.25}$$

All terms related to the inequality constraints vanish. From (2.25) several things can be observed. First,  $\lambda$  gives the change in the objective function value of the primal with a marginal increase in the right hand side of the corresponding constraint, here the demand  $D$  in the nodal balancing equation [Bazaraa et al., 2013]. If  $D$  increases by  $D'$ , the overall cost increase by  $\lambda D'$ . This fact is used to define so called *cuts* in the context of Benders Decomposition. These cuts iteratively define additional constraints to the upper level of the original problem, iteratively narrowing the solution space until the optimal solution is found [Conejo et al., 2016].

Second,  $\lambda$  has the unit of cost or of a price: € per MWh. It represents a so-called *shadow*

price, which in the context of power markets commonly is referred to as *locational marginal price of electricity* (LMPE). If consumers at node  $n$  pay this price  $\lambda_{n,t}$  for the electricity they consume, each electricity provider at the same node receives  $\lambda_{n,t}g_{n,t}$  and each transmission operator of a line  $l$  connected to  $n$  receives  $\lambda_{n,t}K_{n,l}f_{l,t}$ . Then, in summary over the whole time period considered each electricity provider and transmission operator exactly receives her cost. This is the economical equilibrium, commonly understood as *welfare maximization*.

Let us now additionally consider the global constraints (2.6) and (2.7). If these two constraints are binding, adding them to the optimisation program would lead to an increase in the optimal objective function value. They would force the system to expand capacities, which would not have been built in the 'pure' economic case. (2.6) and (2.7) can be considered *normative* constraints. Let us relate  $\mu_{CO_2}$  to (2.6) and  $\mu_{trans}$  to (2.7) and perform the same analysis as above. We find two additional terms added to the right-hand side of (2.25):

$$\begin{aligned} \sum_{n,s} (C_{n,s}\bar{g}_{n,s}^* + \sum_t w_t O_{n,s} g_{n,s,t}^*) + \sum_l C_l \bar{f}_l^* = \\ \sum_{n,t} \lambda_{n,t}^* D_{n,t} + \mu_{CO_2}^* \sum_{n,s,t} \frac{1}{\varepsilon_{n,s}} E_{n,s} g_{n,s,t}^* + \mu_{trans}^* \sum_l L_l \bar{f}_l^* \end{aligned} \quad (2.26)$$

Note, that  $\mu_{CO_2}$  and  $\mu_{trans}$  are lower equal zero. Increasing the corresponding right-hand sides relaxes the corresponding constraints and consequently potentially reduces costs. Furthermore, dual variables are only non-zero when the corresponding constraint is binding, meaning when the limit defined by the constraint is met (see complementary slackness, Equation (2.23)):

$$\mu_{CO_2} \begin{cases} = 0, & \text{if } \sum_{n,s,t} \frac{1}{\varepsilon_{n,s}} E_{n,s} g_{n,s,t} < CAP_{CO_2} \\ < 0, & \text{if } \sum_{n,s,t} \frac{1}{\varepsilon_{n,s}} E_{n,s} g_{n,s,t} = CAP_{CO_2} \end{cases} \quad (2.27)$$

$$\rightarrow \mu_{CO_2} \sum_{n,s,t} \frac{1}{\varepsilon_{n,s}} E_{n,s} g_{n,s,t} = \mu_{CO_2} CAP_{CO_2} \quad (2.28)$$

( $\mu_{trans}$  analogue). Hence, (2.26) can be reformulated to:

$$\begin{aligned} \sum_{n,s} (C_{n,s}\bar{g}_{n,s}^* + \sum_t w_t O_{n,s} g_{n,s,t}^*) + \sum_l C_l \bar{f}_l^* + \\ |\mu_{CO_2}^*| CAP_{CO_2} + |\mu_{trans}^*| CAP_{trans} = \sum_{n,t} \lambda_{n,t}^* D_{n,t} \end{aligned} \quad (2.29)$$

Consequently, invoking global constraints implicitly increases the LMPE and it is again the consumers, who compensate the additional costs originating from manually pushing the system away from its economical optimum. It is also noteworthy, that invoking the global  $CO_2$  constraint allows to derive a price for  $CO_2$  ( $\mu_{CO_2}$ ), which must be chosen in order to achieve a power system design, which only emits as much  $CO_2$  as the constraint allows.

Note, that the LMPE vary in space and time. Consequently, they potentially provide incentives to either shift the demand in time (from expensive to cheaper times) or to make investments at places where the price is comparably high. They are used in some deregulated markets especially in the United States, Singapore and New Zealand. In Section 4, I will use

the concept of LMPE to investigate the impact of regional differences in the cost of capital on the optimal power system design. Furthermore, an alternative cost-allocation mechanism based on tracing the flow of electricity will be introduced in Section 6.2.

## 2.4 Data, Models and Software

Technology	investment (€/kW)	fixed O&M (€/kW/a)	marginal (€/MWh)	lifetime (a)
onshore wind	1182	35	0.015 <sup>a</sup>	25
offshore wind	2506	80	0.02 <sup>a</sup>	25
solar PV	600	25	0.01 <sup>a</sup>	25
OCGT	400	15	58.4 <sup>b</sup>	30
run-of-river	3000	60	0	80
hydrogen storage <sup>c</sup>	1966 <sup>d</sup>	9.2	0	20
central battery <sup>c</sup>	1178 <sup>d</sup>	9.3	0	20
hydro reservoir	2000	20	0	80
transmission	400 €/MWkm	2%	0	40

Table 2.1: Overview of technologies and associated overnight costs.

Throughout this thesis, a comparably simple setup of a Pan-European power system will be used. For each grid model, five different generation sources are considered: onshore wind, offshore wind, solar photovoltaics (PV), open cycle gas turbines (OCGT) and run-of-river (ROR). Electricity generated from these sources can be distributed between nodes via transmission lines and/or idealized high-voltage transmission links (please refer to the appendix for a description of the load flow calculations) and stored in hydrogen (H<sub>2</sub>) and central Li-ion battery storage units. Furthermore, hydro reservoir storage units are considered. These units provide electricity from natural water inflow into their reservoirs. They cannot actively take up electricity from the network. Hydro reservoir storage as well as ROR power plants can only be deployed at buses where capacity exists today and only up to a certain multiple of today's capacity (see Figure 2.1) in order to take care of environmental, political and/or societal acceptance issues.

Each generation and storage technology is characterized by a set of associated annualized costs for investment in power and/or energy and for operation and maintenance (O&M), by marginal costs, lifetime, efficiency, power-to-energy ratio  $r_{n,s}$  and CO<sub>2</sub> emissions. All assumptions with respect to these quantities are summarized in tables 2.1 and 2.2. Except for Section 4, overnight investment costs were converted to net present costs using a discount rate of 7% over the entire lifetime. All costs assumption besides the one for H<sub>2</sub> and battery storage units are estimates for 2030 taken from Schröder et al. [2013].

<sup>a</sup>Choosing small marginal costs for these resources allows to determine the order of curtailment.

<sup>b</sup>Including fuel costs of 21.6 €/MWh.

<sup>c</sup>Costs for storage units were taken from Budischak et al. [2013].

<sup>d</sup>For storage units, investment costs for energy storage are assumed to be included in the investment

Technology	efficiency (fraction)	CO <sub>2</sub> emissions (t/MWh)	$r_{n,s}$ (h)
onshore wind	1	0	
offshore wind	1	0	
solar PV	1	0	
OCGT	0.39	0.19	
run-of-river	0.9	0	
hydrogen storage	$0.75 \cdot 0.58$	0	168
central battery	$0.9 \cdot 0.9$	0	6
hydro reservoir	0.9	0	fixed
transmission	1	0	

Table 2.2: Overview of technologies and their specifications.

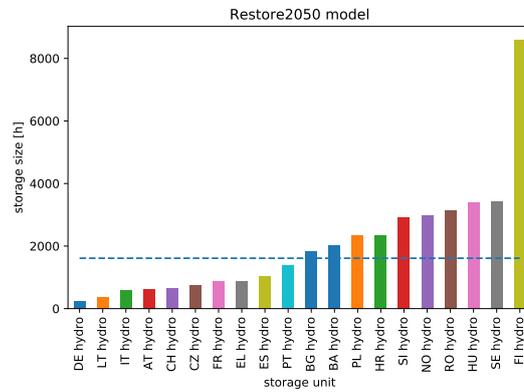


Figure 2.1: Storage size of hydro reservoir storage units in hours at full nominal power; taken from Kies et al. [2016c]; The dashed lines mark the median.

The main difference between the three storage technologies is their ability to store energy over different time horizons. For this thesis, battery storage units are considered as intra-day storage. They are able to store energy equivalent to six hours at full nominal power ( $r_{n,s} = 6$  hours for  $s=Li-ion$ ). Main purpose of these storage units is, for example, to store excess energy from solar PV generated during the day for the evening demand peak. Hydrogen storage units exhibit a much bigger energy storage ( $r_{n,s} = 168$  hours for  $s=H_2$ ). Consequently they are, for instance, able to shift energy from one week to another. The energy storage size of hydro reservoirs storage units are computed from the size of existing dams by Kies et al. [2016b]. Hence,  $r_{n,s}$  for  $s=hydro$  varies between nodes (see Figure 2.1) with a median  $r_{n,s}$  of roughly 1600 hours. Because of this large energy storage and the typical seasonal cycle of the water inflow, hydro storage units play a rather extraordinary role within the power system compared to battery and hydrogen storage units.

---

costs for power.

## 2.5 Coacting Parameters

As described above, power system expansion models depend on economical, (electro-)technical and meteorological parameters. It is the coactions of all these different aspects which determines the optimal design and operation of the power system (see also Section 5). Important parameter features are, for instance, the co-relation of the different volatile resources and the demand, the system's ability to store and distribute electricity and the technology-specific cost. With increasing shares of renewables the dependency of the power system design on the specific characteristics of the renewable resources increases as well [Bloomfield et al., 2018]. Jensen and Greiner [2014] and Schill [2020], for instance, found a phase transition in the need for and use of storage devices at a renewable share of 100%. Kies et al. [2016f] found that demand-side flexibility options can balance fluctuations in the residual load up to a renewable share of 67% and that the presence of these flexibility options significantly increases the optimal share of solar power in a highly renewable European power system (from 19 to 36%). Tranberg et al. [2018] investigated the interplay of the generation sources and different storage technologies based on tracing the flow of electricity. They found, that solar power mainly interacts with short-term storage devices, such as batteries, while wind power rather interacts with hydrogen cavern storage. The reason for this is quite simple: Solar irradiance exhibits a pronounced diurnal cycle. Compared to this, the variability on a weekly scale is small. Furthermore, solar power shows a comparably high co-occurrence with the demand. Let us investigate this with the concept of residual load duration curves [Ueckerdt et al., 2015]. Figure 2.2 shows the residual load duration curves for solar PV, wind power and a combination of the two resources for Oldenburg in 2019 in hourly resolution. Electricity generation was scaled to meet the annual electricity demand. For PV, there is a comparably high share of hours with relatively little residual load (see Figure 2.2 center). This decreases the need to store large amounts of electricity over longer time spans. However, this effect is superimposed by variations on the seasonal scale. The seasonal cycle in solar power is anti-correlated with the seasonal cycle in the demand (at least in Germany): While the demand is highest during the winter season, solar power exhibits its highest potential during summer. This is expressed in a small number of hours of negative residual load during winter and many hours of relatively strong negative residual load during summer. Wind speed, however, does not exhibit any clear diurnal cycle. Hence, variations on the weekly scale explain a higher ratio of the variability. This is expressed on a higher number of hours with relatively high residual load compared to PV (Figure 2.2 top). Furthermore, wind power shows less pronounced differences between the four seasons. There is an increased number of hours of negative residual load during winter compared to the summer season, but this effect is small compared to the seasonal differences in the residual load duration curves for PV. On the positive side of the duration curve, this effect is even smaller. This is caused by a positive correlation of the seasonal cycles of wind power and the electricity demand. The winter months in Germany are potentially windier but also darker and colder. The first increases the potential generation from wind power, the second increases the demand for electricity.

The described differences in PV and wind power adumbrate that a combination of both resources might cancel out some of the specific weaknesses and thus might be beneficial from a power system's perspective. This is the background for studies investigating the *optimal mix* of different generation resources (e.g. Heide et al. [2010], Kies et al. [2016d]). In fact,

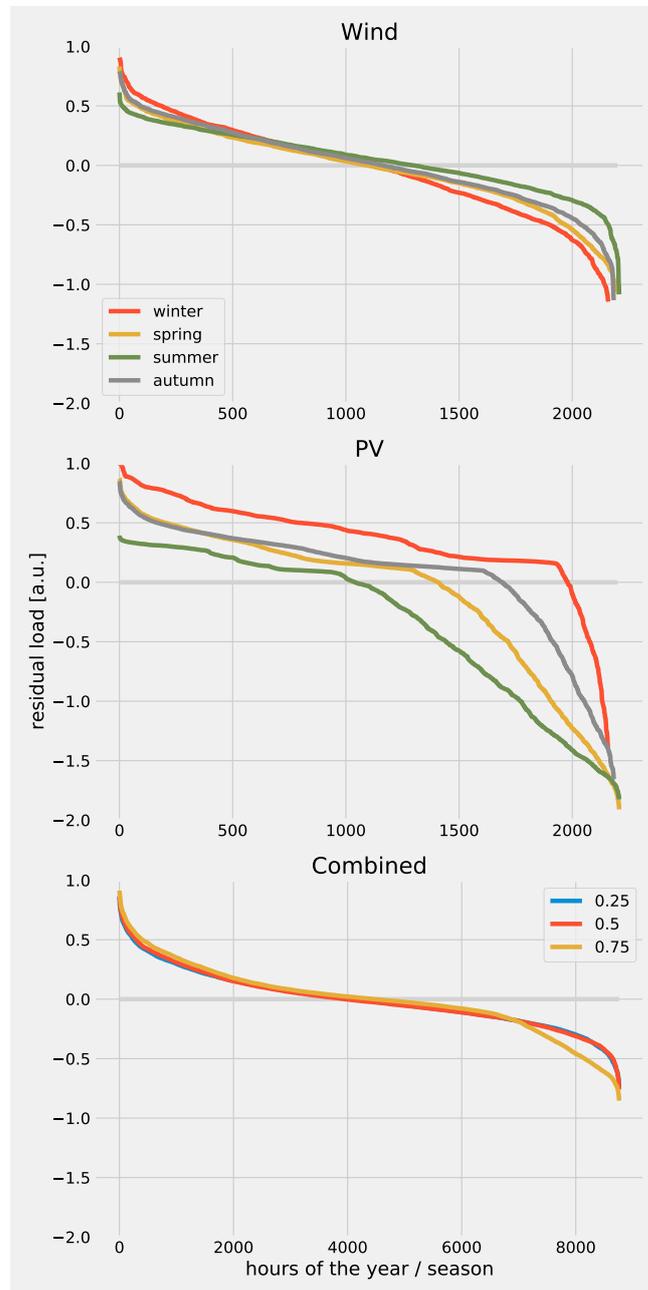


Figure 2.2: Exemplary, synthetic residual load duration curve for Oldenburg 2019 for 100% electricity from wind (top), PV (center) and from a generation mix (bottom); the PV share is given in the legend of the lower figure; renewable feed-in is scaled to meet demand; standard load profile for households.

this optimal mix can be derived from power system expansion models. Hence, it depends on all the different parameters mentioned earlier. For Oldenburg, we find, that a combination of both resources (wind + PV) flattens the residual load duration curve (Figure 2.2 bottom), especially at the extremes at both sides of the duration curve and for relatively little shares of PV compared to wind power.

A similar effect can be observed when generation is accumulated over large geographical areas. Suppose there is only little solar irradiance in Oldenburg due to a thick cloud cover, there might still be bright sunshine a bit further South, say in Burgholzhausen, a village 325 km almost exactly South of Oldenburg (and vice versa). Consequently, one might use this statistical independence of the resources at the two locations by adding both generation time series. This would cancel out part of the variability of the single time series and potentially help to reduce the residual load. This effect is commonly referred to as *spatial smoothing* [Liu et al., 2014, Miettinen et al., 2014, Marrone et al., 2008]. Of course, the statistical dependence of the resources at two different locations, i.e. their correlation, is a function of the distance between the two locations. The point (distance) of statistical independence is called the *correlation length*. For Europe, the correlation lengths for wind and solar power can amount up to several hundreds of kilometers [Martin et al., 2015, Schlott et al., 2018]. Hence, generation assets from volatile renewable resources would optimally be deployed as far away from each other as possible, e.g. by spanning an inter-continental transmission grid [Krutova et al., 2017]. This of course would require huge transmission capacities between all locations and a lot of space at the generation sites for the erection of sufficient generation capacity to meet large parts of the demand of a whole power system. While building these huge transmission capacities might be economically, technically and politically infeasible, the space for power plants is mainly limited by land use restrictions expressed in  $\bar{G}^{max}$  (see Section 2.1). We just found another complex coaction of model parameters implicitly included in power system expansion models with high shares of renewables. I will further elaborate on the importance of the spatial extend and the spatial resolution for power system optimisation in Section 6.1.

Here, let us investigate the spatial smoothing effect by considering the residual load duration curve for the aggregated generation from wind and PV at (i) Oldenburg and Burgholzhausen and (ii) Oldenburg and Almería in Spain approximately 2000 km South-West of Oldenburg for a PV share of 25% in Oldenburg and Burgholzhausen (Bgh 0.25) and 75% in Almería (Al 0.25). Indeed, the spatial smoothing acts as the mixing: We can observe an additional flattening of the load duration curve (Figure 2.3). This flattening is more pronounced for Almería and even more pronounced, when the generation comprises half from solar PV in Almería and half from wind power in Oldenburg (Al opt). In a similar manner, we could, for instance, investigate the effect of sector coupling, storage units and/or demand-side flexibilities.

Of course, the optimal mix and the spatial smoothing are no independent effects. Indeed, the ability to smooth variability strongly affects the optimal mix. And the ability to smooth variability is as much determined by meteorology as it is determined by the available storage and backup technologies. Nevertheless, there is an obvious need to represent meteorological variability as accurate as possible in power system models – including the variability on seasonal to climatic time scales [Bloomfield et al., 2016]. This rises the issue of numerical complexity and tractability, since the number of decision variables and constraints increases

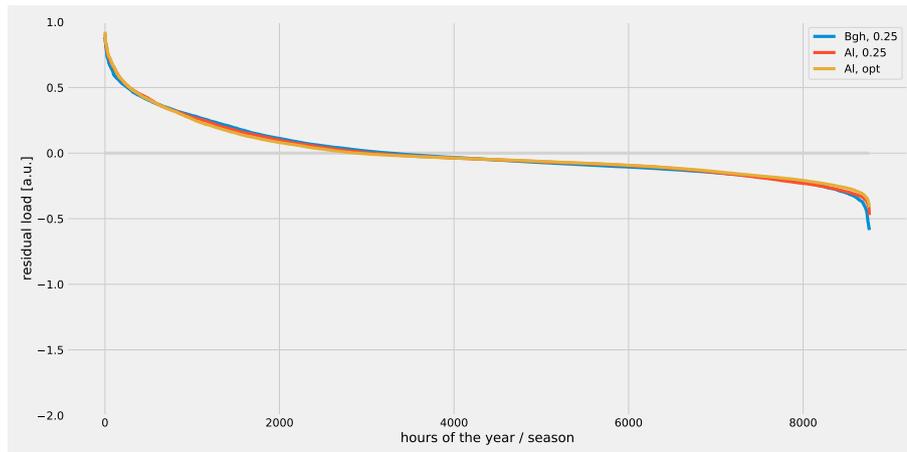


Figure 2.3: Exemplary, synthetic residual load duration curve combined for Oldenburg+Burgholzhausen or Oldenburg+Almería 2019; renewable feed-in is scaled to meet demand; standard load profile for households.

with the length of the parameter time series (compare Sections 2.2 and 5.6.2). The sensitivity of power system expansion models on different representations of the weather (expressed in the availability parameter  $\tilde{G}_{n,s,t}$ ) along with the sensitivities on cost assumptions and model resolutions is the subject of Section 5. The relation between different prevailing weather situations and the continent-wide wind power production will be considered in Section 3.



## Chapter 3

# Research Question 1: How is the relation between weather classes and wind power production

This chapter has been published as:

*Schyska, Bruno U., Couto, António, von Bremen, Lüder, Estanqueiro, Ana and Heinemann, Detlev (2017): Weather dependent estimation of continent-wide wind power generation based on spatio-temporal clustering, Advances in science and research, 14, doi: 10.5194/asr-14-131-2017*

---

Europe is facing the challenge of increasing shares of energy from variable renewable sources. Furthermore, it is heading towards a fully integrated electricity market, i.e. a Europe-wide electricity system. The stable operation of this large-scale renewable power system requires detailed information on the amount of electricity being transmitted now and in the future. To estimate the actual amount of electricity, upscaling algorithms are applied. Those algorithms – until now – however, only exist for smaller regions (e.g. transmission zones and single wind farms). The aim of this study is to introduce a new approach to estimate Europe-wide wind power generation based on spatio-temporal clustering. We furthermore show that training the upscaling model for different prevailing weather situations allows to further reduce the number of reference sites without losing accuracy.

---

### 3.1 Aims and Motivation

A fully integrated European energy market is one of the priority policy areas of the European Commission [e.g. European Commission, accessed 2016-09-22]. Transmission system operators use estimates of the energy production from variable renewable sources within their transmission zones already today. Besides technical aspects, such as the reinforcement of the transmission grid [e.g. Becker et al., 2014, Kies et al., 2016d], also the upscaling algorithms behind these renewable power estimates need to be revised when trading zones are extended – in particular for increasing shares of renewables. In fact, the large-scale integration of variable renewable energy sources (VRES) – such as wind power – introduces additional factors of uncertainty. This uncertainty poses new challenges to the power system operator since it is necessary to keep the balance between production and consumption at every moment, in order to ensure the stability of the power system [Holttinen et al., 2011, Perez-Arriaga and

Batlle, 2012, Estanqueiro, 2008]. In this sense, it is crucial to know the actual and future generation from the VRES within the system. While the future generation is subject of forecasting technologies, this work focuses on the introduction of an upscaling methodology to estimate the Europe-wide actual wind power generation based on spatio-temporal clustering [S. Kisilevich and Rinzivillo, 2010, e.g.]. With the application of upscaling methodologies on the European scale additional potential benefits are expected: Aggregating wind parks with a wide geographical dispersion, for instance, is an effective way to reduce the short term variability and forecast errors by taking advantage of the statistical smoothing effect [Liu et al., 2014, Miettinen et al., 2014, Marrone et al., 2008].

In the current literature, several upscaling approaches can be found: In Ishihara et al. [2007] a typical upscaling function using a bi-exponential function to estimate the cross-correlation is proposed. Pinson et al. [2003] performed a benchmarking of different approaches based on dynamic fuzzy neural networks. In Lobo and Sanchez [2012] the upscaling technique is based on smoothing techniques to construct the predictions of the aggregated wind generation from historical wind speed predictions and the associated wind generation measurements. Recently, Li et al. [2015] proposed a probabilistic approach showing that this type of methodology can provide competitive interval forecasts when compared to conventional statistical approaches. However, all of the upscaling methodologies described above are usually applied to a set of wind parks, and not to the European scale.

As wind is a meteorological quantity, weather conditions may have a strong impact on the wind power variability as well as on the uncertainty of its forecasts [Giebel et al., 2011, Ernst et al., 2007]. Lange and Heinemann [2003] for instance show that the presence of cyclonic systems with strong dynamics – such as cold fronts – can be related to larger errors in the forecast when compared with prevailing weather conditions associated with stationary systems such as anticyclonic systems. A similar methodology was also applied to several wind parks in Portugal demonstrating the weather dependency of the wind power forecast errors [Trancoso, 2012]. Vincent [2010] shows that strong wind variability can be associated with certain weather patterns and Couto et al. [2014] show a strong impact of weather regimes on wind power ramps in Portugal. Consequently, taking into account the underlying role of the synoptic weather patterns could be an important step towards reliable upscaling algorithms.

The objective of this work is to introduce a new upscaling approach for Europe-wide wind power generation based on spatio-temporal clustering (section 3.2.1). The upscaling model will be trained and evaluated for different circulation weather types (CWTs, section 3.2.2) using a set of Europe-wide wind power generation data (section 3.2.3). The training for specific CWTs will be compared to the training over all time steps in the training period in order to investigate its weather dependency and the potential benefit from the weather dependent training (section 3.3). Conclusions will be drawn in section 3.4.

## **3.2 Methodology and Data**

### **3.2.1 Reference Site Selection: Spatio-Temporal Clustering**

Focus of this work is the presentation of a reference site selection scheme based on spatio-temporal clustering. In order to derive a finite set of reference sites to upscale the generation

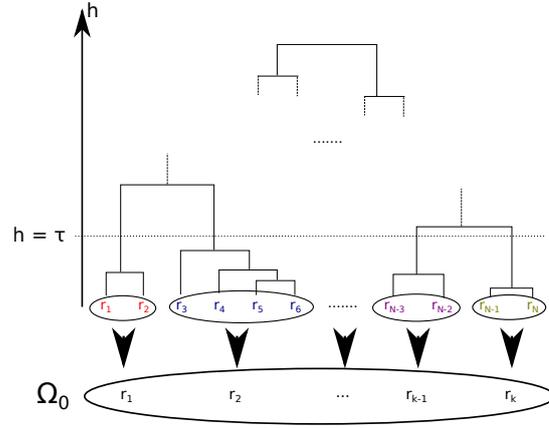


Figure 3.1: Schematic dendrogram for illustration of steps 4 and 5 of the spatio-temporal clustering approach.

of wind power across Europe at a certain point of time the following procedure is applied:

1. Cluster the locations of wind farms (latitude/longitude coordinates) into  $N$  (geographical) clusters via the  $k$ means algorithm [MacQueen et al., 1967].
2. For each of the  $N$  geographical clusters, select the site with the highest wind power capacity. Obtain the set  $\Omega_{geo}$  with size  $|\Omega_{geo}| = N$ .
3. Compute pairwise (temporal) correlations  $\varrho(r_i, r_j) = \varrho(p(r_i, t), p(r_j, t)) \quad \forall r_i, r_j \in \Omega_{geo}$  of the historical generation time series  $p(r_i, t)$  at the  $N$  sites  $r_i, i = 1, \dots, N$  selected in the previous step.
4. Use the correlation information to apply a hierarchical clustering [e.g. Rockach, 2010] with the distance between sites  $r_i$  and  $r_j$  being defined as  $d(r_i, r_j) = 1 - |\varrho_{ij}|$ .
5. Cut the dendrogram obtained from the hierarchical (temporal) clustering at height  $h = \tau$ . Yield  $k = k(\tau) \leq N$  clusters. Here,  $\tau$  is the distance between two clusters. For each cluster, again, select the site with the highest wind power capacity as cluster centres to obtain the final set of  $k$  reference sites  $\Omega_0$ . This step is illustrated in Figure 3.1.

Note that if the average group linkage method is used to agglomerate clusters,  $\tau$  can be interpreted as 1 minus the average intra-cluster correlation. In other words, the final set of reference sites can be determined by choosing the average intra-cluster correlation:

$$\begin{aligned}
 D(A, B) &:= \frac{1}{(|A| + |B|)(|A| + |B| - 1)} \sum_{x, y \in A \cup B} d(x, y) \\
 &= \frac{1}{(|A| + |B|)(|A| + |B| - 1)} \sum_{x, y \in A \cup B} (1 - \varrho(x, y)) \\
 &= 1 - \bar{\varrho}_C
 \end{aligned} \tag{3.1}$$

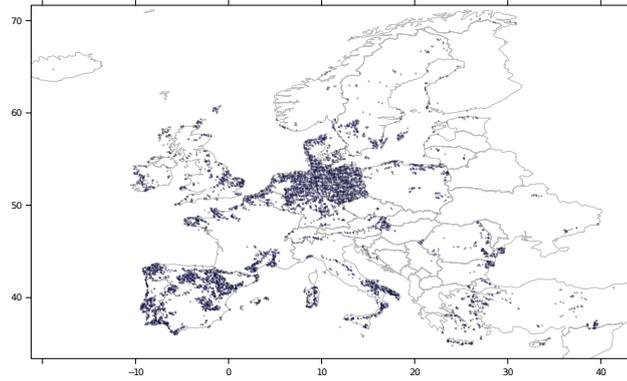


Figure 3.2: Modelled spatial distribution of rated wind power capacity across Europe.

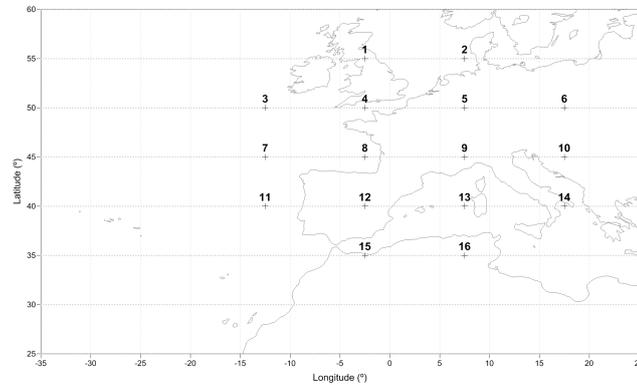


Figure 3.3: Locations of the 16 points used for the circulation weather type identification.

For two clusters (sets)  $A$  and  $B$  and  $C = A \cup B$ , i.e. the cluster resulting from the union of set  $A$  and set  $B$ . Choosing the average intra-cluster correlation as key-parameter to determine the reference sites allows to further investigate the behavior of the clustering approach from a physical-meteorological perspective. This is the major advantage of the proposed methodology compared to, for instance, st-DBSCAN [Birant and Kut, 2007], which does not allow for using different distance measures than the euclidean distance.

### 3.2.2 Upscaling and Evaluation

The upscaling estimate itself for time  $t = t'$  is computed as a weighted sum of the generation measured at the reference sites:

$$E(t = t') = \sum_{r_i \in \Omega_0} w(r_i) p(r_i, t = t') \quad (3.2)$$

Where the weights  $w_i$  are computed from a multiple linear regression of the generation at the  $k$  reference sites  $r_k \in \Omega_0$  on the total Europe-wide generation performed over a pre-chosen training period. Note, that in general the  $w_i$  may vary in dependency of  $\tau$  and  $N$ .

For this study, the upscaling estimate derived from equation 3.2 will be evaluated based on the Pearson correlation and the root mean square error (RMSE) between the upscaling estimate  $E(t)$  and the reference time series for a testing period. Here, the sum of all grid cells of the wind power generation data (section 3.2.3) is used as reference. RMSE values have been normalized to the average hourly wind power production.

In order to investigate the dependency from the prevailing weather situation and the eventual benefit from training the model for specific weather situations, both training and testing will be performed for the nine most common circulation weather types in Europe (see section 3.2.4).

We use five years (2008 - 2012) for training and one year (2013) for testing.

### 3.2.3 Wind Power Generation Data

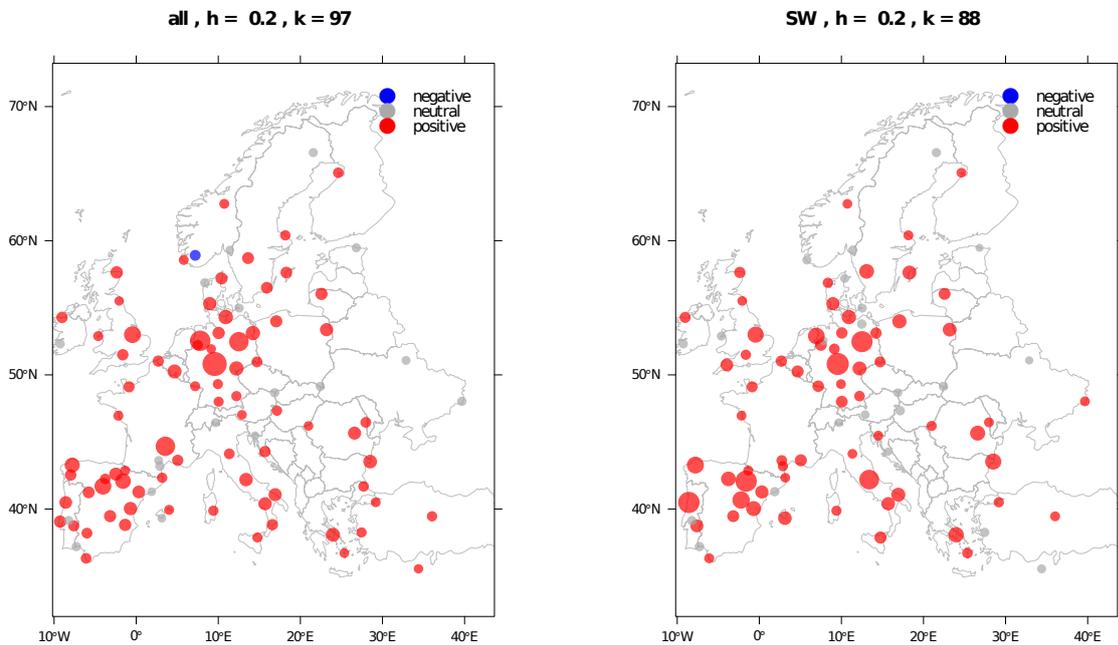


Figure 3.4: Location of the cluster centres and the weights assigned to them by the linear regression (size scale) for  $\bar{\rho}_C = 0.8$  ( $\hat{\tau} = 0.2$ ) and training over all time steps (left) and over the time steps with prevailing CWT SW (right).

The upscaling methodology introduced above is tested for a data set of modeled hourly onshore wind power generation across Europe. This data bases on two data sets: COSMO-EU analysis data provided by the German Weather Service [Doms and Baldauf, 2011] used for the statistical downscaling of MERRA reanalysis data provided by the National Aeronautics and Space Administration of the United States [Bosilovich et al., 2008]. MERRA was used to capture a longer period of time.

The spatial distribution of rated wind power across Europe is modeled as a function of the average (computed over the period considered) wind speed for each location (grid cell)

in Europe. The relation between wind speed and rated power is estimated based on the available data of deployed wind power capacity in Germany. Since this relation is not very distinct, artificial noise has additionally been added:

$$y(r) = a\bar{w}(r) + b + \varepsilon \quad (3.3)$$

Here,  $y(r)$  is the rated wind power at location  $r$ ,  $\bar{w}(r)$  is the average wind speed at the same location,  $a$  and  $b$  are coefficient and intercept fitted from the available data and  $\varepsilon$  is artificial Gaussian noise with zero mean.

The spatial distribution is shown in Figure 3.2. Note, that it does not – and is not meant to – represent the real spatial distribution. Furthermore, offshore locations are not included.

Wind speed is converted to wind power by applying the regional power curve model for the largest German transmission zone developed by Späth et al. [2015]. The procedure described here is similar to the one used by Kies et al. [2016e]. For this study, the years 2008 - 2013 are considered.

### 3.2.4 Circulation Weather Types

Classification of atmospheric circulation into distinct states is a widely used tool for describing and examining weather patterns and their impact on meteorological phenomena, e.g., rainfall [Philipp et al., 2010]. In the literature, several methodologies of weather circulation classification are available [Jenkinson and Collison, 1977, Huth et al., 2008, Philipp et al., 2010, Couto et al., 2014]. In this study, an automatic version of the Lamb weather type classification is applied to MERRA sea level pressure fields in order to obtain a time series of prevailing circulation weather types. This method was initially proposed by Jenkinson and Collison [1977] and thereafter applied by several authors [e.g., Trigo and DaCamara, 2000, Costa et al., 2006].

The algorithm bases on the sea level pressure at the 16 points depicted in Figure 3.3. Assuming geostrophic conditions, westerly and southerly winds can be computed from the meridional and zonal pressure gradient respectively. Doing so, six circulation indices (southerly flow  $SF$ , westerly flow  $WF$ , resultant flow  $FT$ , southerly shear vorticity  $ZS$ , westerly shear vorticity  $ZW$  and total shear vorticity  $ZT$ ) can be computed from the sea level pressure data via:

$$SF = A \cdot 1/4 \cdot (p_5 + 2p_9 + p_{13} - p_4 - 2p_8 - p_{12}) \quad (3.4)$$

$$WF = 1/2 \cdot (p_{12} + p_{13} - p_4 - p_5) \quad (3.5)$$

$$FT = \sqrt{SF^2 + WF^2} \quad (3.6)$$

$$ZS = B \cdot 1/4 \cdot (p_6 + 2p_{10} + p_{14} - p_5 - 2p_9 - p_{13} - p_4 - 2p_8 - p_{12} + p_3 + 2p_7 + p_{11}) \quad (3.7)$$

$$ZW = C \cdot 1/4 \cdot (p_{15} + p_{16} - p_8 - p_9) - \quad (3.8)$$

$$D \cdot 1/4 \cdot (p_8 + p_9 - p_1 - p_2)$$

$$ZT = ZS + ZW \quad (3.9)$$

Southerly and westerly shear vorticity are estimated from the wind shear in the center of the domain. Subscripted numbers indicate the location. The four coefficients  $A$ ,  $B$ ,  $C$  and  $D$  are determined by the central latitude of the chosen raster  $\varphi_0$  (here:  $\varphi_0 = 45^\circ$ ):

$$A = \frac{1}{\cos(\varphi_0)} \quad (3.10)$$

$$B = \frac{1}{2 \cos^2(\varphi_0)} \quad (3.11)$$

$$C = \frac{\sin(\varphi_0)}{\sin(\varphi_0 - 5^\circ)} \quad (3.12)$$

$$D = \frac{\sin(\varphi_0)}{\sin(\varphi_0 + 5^\circ)} \quad (3.13)$$

From the six circulation indices 26 circulation weather types (CWTs) can be deduced as follows:

- If  $|ZT| < FT$  the mean flow dominates over the vorticity (local curvature of the wind field). These CWTs are called directional and named after the eight directions North (N), Northeast (NE), East (E), Southeast (SE), South (S), Southwest (SW), West (W) and Northwest (NW). The flow direction is given by  $\tan^{-1} WF/SF$  if  $WF \leq 0$  and  $\tan^{-1} WF/SF + 180^\circ$  if  $WF > 0$ , respectively.
- If  $|ZT| > 2FT$  the vorticity exceeds the mean flow. The circulation is either cyclonic (L) if  $ZT > 0$  or anticyclonic (H) if  $ZT < 0$
- If  $FT < |ZT| < 2FT$  both, vorticity and mean flow, are equally strong. These CWTs are called hybrid and named after the prevailing circulation, i.e. either cyclonic or anticyclonic, plus one of the eight flow directions.

For this study, the nine most common CWTs in Europe are chosen for evaluation. These are the directional types except for Southeast, the cyclonic type and the anticyclonic type.

## 3.3 Results

### 3.3.1 Cluster Centres and Reference Site Weights

As mentioned above, the number of reference sites varies in dependency of the chosen average intra-cluster correlation. Figure 3.4 shows the locations of the reference sites obtained from the spatio-temporal clustering exemplary for the training over all time steps (left) and the time steps with prevailing Southwestern circulation type (right). The average intra-cluster correlation was exemplary set to  $\bar{\rho}_C = 0.8$ . The size of the dots additionally indicates the weights given to the reference sites by the linear regression. Points with  $|w(r)| < 0.5 \cdot \sigma$  are considered as neutral. Here,  $\sigma$  denotes the standard deviation computed from all weights.

Obviously, the number of reference sites for the CWT SW (3.4 right) is lower (88 to 97). Hence, the correlations of wind power generation at the geographical clusters is higher than

average during time steps of Southwesterly flow – especially on the Iberian Peninsula where the reduction of reference sites is most apparent. Here, wind power production exhibits a relatively coherent spatial structure. This can be related to the passage of large-scale atmospheric phenomena associated with southwesterly circulation, such as cold fronts, able to cover the whole region [Jiménez et al., 2009, Peña et al., 2011]. However, not all of the nine CWTS considered exhibit this higher-than-average correlation. In contrary to southwesterly circulation, some CWTS are usually associated with relatively weak and diffused synoptic scale phenomena. These may cause a less coherent spatial structure of the wind field. Therefore, the number of reference sites for  $\bar{\rho}_C = 0.8$  ranges between 88 for SW and 105 for the Easterly flow type (not shown).

From Figure 3.4 it can also be seen, that the weights given to the selected reference sites vary as well. The reference sites on the Iberian Peninsula get relatively higher weights for the Southwesterly circulation type than for all time steps.

### 3.3.2 Upscaling Evaluation

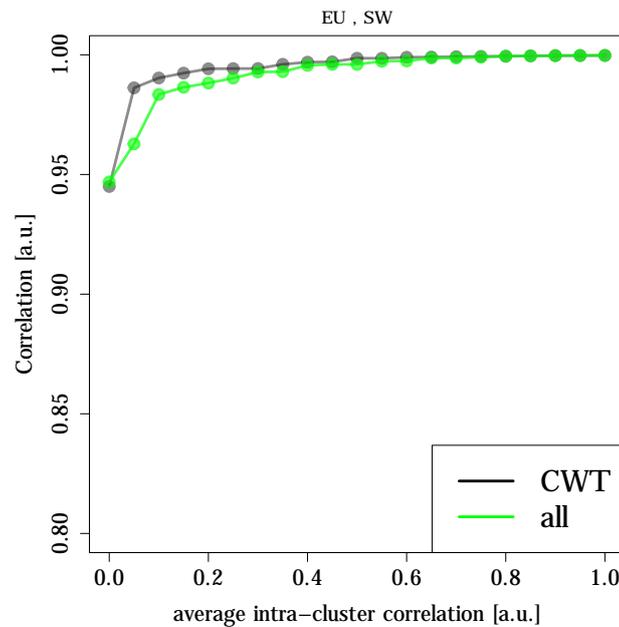


Figure 3.5: Correlation versus the average intra-cluster correlation  $\bar{\rho}_C$  for CWT SW obtained from the specific training for this CWT (black) and from the training over all time steps (green) respectively.

The skill of the methodology introduced in section 3.2 measured by correlation and RMSE is exemplary shown in Figures 3.5 and 3.6 for the Southwesterly circulation type. It can be seen, that very high ( $> 0.95$ ) values for the correlation can be achieved for average intra-cluster correlations above 0.1. For the Southwesterly CWT this corresponds to a number of reference sites  $k = 17$  for whole Europe. For higher  $\bar{\rho}_C$  the correlation asymptotically approaches 1.

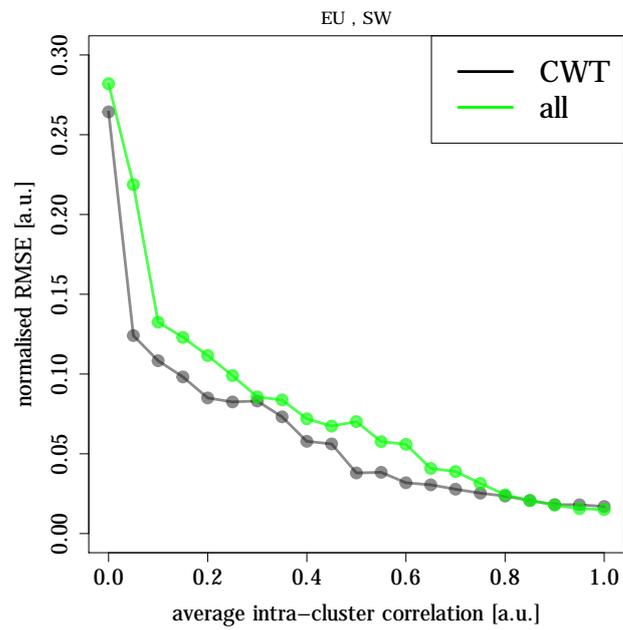


Figure 3.6: As Figure 3.5 but for the RMSE normalised to the average generation.

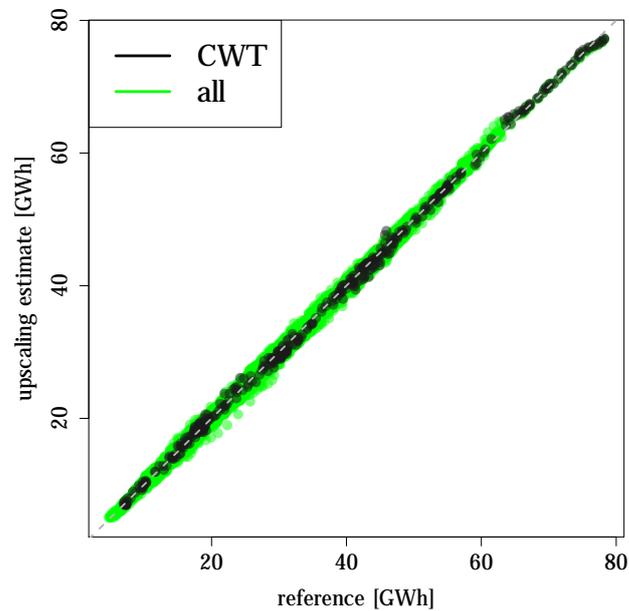


Figure 3.7: Time series of the upscaling estimate [GWh] versus the reference time series [GWh] for all time steps (green) and time steps with prevailing Southwesterly circulation (black).

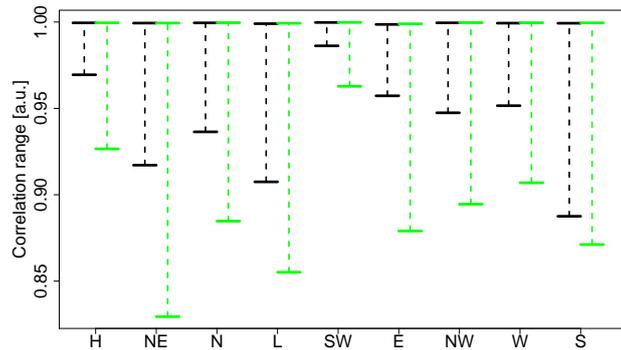


Figure 3.8: Range of correlation values achieved by training the upscaling for the specific CWTs (black) and from training over all time steps (green).

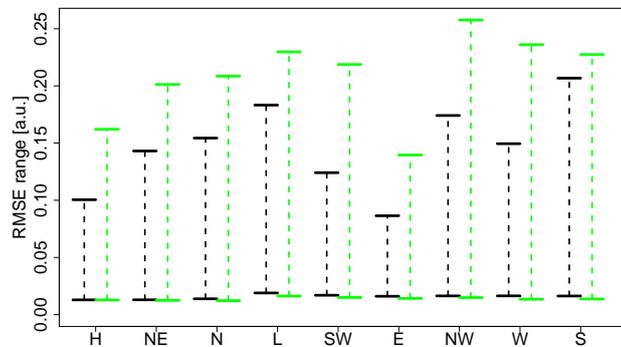


Figure 3.9: As Figure 3.8 but for the RMSE normalised to the average generation.

A similar behaviour is found for the RMSE. For  $\bar{\rho}_C > 0.1$  the RMSE drops below 10% of the average wind power generation in Europe. For average intra-cluster correlations above 0.45 – corresponding to  $k = 41$  – RMSE values below 5% of the average generation can be achieved.

The good agreement between the upscaling estimate and the reference time series can additionally be seen from the scatter plot (Figure 3.7, again for  $\bar{\rho}_C = 0.8$ ). A systematic error only appears for extreme high (above 75 GWh) wind power generation values. Here, the upscaling model systematically underestimates the generation. Furthermore, all these extreme values occur during Southwesterly circulations. This reduces the skill of the upscaling model for this CWT disproportionately strong.

$\bar{\rho}_C = 0$  does not involve any hierarchical clustering. The corresponding data point is considered as non-representative and therefore neglected from the further analysis.

### 3.3.3 Benefit from Training for Weather Types

In general, the Southwesterly CWT is the one, for which the introduced upscaling methodology works best with respect to the correlation (Figure 3.8, black bars). Other CWTs exhibit lower correlations. With respect to the RMSE, the SW type only skills average (Figure 3.9,

black bars). Here, especially the Easterly type benefits from the specific training.

Figures 3.8 and 3.9 show the range of the correlation and the RMSE for all  $\bar{q}_C \in ]0, 1]$  obtained from (i) the training specifically for the particular CWTs in black and (ii) training over all time steps in green. Evidently, the upscaling skill benefits from the specific training. The range of both, correlation and RMSE, can be reduced significantly. It can furthermore be observed that the cyclonic CWT and the Southerly CWT perform worst – with respect to both correlation and RMSE – while the Easterly, Southwesterly and Cyclonic type perform best. The benefit from the CWT specific training is strongest for the Northeasterly and Northwesterly type with respect to correlation and RMSE, respectively.

### 3.4 Discussion and Conclusions

In order to derive a reduced set of reference sites to estimate Europe-wide wind power production, a new spatio-temporal clustering approach has been developed. To test the methodology, model data is used, which is known to be smoother than measured data. Keeping this in mind, we have shown that a rather low number of around 40 reference sites – when chosen carefully – is sufficient to estimate the actual wind power generation across whole Europe with adequate accuracy. We have also shown that it is beneficial to train the upscaling model for different prevailing circulation weather types.



## Chapter 4

# Research Question 2: How do regional differences in cost of capital influence the optimal design of power systems?

This chapter has been published as:

*Schyska, Bruno U. and Kies, Alexander (2020): How regional differences in cost of capital influence the optimal design of power systems, Applied Energy, 262, doi: 10.1016/j.apenergy.2020.114523*

---

In order to reduce greenhouse gas emissions of the power sector, high shares of renewable power sources need to be integrated into existing systems. This will require vast amounts of investments. Cost of the capital needed for these investments are unevenly distributed among European regions. They show a clear North-South and West-East divide, which has not exhibited significant signs of narrowing in recent years. Power system studies investigating a continent-wide European power system, however, usually assume homogeneous cost of capital.

The objective of this paper is to investigate how regional differences in cost of capital affect the result of these studies with respect to the optimal power system design. Our analysis is based on power system optimization with inhomogeneous cost of capital in Europe. We find that assuming homogeneous cost of capital leads to estimates on the levelized costs of electricity in a highly renewable European power system, which are too conservative. The optimal system design is significantly affected compared to an inhomogeneous scenario. In particular, we show that inhomogeneous cost of capital favors overall wind power deployment in the case of Europe, while the investment in solar power decreases.

---

### 4.1 Aims and Motivation

In the European Union, renewable shares are growing rapidly. This requires vast amounts of capital for investments. Contrary to conventional generation, where generation and operation cost (fuel costs, maintenance, etc.) is crucial, renewable energy sources like wind and solar PV require large upfront investments, while fuel costs are non-existent. In addition, most mentioned integration options are capital-intensive or might even require building up entirely new infrastructures such as demand-side management with smart meters or electric

vehicles and their charging infrastructure. This increases the need for investments even further. The European Commission estimates necessary investments of 379 billion Euro p.a. in the European Union after 2021 [European Commission, Accessed: 2018-09-30]. Zappa et al. [2019] estimate the cost for building a 100% renewable European power system until 2050 to be 560 billion Euro p.a.. These costs are to a large extent driven by the cost of the capital needed to make investments. According to Noothout et al. [2016], conditions to invest in renewables vary considerably between regions in Europe. These variations are reflected in different rates at which investors can raise funds. Noothout et al. [2016] measure these rates as the *weighted average cost of capital* (WACC). Regional differences in WACC mainly originate from different financing and tariff-related risks for renewables [Noothout et al., 2016]. Egli et al. [2018] reported that financing conditions for solar PV and wind power projects in Germany have improved significantly over the last years contributing to the reduction of the levelized cost of electricity from these two resources.

In this context, Klessmann et al. [2013] suggested to reduce the capital needs for renewables by applying political means to reduce financing risks in the renewables sector. Temperton [2016] proposed to establish a transnational European facility in order to reduce the cost for financing renewable projects in Europe. And Kitzing et al. [2012] showed that a convergence of policies to support renewables could be observed, which might spur a slight convergence of WACC in the future. However, Brückmann [2018] found that no tightening of this WACC gap has occurred in recent years. An interaction between the WACC and the price for CO<sub>2</sub> emissions has been found by Hirth and Steckel [2016]. Accordingly, higher WACC require higher CO<sub>2</sub> prices to achieve the same reduction target.

Existing power system studies often do not consider regional differences in the cost of capital: Schlachtberger et al. [2017], for instance, assumed spatially homogeneous costs for optimizing a European power system and for investigating the benefits from increased continent-wide transmission capacity limits. Schlott et al. [2018] investigated the impact of climate change on a similar system with the same homogeneous cost assumption. Bearing the results of Noothout et al. [2016] in mind, the assumptions made in these studies appear questionable.

The aim of this study is to investigate the effect of regional differences in cost of capital in Europe on the cost-optimal design of power systems with ambitious CO<sub>2</sub> reduction targets. Unlike other studies in the same field of research of recent years, we directly consider regional differences in the cost of capital. We investigate changes in expenditures for investment and operation compared to a homogeneous reference setup and relate these changes to differences in the cost-optimal deployment of generation capacity. Furthermore, we investigate the impact on overall system costs and the effect of diverging cost of capital. By doing so, our work contributes to a deeper understanding of the effect of input parameter uncertainties on the results of power system optimization models.

## 4.2 Data and Methods

### 4.2.1 Power System Expansion Modelling

In this paper, we investigate the impact of regional differences of the cost of capital on the cost-optimal design of a simplified highly renewable European power system. This cost-optimal design is derived from a greenfield expansion model. No existing generation, transmission and storage assets or retirement of these assets is considered. Instead, the cost-optimal expansion of all assets (starting from scratch) is determined using mathematical programming (Equation (2.1) - (2.15)). The power system considered comprises one node per country (30 in total) and 52 simplified transmission links connecting them. Electricity generation may stem from solar PV, on- and offshore wind, open cycle gas turbines (OCGT), run-of-river power plants and hydro reservoirs. Furthermore, we considered pumped hydro storage units and two generic storage types with fixed power-to-energy ratios (Table 2.1). Generation from coal and nuclear power plants has not been included for the following reasons: Firing coal for the generation of electricity contradicts the European Union's goals to significantly reduce greenhouse gas emissions. This goal is implemented by a global cap on CO<sub>2</sub> emissions (see below). Nuclear power is expected not to be cost-competitive. According to Wealer et al. [2019], investing in nuclear power is uneconomical independent of the cost of capital, electricity prices and specific investment costs. Lazard [2019] found levelized costs for electricity (LCOE) for nuclear power to be between 112 and 183 USD/MWh, which is far above the LCOE for utility scale PV (43-53 USD/MWh) and onshore wind (30-60 USD/MWh). Furthermore, nuclear power exhibits a similar ratio between capital and marginal costs as the renewable resources, i.e. high upfront investments and almost no marginal generation costs. Thus, results likely hold if new developments render nuclear power a cost-efficient alternative.

In our model, the generation capacity of PV, wind and OCGT as well as the transmission capacity were expandable, but we fixed the capacity of hydro dams, run-of-river plants and pumped hydro storage units to the values published by Kies et al. [2016c].

Different formulations of power system expansion models exist. The main difference is whether unit commitment of the generators or discrete expansion steps for the generator's nominal power or the transmission capacity are considered. In both cases, the optimization problem to be solved would be a mixed-integer problem. For this study, we use the pure linear approach, which has been used in several studies before, e.g. [Brown et al., 2018, Schlachtberger et al., 2017, Schlott et al., 2018]. The optimization problem used to derive the optimal system design contains investments in generation, storage and transmission capacity as well as hourly operational costs originating from load dispatch. For the mathematical formulation of this problem see Section 2.1.

For this paper, we assumed a global limit on transmission capacity of three times today's net transfer capacities (3·31.25 TWkm) as an appropriate compromise between cost-optimal extension and technical and social concerns. Although, this assumption is slightly more conservative than the compromise grid defined in Schlachtberger et al. [2017] and Brown et al. [2018] at four times today's values, it allows to capture large parts of the benefits of distributing electricity from renewable resources due to the non-linear decrease in system costs with increasing transmission capacity [Schlachtberger et al., 2017]. Inline with European

emission reduction targets, we define a global CO<sub>2</sub> cap of 5% of the historic level of 1990. In this model, OCGT is the only generation type with non-zero CO<sub>2</sub> emissions. Emissions  $e_{n,s}$  and efficiency  $\eta_{n,s}$  are set to 0.18 tonnes per MWh and 0.39, respectively.

The same methodology has been used to study, for instance, the impact of climate change [Schlott et al., 2018], synergies between sector coupling and transmission [Brown et al., 2018], the benefit of cooperation in a highly renewable European power system [Schlachberger et al., 2017] or the impact of CO<sub>2</sub> constraints [Weber et al., 2019]. We used the software-toolbox *Python for Power System Analysis* [Brown et al., 2017] and the commercial Gurobi solver to solve the optimization problem.

## 4.2.2 Locational Marginal Prices for Electricity

Besides the cost-optimal deployment of generation capacity and the dispatch of electricity, the power system expansion modeling also delivers some information about how prices must develop to achieve this optimal system design via its dual problem. For instance, if a global, i.e. system-wide, CO<sub>2</sub> cap constraint is enforced, the corresponding dual variable  $\mu^{\text{CO}_2}$  to this constraint yields a price for CO<sub>2</sub> emissions. If this price was set to the emissions in the primal problem without CO<sub>2</sub> constraint, the exact same optimization result would be obtained. In the same manner, the dual problem determines the price for electricity for each time step and at each bus within the system, i.e. the LMPE (compare Section 2.3).

The dual problem corresponding to Equation 2.1 - 2.15, again, is a linear problem. By applying the Karush-Kuhn-Tucker conditions, its objective function reads:

$$\max_{\lambda, \mu_{\text{CO}_2}, \mu_{\text{trans}}} \sum_{n,t} \lambda_{n,t} D_{n,t} + \mu_{\text{CO}_2} \sum_{n,s,t} \frac{1}{\varepsilon_{n,s}} g_{n,s,t} E_{n,s} + \mu_{\text{trans}} \sum_l \bar{f}_l L_l \quad (4.1)$$

At the global optimum, primal and dual are in equilibrium, Equation (2.1) equals Equation (4.1), which is equivalent to:

$$\sum_{n,s} C_{n,s} \bar{g}_{n,s}^* + \sum_{n,s,t} O'_{n,s} g_{n,s,t}^* + \sum_l (C_l - \mu_{\text{trans}}^*) L_l \bar{f}_l^* = \sum_{n,t} \lambda_{n,t}^* D_{n,t} \quad (4.2)$$

with  $\mu_{\text{CO}_2}^*$  and  $\mu_{\text{trans}}^*$  being the global CO<sub>2</sub> price and the global price for expanding the transmission lines by one MWkm, respectively. Recall from Section 2.3 that relaxing constraints (2.7) and (2.6) leads to a decrease of the objective function. Hence,  $\mu_{\text{trans}}$  and  $\mu_{\text{CO}_2}$  are less than or equal to zero.  $\lambda_{n,t}$  is the LMPE, which is the price consumers at a node  $n$  and at time  $t$  would have to pay for their electricity demand  $D_{n,t}$  in equilibrium. According to Equation (4.2), these payments cover the following three cost terms: (i) the regional investment in generation and/or storage capacity, (ii) the operational costs for generating electricity locally depending on the respectively available generation sources and (iii) the investment in transmission capacity. Here, the updated operational costs  $O'_{n,s,t} = \left( O_{n,s} - \mu_{\text{CO}_2}^* E_{n,s} / \varepsilon_{n,s} \right)$  also include the costs for CO<sub>2</sub> emissions and the effective costs for transmission reinforcements are given by the sum of the capital costs and the shadow price for transmission related to the global transmission capacity limit (Equation 2.7).

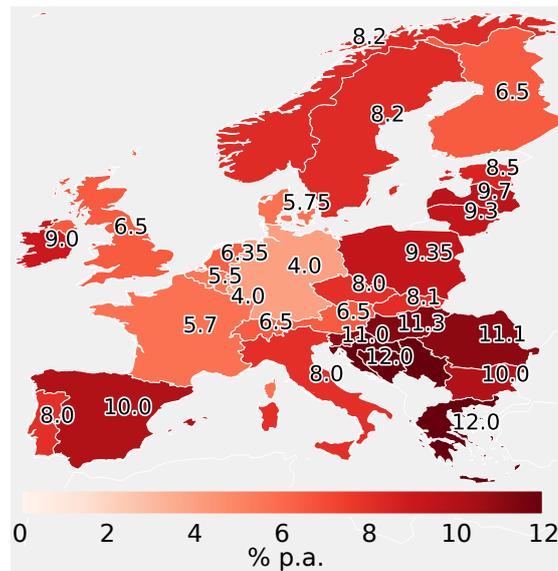


Figure 4.1: Weighted average cost of capital taken from Noothout et al. [2016].

### 4.2.3 Cost of Capital Scenarios

Investors raise funds from different funding sources and at different rates. The cost of capital – when measured by the WACC – is the weighted average of these rates. It comprises the costs of equity (between 6 % and 15 %) and debt (between 1.8 % and 12.6 %). If financial risk was equal across different regions, e.g. by removing information asymmetry and taxes, the cost of equity and debt should be equal as well [Modigliani and Miller, 1958, 1963]. Hence, the WACC reflects varying market conditions between regions [Noothout et al., 2016]. In order to investigate the impact of these varying conditions, we used the WACC obtained from Noothout et al. [2016] for the EU states. The WACC of the four remaining states in the area of interest have been assigned manually from their respective *most similar* neighbouring country. Resulting WACC are shown in Fig. 4.1. Germany has by far the lowest WACC at 4 %, whereas the value peaks in South-Eastern Europe at 12 %. In Noothout et al. [2016], the WACC are only given for wind onshore projects, but reasons given for varying WACC – such as financing environment and policy risks – hold true for all investments in renewables. We therefore assumed one uniform WACC for all generation and storage technologies per country. Because tariff-related risks are the major cause for discrepancies in WACC between European countries [Temperton, 2016], it is unlikely that increasing shares of foreign direct investment or increasing capital investments due to increasing shares of renewables would have a relevant effect on WACC.

For this study, we varied the WACC for a number of scenarios:

1. The *today* scenario used the WACC values shown in Fig. 4.1.
2. The *homogeneous* scenario considered a constant WACC of 7.1% across Europe. 7.1% was obtained as the demand-weighted average of WACC from the today scenario. The homogeneous scenario is used as the main reference throughout this study.

3. For the *inhomogeneous scenario*, the difference between every region and the average WACC has been doubled compared to the today scenario.
4. For nineteen scenarios indexed with the numbers -9 to 9, the WACC was linearly interpolated between the homogeneous scenario (index 9) and the inhomogeneous scenario (index -9) to investigate potential path-dependencies.

When used as the return rate  $r$ , the WACC determines the annuity  $a_{n,s}$  of an asset  $s$  in region  $n$  with lifetime  $l_s$  via

$$a_{n,s} = \frac{r_n (1 + r_n)^{l_s}}{(1 + r_n)^{l_s} - 1} \quad (4.3)$$

From these annuities and the *investment costs*  $c_{n,s}^{\text{inv}}$  the annualized *capital costs*  $c_{n,s}$  of an asset can be computed via

$$C_{n,s} = c_{n,s}^{\text{inv}} \cdot a_{n,s} \quad (4.4)$$

They are given in Euro per MW per year for a generator  $s$  in region  $n$ .

The *investment* refers to the product of the capital costs and the installed capacity  $\bar{g}_{n,s}$ . Similarly, the *operational costs* are defined as the product of the *marginal costs*, i.e. the costs for generating one (additional) unit of electricity, of a generator  $O_{n,s}$  and the actual generation  $g_{n,s,t}$ . Capital costs  $C_{n,s}$  and operational costs  $O_{n,s}$  for all technologies are given in Table 2.1.

We derived a cost-optimal design of a European power system with an ambitious CO<sub>2</sub> reduction target by solving the previously introduced optimization problem for the four aforementioned scenarios. Regional expenditures for investment and operation were measured as:

$$I_n = \frac{\sum_s C_{n,s} \cdot \bar{g}_{n,s}^*}{\sum_t D_{n,t}} \quad (4.5)$$

and

$$R_n = \frac{\sum_{s,t} O'_{n,s} \cdot g_{n,s,t}^*}{\sum_t D_{n,t}}$$

respectively and compared between the simulations. Here, the asterisk indicates the optimal solution obtained from solving Equation (2.1) - 2.15.

In Equation (2.1) through (2.15), investment in the inter-connecting transmission grid is included via cost-optimization as well as through the assumption of a global limit on transmission capacity extension. This reflects that the expansion of the power system often is not purely a technical-economical problem. Instead, it is often hampered by political and social constraints – such as missing public acceptance, for instance – meaning that although transmission grid extension might be cost-optimal, it cannot be realized. Consequently, the costs for transmission expansion can only be estimated with great uncertainty. As shown above one can, nevertheless, derive a shadow price for transmission grid expansion  $\mu^{\text{trans}}$ . However, this shadow price is a *political* price, which cannot directly be compared to the market prices  $C_{n,s}$  and  $O_{n,s}$ . Therefore, we focus on the investment in generation and storage assets in Section 4.3.

#### 4.2.4 Generation and Load Data

We use one year of hourly availability data for onshore wind, offshore wind and solar PV as described by Kies et al. [2016c]. The underlying weather data stems from the MERRA reanalysis [Rienecker et al., 2011] as well as Meteosat First and Second Generation. Feed-in from wind has been modeled using the power curve of an Enercon E-126 at 140 m hub height. AC power from PV modules has been simulated by applying the Heliosat method [Cano et al., 1986, Hammer et al., 1998], the Klucher model [Klucher, 1979] as well as the parameters of a Sunny Mini Central 8000TL converter. The natural inflow to hydro dams and run-of-river power plants is taken from Kies et al. [2017]. It has been modeled as a linear function of the potential energy of the the run-of data obtained from the ERA Interim reanalysis [Dee et al., 2011]. More details on the methodology applied to generate the time series are given by Kies et al. [2016c,g]. Load time series were derived from historical load data provided by ENTSO-E and modified within the RESTORE 2050 project to account for expected increasing shares of e-mobility and electric heat pumps [Kies et al., 2016a].

### 4.3 Results

If the regional distribution of the cost of capital would not affect the optimal system design, investments would change between these scenarios in the same way as there is a difference in the WACC: It would rise in countries where cost of capital increases and shrink where cost of capital decreases compared to the reference. Instead, we observe highest increases in investments in the country, which exhibits the lowest WACC: Germany. Here, investments in generation and storage assets increase by approximately 37.5 Billion Euro per annum or 96 % (Fig. 4.2 top). Increasing investments are also observed for Belgium, Austria, Spain, Greece, Bulgaria, Portugal and Finland. Of these countries, Belgium, Austria and Finland exhibit relatively low cost of capital like Germany, while they are relatively high in the others (Fig. 4.1). This evidences significant changes in system design, expressed by changes in optimal capacity deployment, which in turn influences regional operational costs (Fig. 4.2 bottom).

Compared to the homogeneous scenario, inhomogeneous WACC lead to a strong agglomeration of power generation capacity in Central-Western Europe, especially in Germany, France, Austria and Belgium (see supplementary material in the appendix 7). Consequently, the penetration rate, i.e. the ratio of local electricity generation over local demand, in Germany and Austria increases by 120 % and 50 % respectively (Fig. 4.3). While both are net importers in the homogeneous scenario, they become net exporters in the today scenario. In turn, many countries in Eastern Europe exhibit a higher dependency on imports in the inhomogeneous scenario. This is reflected in penetration rates below one and in a distinct step-wise increase in the locational marginal price for electricity the further the respective country is from the exporting countries in Central-Western-Europe (Fig. 4.4 right).

We assume equal marginal costs for each type of generator, no matter at which node the generator is located. Hence, a rise in operational costs as depicted in Fig. 4.2 (bottom) can only be caused by a replacement of generators with low marginal costs, i.e. wind and PV, with gas power plants and/or the intensified use of gas power. Again, this shift from

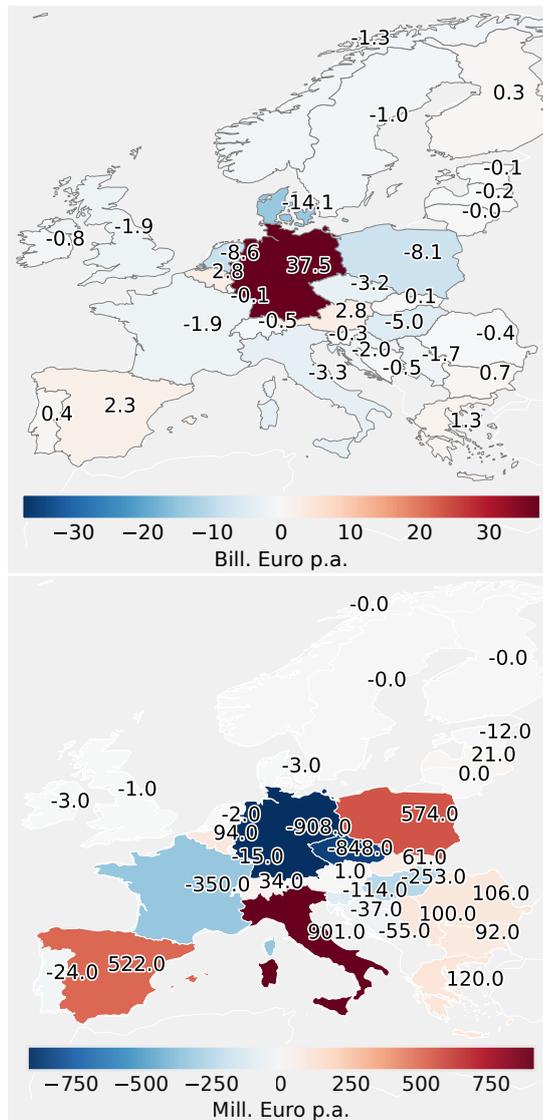


Figure 4.2: Change in regional investment in generation and storage units (top) and operational costs (bottom) between the homogeneous and the today scenario.



one generation technology to another between the today and the homogeneous scenario is most pronounced in Germany. Expenditures for the regional operation of power plants decrease by approximately 908 Million Euro per annum. France and the Czech Republic profit from the cheap electricity generation in Germany (-350 and -848 Million Euro per annum, respectively). And Hungary increasingly imports electricity from Austria, which leads to a decrease in regional operational costs of 253 Million Euro per annum. In Italy, Poland, Spain and the South-Eastern European countries, however, regional expenditures for operation increase due to the intensified deployment and use of less capital-intensive gas power (see Appendix 7). Overall, a slight increase in operational costs is observed due to an increased generation share of offshore wind power in the today scenario compared to the homogeneous scenario.

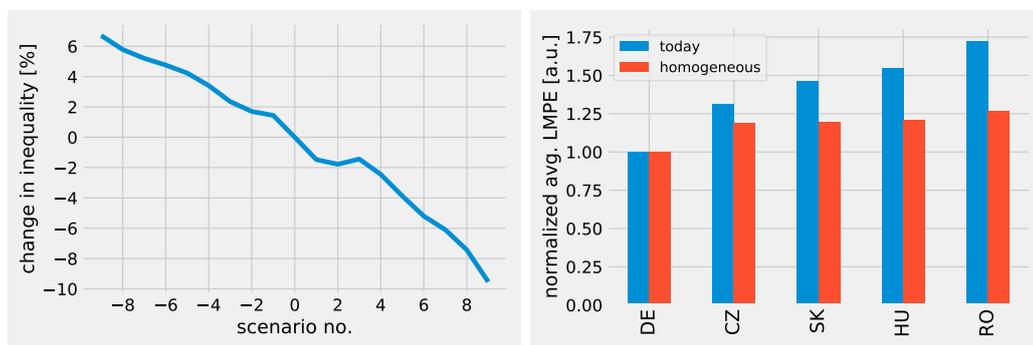


Figure 4.4: Difference in inequality in regional expenditures for electricity as measured by the 20:20 ratio [a.u.] relative to the today scenario (left) and cascade in LMPE for four countries downstream of Germany normalized to the LMPE of Germany [a.u.] (right).

In combination, the mentioned effects have different implications for different countries in Europe concerning the optimal deployment of the different generation technologies and the regional expenditures. In Germany, for instance, below-average WACC would lead to increasing overall investment and capacity deployment, especially of capital intensive wind power (+280 GW, see Appendix 7). Nevertheless, LMPE would fall due to significantly higher penetration rates (+120 %, Fig. 4.3) and a lower dependency on gas power (-11 GW) and imports (Fig. 4.5). The Czech Republic profits from its proximity to the 'export country' Germany. The fact that Germany exports electricity reduces the need for local investments and the expenditures for operation in the neighbouring Czech Republic. Transmission cost are low, because the two countries are directly connected, and consequently the LMPE decreases. A similar effect can be observed in Sweden and Denmark. In Hungary, LMPE rise, although the expenditures for investment and operation and the penetration decrease here as well. This is due to the fact that Hungary is located further away from the exporting countries (Fig. 4.4 right). Thus, LMPE need to cover more investment in transmission capacity upstream.

Wind power in general – both onshore and offshore – profits from relatively low WACC in Germany as, to a smaller extent, in Belgium, France and Finland. Besides relatively low cost of capital, these countries exhibit relatively good wind resources and, at least in the

case of Germany, high demand for electricity and a favourable topological position within the network. This leads to the aforementioned additional 280 GW onshore wind power in Germany, 24 GW in France, 7 GW in Belgium and 4 GW in Finland in the today scenario compared to the homogeneous reference. Additional offshore wind power deployment only occurs in Germany: +30 GW. In most other countries, onshore wind power deployment drops: Denmark -77 GW, Poland -65 GW, the Netherlands -47 GW, Italy -39 GW. In sum, continent-wide onshore and offshore wind power installations increase by 8 GW and 21 GW, respectively. Figures showing the regional changes in nominal power and investment per generation source can be found in the Appendix 7.

In contrast to wind power, continent-wide solar PV installations decrease by 24 GW between the homogeneous scenario and the today scenario. The main reason for this is the reduction of deployment of PV power in countries with high cost of capital, especially the Czech Republic (-30 GW), Hungary (-22 GW), Croatia (-16 GW) and Serbia (-11 GW). As described above, this reduction is not driven by the costs of capital alone, but triggered by the relative proximity to countries with relatively low WACC, i.e. Germany and Austria. In these countries increased deployment of PV power can be observed (Germany +6 GW, Austria +38 GW). In Belgium PV power deployment rises by 26 GW. In Spain, PV replaces 10 GW of even more capital-expensive wind power.

It has been shown by Tranberg et al. [2018] that solar PV and gas power are suitable complements. In countries with relatively good solar resources, gas power plants – potentially together with battery storage units – are used to cover load peaks and to bridge times of low sunshine. Consequently, as PV power deployment reduces, the continent-wide deployment of gas power plants decreases by 10 GW in the today scenario compared to the homogeneous scenario as well.

It is also noteworthy that, although operational costs only decrease slightly, overall system costs significantly dependent on the spatial distribution of the WACC. Compared to the homogeneous scenario, inhomogeneous WACC lead to a reduction of levelized costs of electricity of approximately 2.5 % in the today scenario and of up to more than 5 % in the inhomogeneous scenario (insert of Fig. 4.5). This reduction in LCOE is due to the complex interaction of regional demand, costs, the quality of the volatile renewable resources and the capabilities of transmitting and storing electricity. We describe this complex interaction by investigating the levelized costs for electricity separately for each generator type. It is defined as the sum of generator type specific investment and operational costs divided by the electricity generation. Since the implemented CO<sub>2</sub> cap (Equation (2.6)) is reached in each scenario, the generation from gas power plants – the only generator type with non-zero CO<sub>2</sub> emissions considered – does not change between the scenarios. Additionally, a decrease in gas power capacity installation can be observed (Fig. 4.6 left). Hence, the utility rate for gas power increases in the today scenario compared to the homogeneous scenario. However, because gas power is increasingly deployed in countries with relatively high cost of capital, levelized cost for gas power only decrease slightly by less than 1 %. A similar effect can be observed for solar PV: As mentioned above, PV power suffers from the relatively high cost for capital in the countries with relatively good solar resources. This leads to an increase in levelized costs for PV power of approximately 6 %. Consequently, the decrease in LCOE must be driven by a reduction in the levelized costs for wind power. Indeed, the levelized cost for onshore wind power decreases by around 8 % and for offshore wind power by 15 %

between the today and the homogeneous scenario caused by the co-occurrence of high demand, relatively low costs and good wind resource quality in the Central-Western European countries mentioned already. A slight increase in the levelized costs for storage units (batteries +1 %, H<sub>2</sub> +2 %) counteracts this effect. For storage units, operational costs have been determined from the product of LMPE and energy uptake as suggested by Pawel [2014].

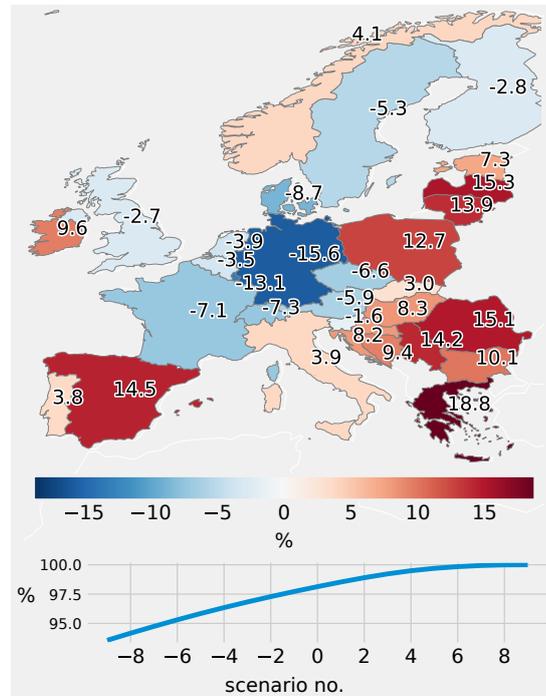


Figure 4.5: Regional relative change in locational marginal prices for electricity between the homogeneous scenario and the today scenario. The insert shows the system costs normalized to the systems costs of the homogeneous scenario.

However, the reported reduction in costs is unevenly distributed among countries. In 17 out of 30 countries, LMPE rise due to a relative increase of cost of capital leading to higher local investments, higher shares of generators with low capital but high marginal costs and a higher dependency on imports, i.e. investment in transmission. This causes a growth in inequality of regional expenditures for electricity of up to 10 % in the today scenario compared to the reference scenario (Fig. 4.4 left). Here, inequality is measured as the so-called *20:20 ratio* defined as the ratio between the 20th and 80th percentile of the levelized nodal expenditures for electricity. One main reason for this rising inequality are higher LMPE in countries with weaker economies, e.g. Greece, Romania, Serbia, Latvia and Lithuania (Fig. 4.5). In these countries the Human Development Index is below the value of, for instance, Germany and France [Kovacevic et al., 2018]. Higher LMPE would put an additional burden on electricity consumers in these countries, which potentially suffer from economic hardship already and can, therefore, hamper the acceptance of renewables and the mitigation towards climate goals.

Overall, it has been demonstrated that diverging WACC lead to an increased inequality in

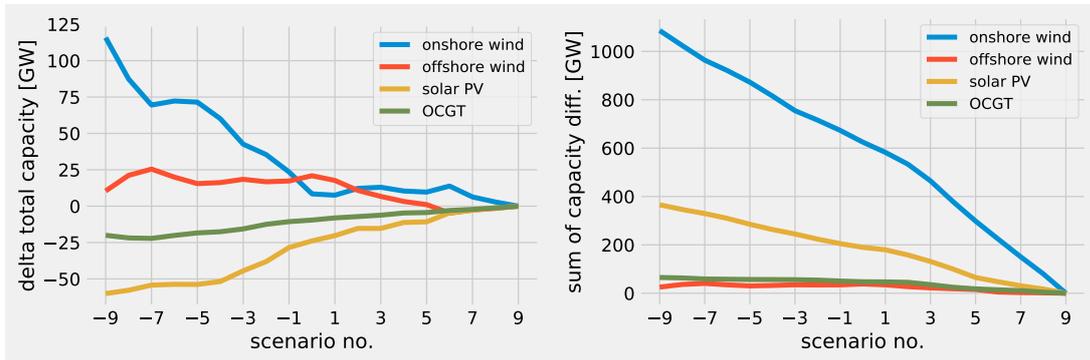


Figure 4.6: Left: Change in overall power generation capacity deployment for various generation sources; Right: Sum of absolute nodal differences in power generation capacity deployment for various generation sources.

regional expenditures for electricity and a reduction of levelized costs of electricity – driven by changes in regional investment and resulting changes in regional generation and electricity penetration. We also explained that wind power profits from low cost of capital in wind-rich, high-demand countries such as Germany, France and Belgium. If WACC diverge further, this effect accelerates, leading to additional 116 GW onshore wind power capacity for the inhomogeneous scenario compared to the homogeneous case (Fig. 4.6 left). Offshore wind power, however, remains more or less stable around an additional 11 GW. The deployment of gas and PV capacity simultaneously falls. Besides these changes in overall capacity expansion, there is a significant redistribution of deployment between countries (right panel of Fig. 4.6). This redistribution is most pronounced for onshore wind power: Compared to the homogeneous scenario, more than 1 TW are shifted from one country to another, expressing the agglomeration of wind power in Germany, France and Belgium mentioned earlier. Additionally, the inequality in regional expenditures for electricity rises up to 16 % in the inhomogeneous case compared to the homogeneous scenario (Fig. 4.4 left).

## 4.4 Discussion

Inline with Hirth and Steckel [2016] our study emphasizes the importance of cost of capital in the context of fostering the integration of renewable generation sources into future power systems. We show that regional differences in weighted average cost of capital lead to significant changes in the optimal design of a European power system with an ambitious CO<sub>2</sub> reduction target when compared to a setup with homogeneous WACC. The latter is an often made assumption in power system modeling and – as we explained – potentially leads to wrong conclusions concerning the optimal system design and the spatial distribution of costs. Schlachtberger et al. [2017] for instance assumed a constant return rate of 7 % across Europe. They reported an optimal cost share for solar PV of at least 30 % (depending on the global transmission capacity limit) in Hungary. Considering the relatively high WACC in Hungary we find an optimal share of solar PV of less than one per cent of the annualized system costs. Similarly, Schlachtberger et al. [2017] found a PV cost share of 40 % to

approximately 75 % in Austria. In our study, Austria exhibits a relatively low WACC. Together with its position within the network and good solar resources this leads to a PV cost share of 91 %. Furthermore, the assumption of homogeneous WACC across Europe leads to too conservative estimates of the levelized costs of electricity, mainly caused by too high estimates of the levelized costs for wind power.

The relative increase of wind power in Europe in the case of inhomogeneous WACC also modifies the variance of the generation time series. Wind power in general exhibits lower diurnal variability and a seasonal cycle opposite to the seasonal cycle of solar PV [Heide et al., 2010]. Hence, modifications of the generation mix lead to different requirements for flexibility as well. Consequently, our results – at least partially – change the interpretation of other studies investigating the optimal generation mix, the need for flexibility options and/or the interplay of generation and storage under the assumption of homogeneous cost of capital across Europe: e.g. [Brown et al., 2018, Schlott et al., 2018, Tranberg et al., 2018]. Schlott et al. [2018], for instance, found an increasing importance of PV power in Europe under different climate change scenarios. This effect might be weakened when the relative high costs for investments in the Southern European countries are taken into account. Brown et al. [2018] describe how the integration of battery electric vehicles (BEV), long-term thermal energy storage (LTES) and power-to-gas units (P2G) helps smoothing the variability from solar and wind power generation in a sector-coupled European power system. In particular, BEV interact with the diurnal variability of solar power and load, while LTES and P2G mainly balance the synoptic to seasonal variations. These findings are confirmed by Tranberg et al. [2018]. Since synoptic variations are more pronounced for wind power (for solar power the diurnal cycle dominates), an increased share of wind power might favour LTES and P2G while decreasing the importance of short-term storage, such as batteries and BEV.

Some critical remarks: Although our model setup is inline with a number of similar studies, the concept of foreign direct investment (FDI) is not considered. Wall et al. [2019] describe how policies influence FDI in renewable energies. Accordingly, increasing FDI has supported the global expansion of renewable energies. The main source of FDI is Europe, in particular Germany and Spain, investing in renewable energy projects in the remaining European countries [Hanni et al., 2011]. Such effects cannot be covered by the chosen model design. However, unlimited FDI could indeed be simulated by assuming the minimum WACC within the region of interest for all regions, meaning that all capital needed to make investment would be acquired in the region of minimum WACC. This resembles the setup of the homogeneous scenario and, thus, does not affect the general findings of this work. Furthermore, differences in WACC for renewable energy are mostly caused by tariff-related risks which apply for foreign direct investment as well [Temperton, 2016]. As mentioned in Section 4.2.3, WACC reflects varying market conditions between regions. Changing amounts of investment should not alter this. Concerning the generation and storage technologies available for expansion, we follow the conservative assumptions made in Brown et al. [2018]. In particular, other dispatchable zero emission technologies such as biomass and geothermal are not taken into account. Combined cycle gas turbines could in general be included. But since they would only partially replace OCGT, keeping the share of variable renewable generation sources unchanged, the inclusion would not affect the overall findings. Consequently, we excluded them in order to keep the model setup as simple as possible and the computation time appropriate

as recommended by DeCarolis et al. [2017].

## **4.5 Conclusion**

Because power system modeling is an important tool for policy advise and system planning, the validity of the underlying assumptions is of crucial importance. Additionally considering the overall importance and urgency of establishing low-emission power systems to tackle climate change and the fact that no homogenization in WACC has been observed in Europe, we stress that the regional inhomogeneity in WACC needs to be taken into account in future studies or should at least be considered as a potential scenario. Furthermore, the uncertainty of the input parameters in power system models should be respected more carefully. Future work might focus on investigating the effects of this uncertainty on the outcome of power system optimization models.



## Chapter 5

### Research Question 3: How large is the sensitivity of power system expansion models?

This chapter has been published as:

*Schyska, Bruno U., Kies, Alexander, Schlott, Markus, von Bremen, Lueder and Medjroubi, Wided (2021): The Sensitivity of Power System Expansion Problems, Joule, 5(10), doi: 10.1016/j.joule.2021.07.017*

---

Power system expansion models are a widely used tool for planning power systems, especially considering the integration of renewable resources. Studies using these models form the basis for far-reaching political decisions. The backbone of power system models is an optimisation problem, which depends on a number of economic and technical parameters. Although these parameters contain significant uncertainties, a consistent way to quantify the sensitivity to these uncertainties does not exist. Here we analyse and quantify the sensitivity of power system expansion models to different model parameters and model definitions based on a novel misallocation metric. While measuring the sensitivity as a simple cost term, this metric can contribute to openness and transparency in power system research. We demonstrate its applicability by three prominent test cases: the definition of capital cost, different weather periods and different spatial and temporal resolutions.

---

#### 5.1 Aims and Motivation

As described in Section 1, many solutions proposed to ease the integration of renewable resources into existing power systems are the result of studies using power system expansion models (PSEM, Section 2). Over the last years, PSEM have become progressively more complex. They include more and more aspects of power systems in an increasing level of detail. Furthermore, they depend on a number of uncertain parameters. Assumptions made for costs or the availability of weather-dependent generation sources as well as the reduction of the model resolution to make models tractable introduce uncertainty Trutnevyte [2016], Mavromatidis et al. [2018], Schlachtberger et al. [2018]. Recently, Nacken et al. [2019] applied a method called *modelling to generate alternatives* (MGA) to a future German energy supply and showed that it produces a number of significantly different energy scenarios. MGA is based on changing the PSEM structure by setting the cost-optimal objective value plus

some slack as a new constraint and exploring the resulting solution space Brill Jr et al. [1982], Price and Keppo [2017], DeCarolis et al. [2016]. Neumann and Brown [2019] used a similar method to explore the near-optimal solution space of a cost-optimised European power system. They observed a high variance in the deployment of individual components near the optimal solution. Based on a *global sensitivity analysis*, Moret et al. [2017] found that the uncertainty of economic parameters has the highest influence on the results of an energy model. Similarly, Shirizadeh et al. [2019] investigated the robustness of a fully renewable power system model of France to uncertainties in future cost of generation technologies. They found that, although the optimal generation mix clearly depends on the respective cost for the different technologies, overall system costs are relatively insensitive. For an overview of methods applied in the context of uncertainty in power system modeling see Yue et al. [2018].

As the results of power system studies often are the basis for far-reaching political decisions and societal discussions, data and methods behind these results must be trustworthy and challengeable Pfenninger [2017b]. Hence, quantifying and documenting uncertainties and their implications on power system studies is crucial. However, a clear and concise way to do this does not yet exist.

The aim of this paper is the analysis of the sensitivity of a common power system expansion model to different model parameters and designs. We quantify this sensitivity via a novel misallocation metric. It expresses the sensitivity as the additional cost arising from misallocating generation capacities and allows to cross-validate input data and to find representative data sets. By quantifying the sensitivity to three prominent parameter scenarios this study contributes to increasing transparency in power system research.

## 5.2 Defining a misallocation metric

Consider a linear program of the form:

$$\begin{aligned} \min_x \quad & c^T x \\ \text{s.t} \quad & Ax \geq b \\ & x \geq 0 \end{aligned} \tag{5.1}$$

In order to solve this program, the parameter matrix  $A$ , the parameter vector  $b$  and the objective coefficients  $c$  need to be defined. They determine the actual problem. If  $A$  and/or  $b$  and/or  $c$  are modified, the solution of the linear program changes depending on the sensitivity to the respective parameters (compare Section 2.5).

Now, let  $x_{\alpha_i}^*$  be a realisation of the random variables  $\{x_1, x_2, \dots, x_{|x|}\}$  solving the linear program under a given set of parameters (in the following referred to as *scenario*)  $\alpha_i$ . Furthermore, let  $\Gamma_0^{\alpha_i}$  be the optimal (minimum) value of the objective function, i.e.

$$\Gamma_0^{\alpha} = c^T x_{\alpha_i}^* \tag{5.2}$$

If one is interested in the effect of using another different set of parameters expressed as scenario  $\alpha_j$ , one could, for instance, measure the difference in the objective function value

$\Delta\Gamma = \Gamma_0^{\alpha_i} - \Gamma_0^{\alpha_j}$  or the Euclidean norm of the optimal values of the decision variables,  $\Delta x^* = \|x_{\alpha_i}^* - x_{\alpha_j}^*\|$ . The difference between the two solutions can then be distinguished into four cases:

1. small  $\Delta\Gamma$  and
  - a) small  $\Delta x^*$ : In this case, the error is obviously small. The linear problem can be considered *insensitive* to the choice of the two scenarios considered.
  - b) large  $\Delta x^*$ : Here, drawing a clear conclusion is difficult. On the one hand, the error when deciding for one solution might indeed be large. Then, the difference in the objective function value would be small only by chance and we could consider the linear problem to be *sensitive* to the choice of the scenario. On the other hand, the error might be small because the solution space is flat near the optimal point or a secondary optimum exists. The linear problem, again, is *insensitive*.
2. large  $\Delta\Gamma$  and
  - a) small  $\Delta x^*$ : This case is possible if the linear program is very sensitive towards changes in some decision variables or if the parameters strongly differ between both scenarios. We consider the linear problem to be *insensitive*.
  - b) large  $\Delta x^*$ : In this case, the error would obviously be large. The linear problem is *sensitive* towards the input dataset.

In order to determine the actual error originating from considering only one of the two scenarios (e.g.  $\alpha_i$ ), one could set the solution  $x_{\alpha_i}^*$  as lower bound to the respective other linear program from scenario  $\alpha_j$  and vice versa. If the program is insensitive, this should not cause any large additional cost. However, if the program is sensitive, this should cause large additional cost because large adaptations to  $x_{\alpha_i}^*$  are necessary in order to make it a solution of the linear program from scenario  $\alpha_j$ . Let us denote the optimal value of the objective function with lower bounds defined by the optimal solution of the linear program  $\Gamma_{\alpha_i}^{\alpha_j}$  and  $\Gamma_{\alpha_j}^{\alpha_i}$ , respectively. The additional cost, caused by constraining the solution downwards is then given by  $\Gamma_{\alpha_j}^{\alpha_i} - \Gamma_0^{\alpha_i}$  and the overall sensitivity of the linear program to the choice of the scenario can be quantified by the following *misallocation metric*:

$$M_{\alpha_j}^{\alpha_i} = \Gamma_{\alpha_j}^{\alpha_i} - \Gamma_0^{\alpha_i} + \Gamma_{\alpha_i}^{\alpha_j} - \Gamma_0^{\alpha_j} \quad (5.3)$$

As stated above,  $x_{\alpha_i}^*$  refers to the solution of the problem with minimum lower bounds, i.e. the unconstrained problem.  $\Gamma_0^{\alpha_i}$  denotes the corresponding value of the objective function. In order to compute the constrained solution  $x_{\alpha_i, \alpha_j}^*$  and  $\Gamma_{\alpha_j}^{\alpha_i}$ , the solution of the corresponding counter-scenario is set as lower bounds to at least some of the decision variables, which means that the following constraint is added to the linear program:

$$x_i \geq x_{\alpha_j}^*, x_i \in \bar{x} \subset x \quad (5.4)$$

In this paper, these additional constraints are applied to variables representing long-term investment decisions, i.e. generation, storage and transmission capacities.

$M$  is driven by the difference in  $\Gamma_0^{\alpha_i}$  and  $\Gamma_{\alpha_j}^{\alpha_i}$ . When this difference is small, the sensitivity is also small.  $M$  fulfills the properties of a pseudometric, it is positive definite

$$M_{\alpha_j}^{\alpha_i} \geq 0 \quad (5.5)$$

because  $\Gamma_0^{\alpha_i} \leq \Gamma_{\alpha_j}^{\alpha_i}$ , symmetric to the order of the scenarios

$$M_{\alpha_j}^{\alpha_i} = M_{\alpha_i}^{\alpha_j} \quad (5.6)$$

and fulfills the triangle inequality

$$M_{\alpha_j}^{\alpha_i} \leq M_{\alpha_k}^{\alpha_i} + M_{\alpha_j}^{\alpha_k} \quad (5.7)$$

## 5.3 Results

### 5.3.1 The Sensitivity to the Capital Cost of Generation Capacities

One major parameter to PSEM are the capital cost for all generation, storage and transmission assets. Let us, as a first example, assume we would like to investigate the sensitivity of the model to different specifications of these capital cost. In order to do so, we define two scenarios: The first scenario (*hom*) assumes a homogeneous distribution of the cost of capital across the European countries. The second scenario (*dia*) takes regional differences into account. The cost of capital is set as reported from the *diacore* project Noothout et al. [2016]. In this scenario, the rate of return on capital ranges from 12% in the South-Eastern European Countries to only 4% in Germany (Figure 4.1).

Schyska and Kies [2020] have shown that these two scenarios lead to significantly different solutions for the cost-optimal generation capacity layout  $x_{hom}^*$  and  $x_{dia}^*$  and the system cost  $\Gamma_0^{hom}$  and  $\Gamma_0^{dia}$ . In particular, the optimal solution for the *dia* scenario contains a larger share of offshore wind power, while the share of onshore wind power, solar photovoltaics (PV) and OCGT decreases compared to the *hom* scenario. After computing  $\Gamma_{dia}^{hom}$  and  $\Gamma_{hom}^{dia}$ , one finds

$$\frac{M_{dia}^{hom}}{\sum_{n,t} d_{n,t}} = 4.6 \frac{\text{EUR}}{\text{MWh}}$$

with  $d_{n,t}$  being the demand at node  $n$  and time  $t$ . This is 6.5 % (6.4 %) of  $\Gamma_0^{hom}$  ( $\Gamma_0^{dia}$ ). Note that the difference in  $\Gamma_0^{hom}$  and  $\Gamma_0^{dia}$  suggests a smaller error of only 1.4  $\text{EUR}/\text{MWh}$ . By applying our new metric we are able to show that the sensitivity of the power system expansion problem to the regional distribution of the cost of capital is indeed much higher (more than 3-fold) than this difference in the levelized costs suggests.

This higher sensitivity can be explained by taking a look at the shape of the solution space: In general, the solution space for the upper level problem of the expansion problem is steeper than for the lower level problems. This means that slight changes in the capacity layout may lead to significantly different investment cost, while there potentially exist many ways to solve the operational problem with similar cost. This effect is enhanced, if additional regional differences in the cost of capital are considered. Building an offshore wind park

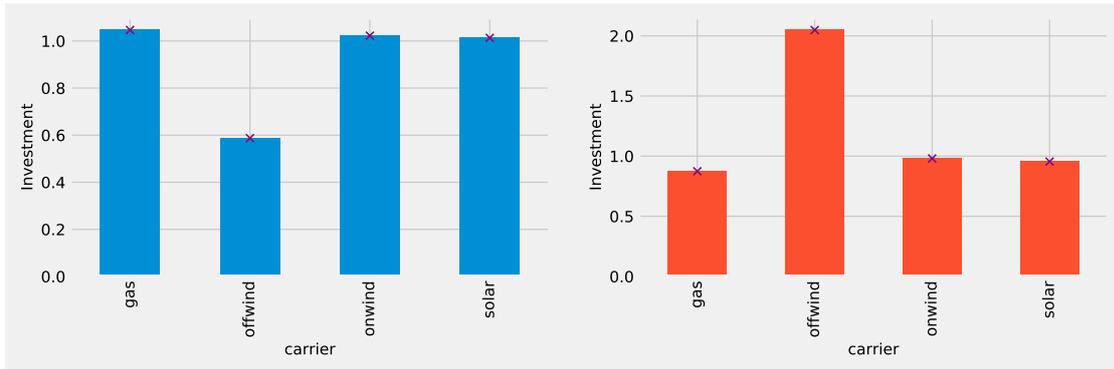


Figure 5.1: Investment in generation capacity relative to the unconstrained solution for the homogeneous scenario constrained with the solution of the inhomogeneous case (left, blue) and for the inhomogeneous scenario constrained with the solution of homogeneous case (right, orange), crosses indicate the minimum investment for each generation source.

in Germany or in Greece, for example, makes a bigger difference now as it made in the homogeneous case. For our first example, we modified the regional distribution of capital cost but kept the nodal loads and the weather time series, which determine the availability of the volatile renewable resources, unchanged. Consequently, we find that both solutions of the unconstrained problems  $x_{dia}^*$  and  $x_{hom}^*$  also solve the operational problems of the respective other problem. This is reflected in the fact, that no changes to the capacity layouts are necessary (Figure 5.1). However, since the two capacity layouts are quite different (as reported above) the investment cost in the constrained cases increase by 6 % and 5 %, respectively, compared to the unconstrained cases due to the different cost assumptions in the two scenarios (Figure 5.2).

### 5.3.2 The Sensitivity to the Capacity Factor Time Series

Second, we investigate the sensitivity of the PyPSA-Eur model to the capacity factor time series  $\bar{g}_{n,s,t}$ . These time series specify the temporal availability of all volatile resources in the power system at any node. They are given in units of the installed capacity, i.e.  $\bar{g}_{n,s,t} \in [0, 1]$ . For wind and solar power the capacity factor is determined by the prevailing weather situation and, as weather changes from time to time, capacity factors vary as well, from hour to hour but also from year to year and from decade to decade due to climate variability and climate change. Depending on the chosen weather period, the power system optimisation might, consequently, lead to different optimal capacity shares. The years 2000-2006, for instance, exhibit a higher average solar power capacity factor over almost entire Europe except for the Iberian Peninsula as predicted for the years 2094-2101. In contrast, the average onshore wind power capacity factor is lower in Central-Western Europe and higher especially in the South-East (Figure 5.3). For this investigation, data from Schlott et al. [2018] has been used. It includes time series of the capacity factors for onshore and offshore wind, solar PV and run-off river as well as of the natural inflow into hydro dams from the climate model *CNRM* Volldoire

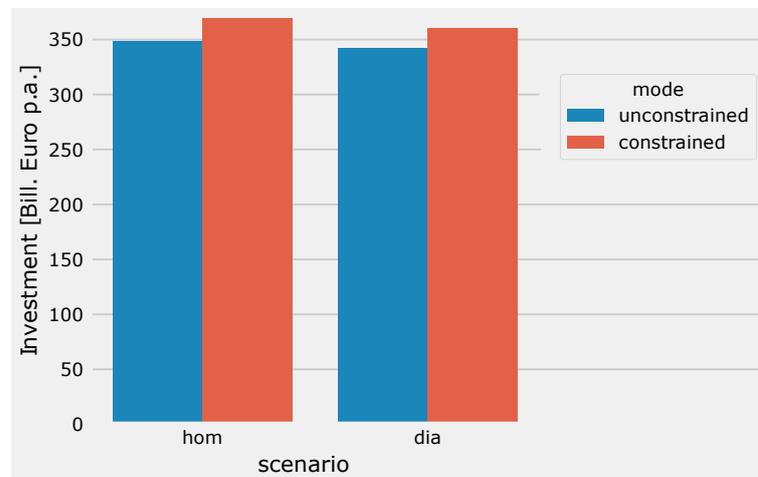


Figure 5.2: Total investment in generation, storage and transmission capacities in the *hom* scenario (left) and the *dia* scenario (right) and the unconstrained (blue) and constrained case (red) [Bill. Euro].

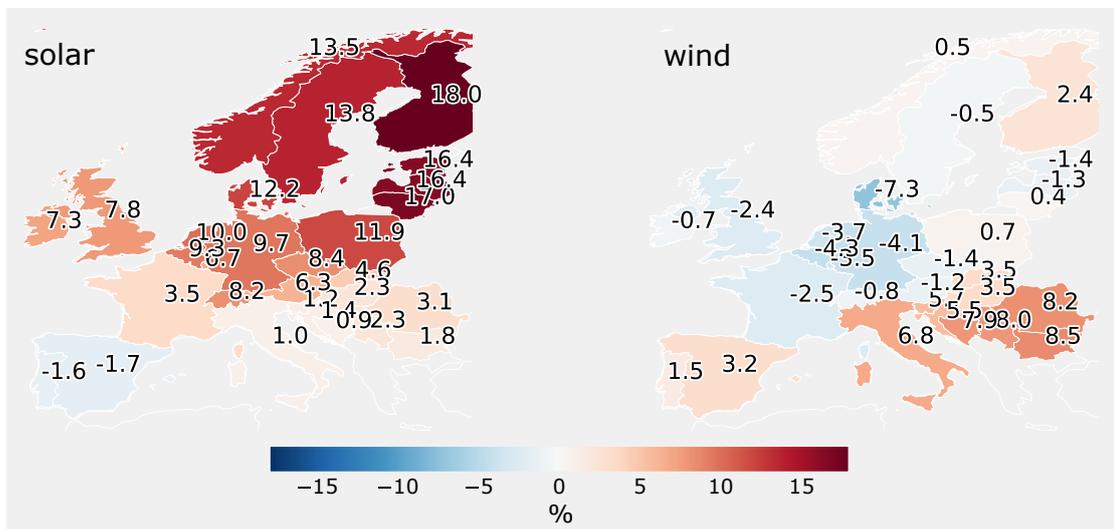


Figure 5.3: Relative difference in solar power capacity factors (left) and onshore wind power capacity factors (right) between the years 2000-2006 and 2094-2101 [% of 2000-2006].

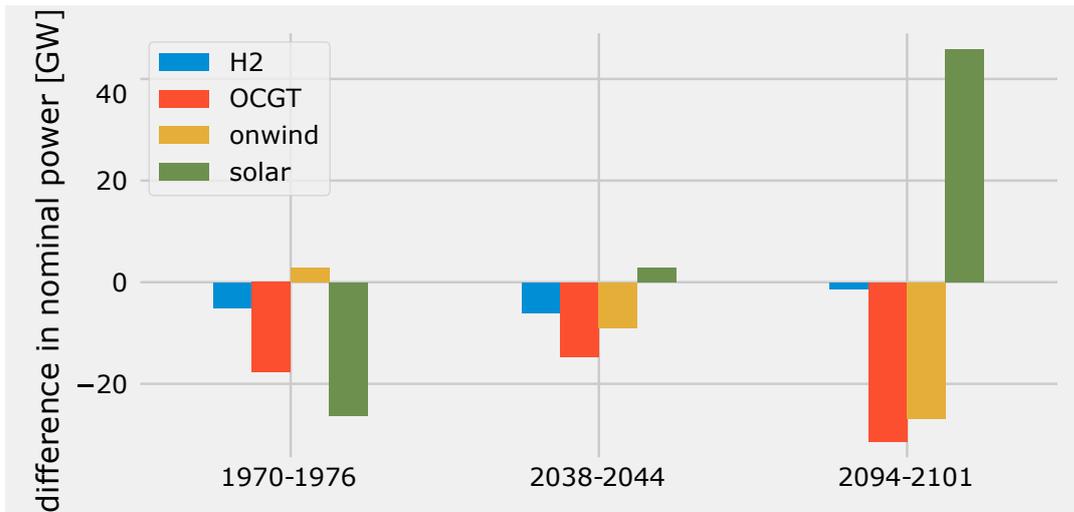


Figure 5.4: Difference in the cost-optimal capacity deployment for three different time slices compared to the years 2000-2006 [GW].

et al. [2013] downscaled to a higher spatial resolution within the EURO-CORDEX project Jacob et al. [2014] for each node of the PyPSA-Eur model in its one-node-per-country setup. We use four time slices of length 6-7 years in 3-hourly resolution: 1970-1976, 2000-2006, 2038-2044 and 2094-2101.

The average capacity factor, however, is only one aspect of the quality of renewables resources. The temporal variability, the spatial and temporal co-occurrence of different resources, the correlation with the electricity demand, the system's ability to distribute generation over large areas and the possibility to interact with different kind of storage technologies are equally important. The complex interaction of all these aspects determines the optimal capacity layout of a power system. Indeed, optimising the PyPSA-Eur model based on either the period 2000-2006 or the period 2094-2101 leads to different optimal investment in generation capacities: Compared to the years 2000-2006, the period 2094-2101 leads to an increase in investment in solar power capacities of 48 % and a decrease in onshore wind power capacity by 4 % (Figure 5.4). Offshore wind is only marginally deployed and can be neglected in both cases. Interestingly, the levelized cost for all periods differ only slightly. The differences range from 0.16 EUR/MWh for the combination (2038-2044, 2094-2101) to 0.56 EUR/MWh for the combination (1970-1976, 2094-2101). Schlott et al. [2018] found similar results for CNRM. Other climate models lead to more diverse results in the objective function Schlott et al. [2018].

Let us now compute the sensitivity metric. Therefore, we set the solution for the period 2000-2006 as lower bound to the linear problem with the capacity factors taken from 2094-2101 and vice versa. Adding up the differences in the (two) constrained and unconstrained solutions (Equation (5.3)) leads to an overall sensitivity of 5.1 EUR/MWh (Figure 5.5). This is the highest sensitivity for all possible combinations of the four weather periods considered. The second highest sensitivity is observed for the combinations (1970-1976, 2000-2006) and (2000-2006, 2038-2044). And the combination (1970-1976, 2038-2044) exhibits the

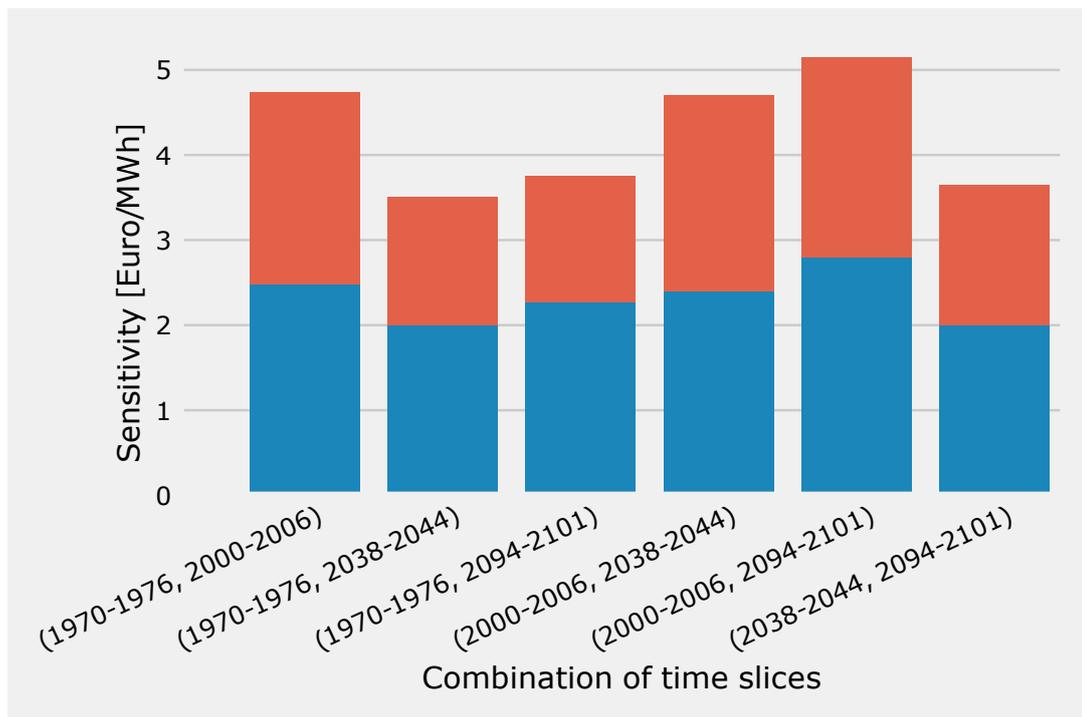


Figure 5.5: Sensitivity to different weather periods quantified by the misallocation metric  $M$  [EUR/MWh]. Blue bars indicate the difference between the constrained and the unconstrained problem of the respective earlier period, the red bar indicates the difference for the later period. Note that for all combinations the difference in the unconstrained solutions is below 0.6 EUR/MWh.

lowest sensitivity of 3.5 EUR/MWh. These findings additionally allow to make an inference on the suitability of using a specific climate period for power system investigations. Since the differences between 2000-2006 and the other time periods is the largest, the period 2000-2006 seems not to be representative and hence not a good choice for power system modelling.

In contrast to the introductory example of modifying the regional distribution of capital cost, we now fixed the cost and varied the availability time series for the different volatile generation sources. In this case, the capacity layout obtained from one optimisation is no longer necessarily able to ensure the supply under the constraints of another optimisation problem based on another weather period. If this would be the case, setting one solution as lower bounds to another scenario would not cause any major additional cost, despite possibly some relatively small changes in the operational cost due to the less effective use of generation, storage and transmission capacities (e.g.). The problem would be insensitive to the differences in the capacity factors from the two scenarios considered,  $M$  close to zero. However, for the weather periods (2000-2006) and (2094-2101) this is not the case. As mentioned above, optimising the power system based on (2094-2101) leads to a more solar dominated system, adopted to climate change. Setting this relatively high solar PV

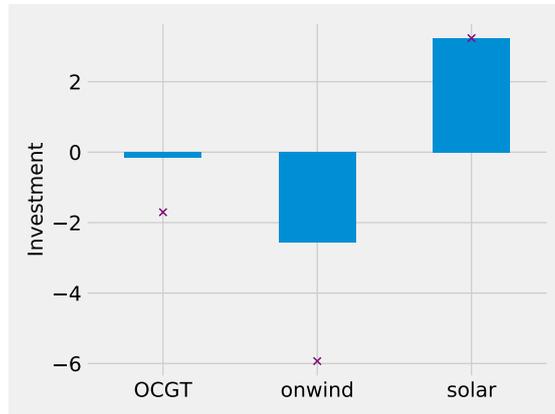


Figure 5.6: Difference in investment in generation capacity compared to the unconstrained solution for the weather period (2000-2006) constrained with the solution of the weather period (2094-2101). Crosses indicate the minimum investment for each generation source.

capacity as lower bound to the optimisation based on (2000-2006) causes major changes in the optimal capacity deployment of the other generation sources as well: The optimal investment in onshore wind power capacity decreases by more than 2.5 Bill. Euro, the investment in OCGT capacity by approximately 0.2 Bill. Euro (Figure 5.6). This leads to an increase in the *levelized cost of electricity* (LCOE) of 8 % of the constrained solution compared to the unconstrained solution. This increase in LCOE is the dominating term in Equation (5.3) and determines the sensitivity of the investigated linear problem to the scenarios considered.

### 5.3.3 The Sensitivity to Reduced Spatial and Temporal Resolution

As a third demonstration example, we investigate the sensitivity of the linear program (2.1) to the temporal and spatial resolution. In order to reduce the spatial resolution, the original PyPSA-Eur network has been scaled down to 45, 64, 90 and 128 nodes using the network clustering approach introduced by Hörsch and Brown [2017]. The temporal resolution has been reduced by averaging the parameter time series over consecutive time spans of length  $\tau \in \{3, 6, 12, 24\}$  hours as described in Section 5.6.2.

The misallocation metric  $M$  for all possible combinations of these different parameter sets sizes exhibits a clear pattern (Figure 5.7). Basically, it can be divided into three different blocks: two blocks of (relatively) low sensitivity where  $M \leq 4.2$  EUR/MWh and one of (relatively) high sensitivity where  $M \geq 7.2$  EUR/MWh. The first block of low sensitivity contains all combinations of scenarios with a temporal resolution higher than 6 hours, i.e.  $(N, 1H)$ ,  $(N, 3H)$  and  $(N, 6H)$ , independent from the spatial resolution  $N$ . The second block of low sensitivity contains all combination of scenarios with a temporal resolution smaller than 12 hours – again independent from the spatial resolution. And the block of high sensitivity contains all combinations of scenarios where one scenario has high ( $\leq 6$  hours) temporal resolution and the other scenario has low temporal resolution ( $\geq 12$  hours).

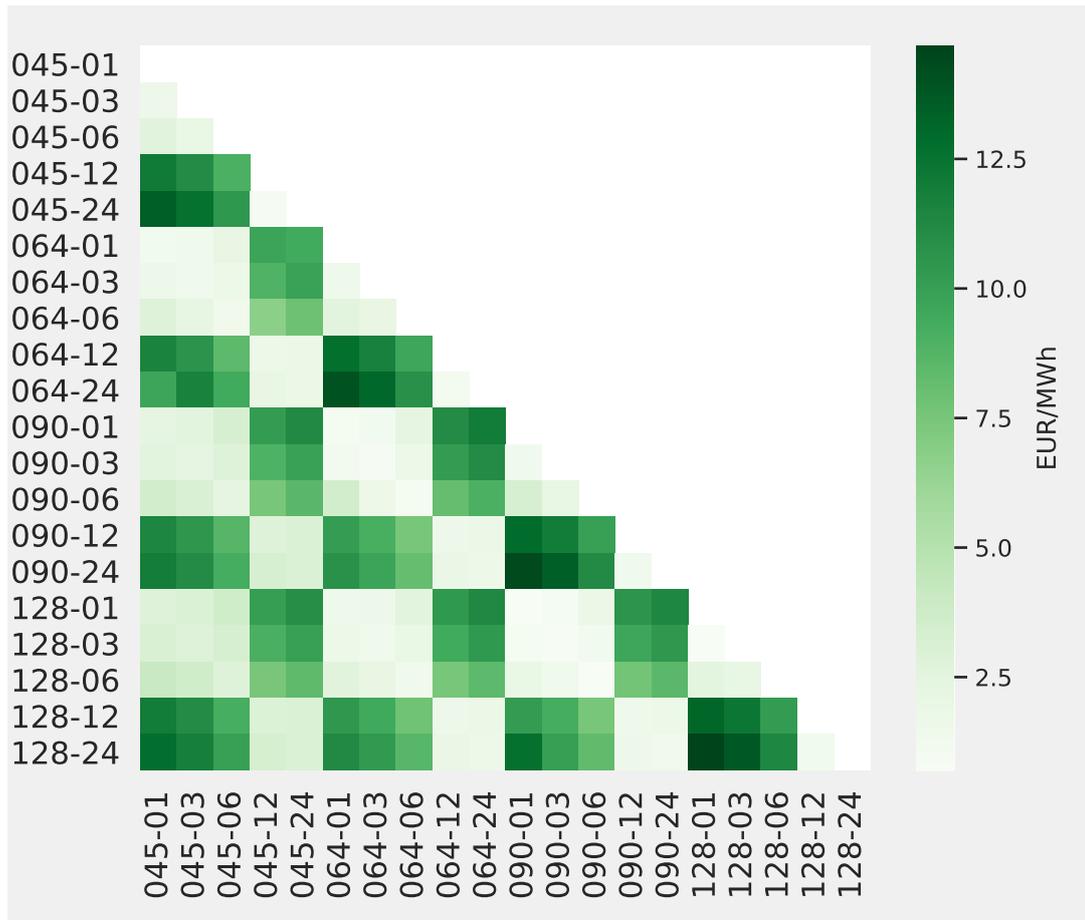


Figure 5.7: Sensitivity heatmap [EUR/MWh]; in the tick labels, the number left of the hyphen indicates the number of nodes, the number right of the hyphen indicates the temporal resolution.

From this definition of blocks one can already see that the expansion problem is much less sensitive to changes in the spatial resolution as it is to changes in the temporal resolution. For instance, the sensitivity of the problem with hourly temporal resolution to increases in the spatial resolution from 45 nodes up to 128 nodes is below 4 EUR/MWh. In contrast, the sensitivity of the 128 node setup to reductions in the temporal resolution from hourly to minimum 12-hourly reaches a maximum value of almost 13 EUR/MWh.

The reasons for this are twofold. First of all, increasing the spatial resolution does not necessarily lead to a higher degree of information in the time series, and vice versa. Consequently, the results obtained from models with different spatial resolutions do not differ much. This phenomenon is of meteorological nature: Hourly wind power and solar PV capacity factor time series exhibit large correlation lengths. Consequently, aggregating nodes, which are geographically close, does not lead to a significant loss of information about the temporal characteristics of the aggregated nodes. For a detailed investigation of the correlation lengths of wind and solar PV in the PyPSA-Eur model see Section 6.1. In contrast, modifying the temporal resolution potentially leads to significant differences in the optimal capacity deployment, especially when the temporal resolution 'jumps' from one of the blocks we defined above to another. Main reason for this is, that downsampling the time series via averaging removes part of the temporal variability. In general, a rolling window averaging can be understood as a filter. For instance, averaging a time series with a rolling 24 hour window filters out most of the sub-24 hour variability of the time series, including the diurnal cycle (if present). If such a filter is applied to both the capacity factor time series and the demand time series, the residual load gets implicitly filtered as well. As a consequence, any storage technology meant to flatten the sub-12 hour variability of the residual load time series would no longer be needed (because there is no sub-12 hour variability).

In general, storage technologies can be assigned to a characteristic variability in the residual load time series via their energy-to-power ratio  $r$ , i.e. the number of hours they can store (dispatch) electricity at full power when starting from empty (full) storage. In the setup used here, batteries are characterized by an energy-to-power ratio of  $r = 6$  hours. They are meant to balance discrepancies between demand and availability which occur on the intra-day scale. As described above, these discrepancies disappear when the demand and capacity factor time series are downsampled to a lower resolution. Consistently, no battery storage devices are optimally deployed in the model setups with a temporal resolution below 6 hours (Figure 5.8). In contrast, hydrogen cavern storage units exhibit an energy-to-power ratio of  $r = 168$  hours, making them a *weekly* storage. As the weekly variability is still present in the downsampled time series, hydrogen storage devices are still useful.

In broad terms, filtering the high-frequency part of a time series' variability can be understood as removing scatter. This in turn also increases the correlation between the time series, again not only between the availability time series but also between the availability and the demand. Apparently, this rise in correlation mainly increases the system-friendliness of solar PV. Its investment share grows from approximately 40 Bill. Euro for the 3-hourly time series to more than 60 Bill. Euro for the 24-hourly time series (Figure 5.8). In turn, the importance of offshore wind power, which is mainly used to cover the baseload in the highly resolved model, decreases, because the filtered time series no longer contain any non-baseload part. The offshore wind power share drops from approximately 23 Bill. Euro to zero. Overall, downsampling time series leads to reduced cost and a significantly different

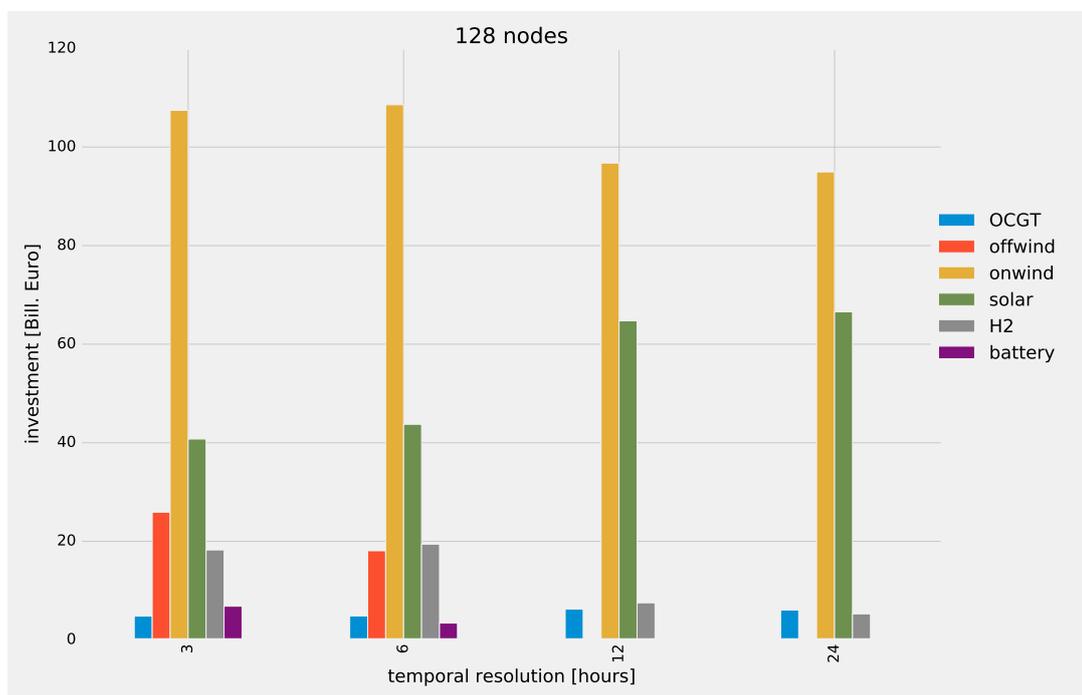


Figure 5.8: Optimal investment in generation and storage capacity [Bill. Euro] for the 128 node network and for different temporal resolutions of the exogeneous parameter time series.

capacity mix. Setting this capacity mix as lower bound to the highly resolved model causes large additional costs, mainly because the model is forced to deploy much more solar PV as it would optimally deploy. Vice versa the offshore and battery storage investment exceeds its optimal value. Overall, this is expressed in a high sensitivity.

However, there is one effect counteracting this phenomenon. This effect appears when the spatial resolution is modified in addition to the temporal resolution. In this case, averaging takes not only part in the temporal dimension but also in the spatial dimension. More precisely, models with a higher spatial resolution experience less averaging on the spatial scale than models with a coarser spatial resolution – assuming that the models' resolutions are in any case below the resolution of the underlying weather data. This potentially leads to higher capacity factors in the highly resolved case. When transmission capacity is sufficiently available and/or the network is sufficiently meshed, higher capacity factors require less generation capacity as the model with lower spatial resolution. Setting these relatively low capacities as lower bounds to the coarser model does not lead to any additional costs because the optimal capacities are above these bounds anyhow. The lower bounds are non-binding. Consequently, the sensitivity is determined by the additional cost arising from setting the optimal capacities of the coarser model as lower bounds to the finer resolved model. Apparently, these additional cost are small compared to the costs arising from modifying the temporal resolution. When the spatial resolution is not modified, both differences in the Equation for the sensitivity metric (Equation (5.3)) are non-zero. This causes the sensitivity between two models of the same spatial but different temporal resolutions, i.e.  $(N, 1H)$  and  $(N, 24H)$  to be larger as between two models of different spatial and temporal resolutions  $(N, 1H)$  and  $(M, 24H)$ .

## 5.4 Discussion

In this study, we introduced a novel method to study the sensitivity of power system optimisation models to different input data scenarios. Core of this method is a metric which is based on setting the decision derived from using one input data set as the lower boundary to the PSEM solving the same program with another parameter scenario. In the sense of modifying and re-solving the original optimisation problem it is comparable to the methods applied by Nacken et al. [2019] and Neumann and Brown [2019]. However, we quantify the sensitivity by one number – the additional cost arising from misallocating generation, storage and transmission capacities caused by using information for long-term planning which differs from the information the model experiences in short-term operation – instead of exploring it visually.

In order to test this methodology, we used a simplified setup of a European power system model. For instance, we limited the available technologies for electricity generation to OCGT, wind, solar and hydro power. Other technologies such as nuclear or combined-cycle gas turbines have not been considered. Furthermore, no coupling of the electricity sector to others sectors has been modeled. However, the explanations for the described sensitivities are rather general. We believe that including more technologies and/or incorporating sector coupling would not influence these general findings and the general applicability of the proposed method.

In future research, it seems reasonable to compare and combine the proposed methods with the MGA approach of Nacken et al. [2019] or the methods to investigate the shape of the solution space proposed by Neumann and Brown [2019] to study the uncertainty of energy system models. Furthermore, the sensitivity to modifications in the temporal resolution could be further investigated by applying the approach of coupling design periods introduced by Gabrielli et al. [2017] or the time series aggregation approach based on hierarchical clustering with connectivity published by Pineda and Morales [2018].

## 5.5 Conclusions

From the results described above we draw the following conclusions:

1. In summary, we found an on average sensitivity to the choice of the underlying weather data. Our results indicate that the period 2000 through 2006 is not suitable for deriving general conclusions about the optimal design of the European power system. It led to the highest misallocation of generation and storage capacities compared to the other periods considered. This finding emphasizes the importance of using representative weather data sets and demonstrates how our metric can be used to identify them.
2. Similarly, the capital cost of generation assets should be defined according to the state of the art. The sensitivity to the geographical distribution of the cost of capital was found to be as high as the sensitivity to the capacity factor time series.
3. As long as the temporal resolution of the underlying time series does not include any information about microscale meteorological processes, the spatial resolution of the power system model is of minor importance. The sensitivity to increases and decreases in the number of nodes is relatively small. Modeling the European power system with only a few dozens of nodes seems reasonable. This may change when more detailed information about the regional distribution of the demand and demand-side flexibilities (e.g.) are included.
4. In contrast, the temporal resolution of the underlying time series must be chosen carefully, especially with storage devices involved. The power system model shows the highest sensitivity to modifications of the temporal resolution across the characteristic storage horizon of the storage devices. As a conclusion, the temporal resolution should be chosen such that the variability, which the storage devices are supposed to balance, is well represented. Particularly, the temporal resolution should be greater than 6-hourly when daily storage units – such as batteries – are considered. Contrarily, time series with daily resolution might be appropriate when only weekly and/or seasonal storage types are part of the model.

We showed that common PSEM exhibit significant sensitivities. Considering the potential political and societal impacts of power system studies, it appears crucial to quantify and report these model sensitivities and uncertainties along with the model results.

## 5.6 Experimental Procedures

### 5.6.1 Power System Model and Data

In this study, we investigate the sensitivity of a common power system expansion problem to (i) different capital cost assumptions (ii) the capacity factor time series for the available volatile renewable resources, (iii) different temporal and spatial resolutions as well as (iv) different model formulations for coupled and decoupled representative periods of different lengths. For a mathematical formulation of PSEM please see Section 2.1.

Upper bounds for system-wide transmission capacities and CO<sub>2</sub> emissions are defined in Equation (2.7) and (2.6), respectively. For this paper, we assumed a global limit of three times today's net transfer capacities (3·31.25 TWkm) as an appropriate compromise between cost-optimal extension and technical and social concerns. In line with European emission reduction targets, CO<sub>2</sub> emissions are limited to 5 % of the historic level of 1990.

For this study, we use the *PyPSA-EUR* model published by Hörsch et al. [2018a]. In its full spatial and temporal resolution this model consists of 3567 substations and 6047 transmission lines and covers one year in hourly resolution. It includes time series of capacity factors for onshore and offshore – where applicable – wind power as well as solar PV power and time series of electricity demand for each substation. Furthermore, time series for the inflow into hydro reservoirs and runoff river power plants, based on a potential energy approach Kies et al. [2016g], and upper bounds for the extendable generation capacity per renewable technology and substation are included.

Capacity factor time series are commonly derived from reanalyses data sets [Jurasz et al., 2020]. In *PyPSA-Eur*, time series for wind power capacity factors and the inflow to hydro-electric power plants are derived based on the ERA5 reanalysis [Copernicus Climate Change Service (C3S), 2017]. Onshore and offshore wind power capacity factors have been computed using the power curves of a 3 MW Vestas V112 with 80 m hub height and the NREL Reference Turbine with 5 MW at 90 m hub height, respectively. Solar PV capacity factor time series have been computed from the Heliosat (SARAH) surface radiation data set [Pfeifroth et al., 2017] using the electric model of Huld et al. [2010] and the electrical parameters of the crystalline silicon panel fitted in the same publication. All solar panels are assumed to face south at a tilting angle of 35 degrees. Hourly electricity demand for all European countries has been obtained from the *European Network of Transmission System Operators* (ENTSO-E) [ENTSO-E, 2012] and assigned to substations via a linear regression of the GDP and the population. Upper limits of generation capacities have been derived by restricting the available area to *agricultural areas* and *forest and semi natural areas* given in the *CORINE Land Cover* data set [Corine Land Cover, 2017] and by excluding all nature reserves and restricted areas [Natura 2000, 2016]. From the available area, the maximally extendable generation capacity has been computed via fixed densities of 3 MW per square kilometer for onshore wind and 1.45 MW per square kilometer for solar PV, respectively. For further details on the data set and the underlying methodology please see Hörsch and Brown [2017].

From this data set, the parameters for the corresponding PSEM (2.1)-(2.15) have been defined. Therefore, we fixed the nominal power of all hydro power plants and pumped hydro storage units to the values reported by Kies et al. [2016b], while the nominal power of wind, solar PV and open-cycle gas turbine (OCGT) power plants can be expanded within given

bounds. Additionally, we consider two generic storage types with fixed power-to-energy ratio  $r$ :

1. batteries:  $r = 6$  h
2. hydrogen storage:  $r = 168$  h

Their nominal power can be expanded as well. For each technology the investment and operational costs depicted in Table 2.1 have been used.

In full resolution, this model can hardly be solved (compare Section 2.2). Therefore, we use the network clustering algorithm introduced by Hörsch and Brown [2017] to derive clustered versions of the original data set with 45, 64, 90 and 128 substations, respectively. The time series aggregation method described in section 5.6.2 is then applied to these clustered networks.

Although most PSEM aim at finding optimal solutions for power system design, they may significantly vary in structure and in scope. For a list of models see for instance the Open Energy Platform (<https://openenergy-platform.org/factsheets/models/>).

## 5.6.2 Reducing the Spatial and Temporal Resolution of PSEM

In general, two types of time series aggregation methods can be distinguished. The first one aims at decreasing the number of time steps by reducing the resolution of the parameter time series. The *downsampling* approach described below, for instance, can be assigned to this class.

The second class aims at decreasing the number of time steps, while keeping the temporal resolution unchanged. In this way, as much of the temporal variability as possible shall be conserved. Usually, this is achieved by selecting a limited number of representative *design periods* from the original time series. Depending on the periods' lengths the variability on different temporal scales can be retained. This, of course, breaks the natural order of the time steps and, consequently, no variability on time scales longer than the periods' lengths can be pictured. Hence, ways need to be found, which allow to model the variability on long time scales (months - seasons), which is represented by the natural inflow into hydro power plants or the seasonal cycle in electricity demand, e.g.. For an overview of these methodologies see Pfenninger [2017a] and Kotzur et al. [2018].

In order to account for different time step intervals, weightings need to be defined for each time step considered in the expansion problem: first, in the objective ( $w_t$  in Equation (2.1)) and second, in the definition of the storage units' state of charge ( $\omega_t$  in Equation (2.4)).

For this study, we applied a simple downsampling technique. It averages the original exogenous parameter time series over consecutive time spans of length  $\tau$ . Hence, it yields  $T/\tau$  time steps at constant intervals. The snapshot weightings  $w_t$  and  $\omega_t$  are set to  $\tau$ .

The spatial resolution of the PSEM is modified by applying the network clustering approach introduced by Hörsch and Brown [2017]. The original model is clustered to 45, 64, 90 and 128 nodes.

### 5.6.3 Computing the misallocation metric

For each of the parameter sets  $\alpha_i$  the expansion problem (2.1)-(2.15) is first solved without any lower bounds to the nominal power. The resulting solution vector for the cost-optimal generation capacities  $G_{n,s}^*$  is then set as the lower bound to the nominal power for the respective partner problem  $\alpha_j$ :

$$[G_{n,s}]_{\alpha_j}^{\alpha_j} \geq [G_{n,s}^*]_0^{\alpha_i} \quad (5.8)$$

Following this procedure in both directions delivers the terms of Equation (5.3).

In case the number of substations of the two parameter sets differs, i.e.  $N_i \neq N_j$ , the lower bounds for each parameter set are computed from the corresponding cluster of buses of the other parameter set: Let  $\mathcal{N}_i = \{S_{i,1}, S_{i,2}, \dots, S_{i,m}, \dots, S_{i,N_i}\}$ ,  $\mathcal{N}_j = \{S_{j,1}, S_{j,2}, \dots, S_{j,k}, \dots, S_{j,N_j}\}$  be the two sets of clusters of buses derived from the original full-resolution data set with  $|\mathcal{N}_{i/j}| = N_{i/j}$ . In the clustered networks, each of these clusters  $S$  is merged into one single bus  $n(S_{i,m})$ ,  $n(S_{j,k})$ . Then, the lower bound to a generator of technology  $s$  at a bus of set  $\mathcal{N}_i$  is set to the weighted sum of the optimal capacity of the buses of set  $\mathcal{N}_j$  and vice versa:

$$\left[ G_{n(S_{i,m}),s}^{\min} \right]_{\alpha_i}^{\alpha_j} = \sum_{k=1}^{N_j} w_k \left[ G_{n(S_{j,k}),s}^* \right]_0^{\alpha_i}, \quad \forall S_{i,m}, n(S_{i,m}) \quad (5.9)$$

where the weights  $w_k$  are determined from the number of common nodes of the two clusters  $S_{i,m}$ ,  $S_{j,k}$ :

$$w_k = \frac{|S_{i,m} \cap S_{j,k}|}{|S_{j,k}|} \quad (5.10)$$

Here,  $|S_{i,m} \cap S_{j,k}|$  is the number of nodes, which appears in both clusters, i.e. the clusters' intersection.



## Chapter 6

# Extending Research Questions and Outlook

### 6.1 Correlation Lengths of Wind and Solar Power

In meteorology, there is a clear relation between a phenomenon's characteristic temporal and spatial scale. Usually, one differentiates between the *microscale*, the *mesoscale* and the *synoptic scale*. While the microscale basically includes turbulent motions acting within second to minutes and with a spatial extent of millimeters to centimeters, the mesoscale covers phenomena like thunderstorms, hurricanes, fronts and convective systems occurring within minutes to days and on several kilometers extent. The synoptic scale includes high and low pressure systems possibly lasting for up to several days on 100 to 1000 kilometers extent Grue et al. [2012]. Hence, hourly weather time series – as used throughout this thesis – only contain the variability introduced by mesoscale and synoptic processes. Sub-hourly microscale processes, e.g. the sub-second perturbations of the grid caused by turbulent wind power found by Haehne et al. [2019], are filtered out. As mesoscale and synoptic processes act on relatively large spatial scales, the correlation lengths of wind speed can consequently be up to several hundreds kilometers. For wind speed, Martin et al. [2015] found a correlation length of 273 km in Canada and 368 km in Australia. They computed these correlation lengths for the high frequency, stochastic, part of the time series by applying a high-pass filter and by removing the seasonal cycle prior to estimating the correlation length. However, PSEM have to cope with both aspects of the time series, the high-frequency (stochastic) *and* the low-frequency (deterministic) part. It is the interplay of these two aspects which determines the need for balancing and the optimal capacity share of the respective resources. Low frequency variations, in general, exhibit an even higher correlation length as the stochastic time series. For Europe, Schlott et al. [2018] estimated a correlation length in wind speed of 300 to 700 km which is likely to increase in Northern-Central Europe and to decrease around the Mediterranean towards the end of the century due to climate change. Without applying any data pre-processing we find that the spatial extent of the PyPSA-Eur model (5000 km max) is not sufficient to determine the correlation length of the wind power capacity factor (Figure 6.1 left). We do so by integrating correlation over distance via

$$\xi(r_n) = \sum_{k=2}^n \frac{1}{2} (r_k - r_{k-1}) (\rho_{r_k} + \rho_{r_{k-1}}) \quad (6.1)$$

for the PyPSA-Eur 1024 node setup. For calculating the correlation length, the pairwise distance  $r$  and correlation  $\rho$  between all nodes have been computed. Correlation data has then been sorted according to node separation. As  $\xi(r_n)$  does not saturate until the largest

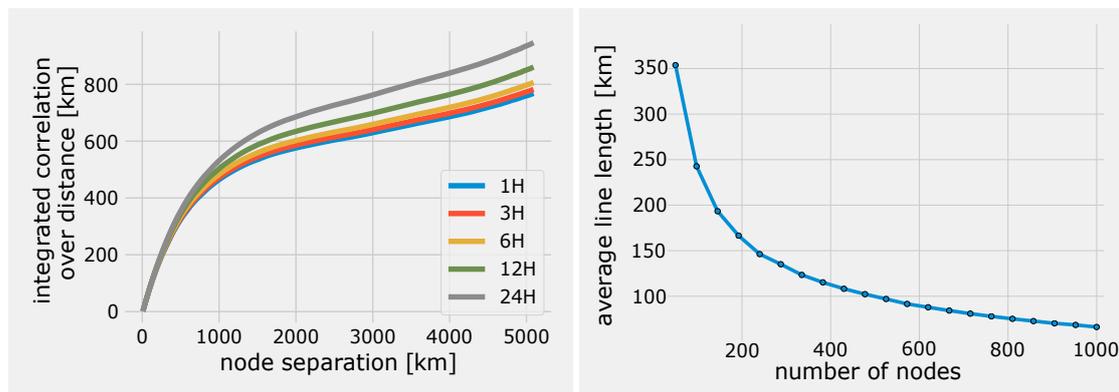


Figure 6.1: Left: Integrated correlation data of wind power capacity factor time series for the PyPSA-Eur 1024 node setup and different temporal resolutions; the value at largest node separation is referred to as *correlation length*; correlation data is integrated over separation using the trapezoid rule, for further details see Martin et al. [2015]. Right: Average line length versus number of buses obtained from clustering the original 380-kV network.

separation, the derived correlation length of approximately 670 km (the integrated correlation at the largest node separation) still is an underestimate of the correlation length which the model experiences.

Compared to wind power, the time series of the solar power availability exhibit a larger deterministic component: the diurnal cycle. On the other hand, incremental changes might be larger. The transition from cloud (shadow) to sun (light) is potentially faster than the transition from windy times to less windy times. As sunrise and sunset occur at the same time over large geographical areas, the correlation length of solar power is even larger as the correlation length of wind power.

For the PyPSA-Eur network, the average distance between two nodes varies between 60 km for 1000 nodes and 350 km for the 45 node setup (Figure 6.1 right). These distances are far below the estimated correlation length and although the amount of meteorological information lost depends on the distance of the aggregated nodes, the loss of information – at least of the kind of information which is relevant for investment decisions – when buses are aggregated is comparably small. As described in Section 5.3.2, this is expressed in a relatively small sensitivity to the spatial resolution. Presumably, this would only change when time series with a higher temporal resolution would be used or when an even smaller number of buses would be considered. Time series with a higher temporal resolution would include a higher share of high frequency variability originating from microscale meteorological phenomena. As these phenomena act on smaller spatial scales, the correlation lengths would decrease, too. When the number of buses would be reduced further, the distance between the aggregated nodes might exceed the correlation length.

Additionally to their temporal variability, time series can be described by their amplitude. In the context of renewable resource assessment, the quality of the resource is commonly described by the average or the sum of the capacity factor time series, the latter being

referred to as *full load hours*. As shown in Section 5.3.2, this quantity varies as well but with an ever lower frequency: on seasonal to climatological scales. The spatial variance of the full load hours is to a large extent determined by the orography and the latitude. Locations close to the shore, for instance, generally exhibit higher wind power full load hours as locations upcountry. Locations far North are less sunny as locations in the South and, hence, exhibit lower solar power full load hours. Consequently, there is no clear relationship between the distance between two nodes and the difference in the full load hours and the effect of aggregating nodes is hard to assess. It depends on the specific location.

## 6.2 Tracing the Flow of Electricity

In Section 2.3 I showed how the concept of Locational Marginal Pricing can be used to allocate cost in a power system. Here, another mechanism based on tracing the flow of electricity shall be introduced: The aim of flow tracing is to express one node's generation  $G_{n,t} = \sum_s g_{n,s,t}$  as a function of the nodal loads  $D_{n,t}$ , i.e. the loads of all other nodes in the network. Starting from the nodal balancing equation

$$G_{n,t} - D_{n,t} = F_{n,t}^{out} - F_{n,t}^{in} \quad (6.2)$$

we write

$$F_{n,t}^{in} + G_{n,t} = F_{n,t}^{out} + D_{n,t} \quad (6.3)$$

and define the nodal through-flows

$$\begin{aligned} \Phi_{n,t}^d &:= F_{n,t}^{out} + L_{n,t} \\ \Phi_{n,t}^u &:= F_{n,t}^{in} + G_{n,t} \end{aligned} \quad (6.4)$$

where  $^d$  and  $^u$  denote the downstream and upstream case respectively [Bialek, 1996, Tranberg et al., 2015]. From Equation 6.3 it can be seen, that  $\Phi_{n,t}^d = \Phi_{n,t}^u$  at any time.

Let  $\mathcal{N}_n$  be the set of neighboring nodes of node  $n$ . We can then define two time dependant subsets of neighboring nodes downstream of node  $n$   $\mathcal{N}_{n,t}^d \subset \mathcal{N}_{n,t}$  and upstream of node  $n$   $\mathcal{N}_{n,t}^u \subset \mathcal{N}_{n,t}$ , respectively. Considering the downstream case, we can write for the load of node  $n$

$$\begin{aligned} D_{n,t} &= \Phi_{n,t}^d - F_{n,t}^{out} \\ &= \Phi_{n,t}^d - \sum_{j \in \mathcal{N}_{n,t}^d} |F_{n-j,t}| \\ &= \Phi_{n,t}^d - \sum_{j \in \mathcal{N}_{n,t}^d} a_{nj,t} \Phi_{j,t}^d \end{aligned} \quad (6.5)$$

with  $a_{nj,t} = |F_{n-j,t}|/\Phi_{j,t}^d$ . In matrix form this yields

$$\begin{aligned} \vec{D}_t &= \mathbf{A}_t^d \vec{\Phi}_t^d \\ \text{or} & \\ \vec{\Phi}_t^d &= [\mathbf{A}_t^d]^{-1} \vec{D}_t \end{aligned} \quad (6.6)$$

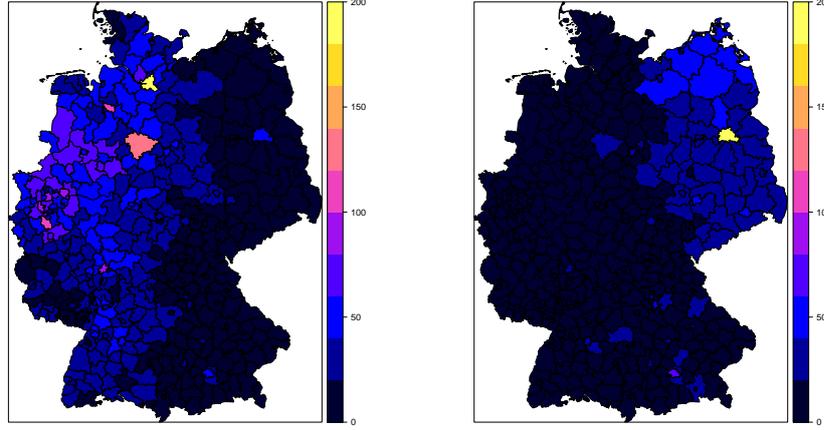


Figure 6.2: Electricity exports to German regions in an idealised German electricity grid from the North Sea (left) and Baltic Sea (right) [ $\sqrt{GWh}$ ]; isolated scenario.

where  $[\mathbf{A}_t^d]^{-1}$  denotes the inverse of matrix  $\mathbf{A}_t^d$  given by

$$A_{ij,t}^d = \begin{cases} 1 & , \text{if } i = j \\ -a_{ij,t} = -\frac{|F_{i-j,t}|}{\Phi_{j,t}^d} & , \text{if } j \in \mathcal{N}_{n,t}^d \\ 0 & , \text{else} \end{cases} \quad (6.7)$$

Here again,  $|F_{i-j,t}|$  is the flow from link  $i$  to  $j$ . Equation 6.6 expresses the nodal through-flows as a function of the nodal loads. For one node  $n$  we get

$$\Phi_{n,t}^d = \sum_{k=1}^N [\mathbf{A}_t^d]_{nk}^{-1} D_{k,t} \quad (6.8)$$

To get a function for the nodal generation, we consider the generation of node  $n$  as an inflow and write:

$$|F_{\rightarrow n,t}| = G_{n,t} = \frac{G_{n,t}}{\Phi_{n,t}^d} \Phi_{n,t}^d = \frac{G_{n,t}}{\Phi_{n,t}^d} \sum_{k=1}^N [\mathbf{A}_t^d]_{nk}^{-1} D_{k,t} \quad (6.9)$$

From Equation 6.9 we see, that each node  $k$  contributes

$$\frac{G_{n,t}}{\Phi_{n,t}^d} [\mathbf{A}_t^d]_{nk,t}^{-1} D_{k,t} =: \Gamma_{nk,t} \quad (6.10)$$

to the nodal generation of node  $n$ . Note, that  $\Gamma_{nk,t}$  has the unit of power. This approach has been introduced by Tranberg et al. [2015] based on the derivations of Bialek [1996] and used by Tranberg et al. [2018, e.g.] to investigate the interplay of different generation and storage assets in a power network. Here, we will exemplarily use this method to trace the flow of electricity in an idealised German electricity network. The results of this investigation could for instance be used for the definition of grid development plans and/or for allocating

the cost for transmission grid expansion to specific generators and/or consumers. Let us consider the same data set as used by Kies et al. [2016e], once in an isolated scenario, i.e. without any connection to neighboring countries, and once in an embedded scenario, where Germany is embedded in a European system. Here, each country is represented by one node.

In this setup, the electricity generated in the North Sea mainly distributes across the Western part of Germany, while the electricity from the Baltic Sea mainly flows towards the Eastern part including the major load centers Berlin and Munich (Figure 6.2). A similar divided picture can be found for the other European countries (Figure 6.3). While the electricity from the North Sea mainly contributes to fulfill the demand in France and the Netherlands, electricity from the Baltic Sea is mainly (outside Germany) consumed in Poland.

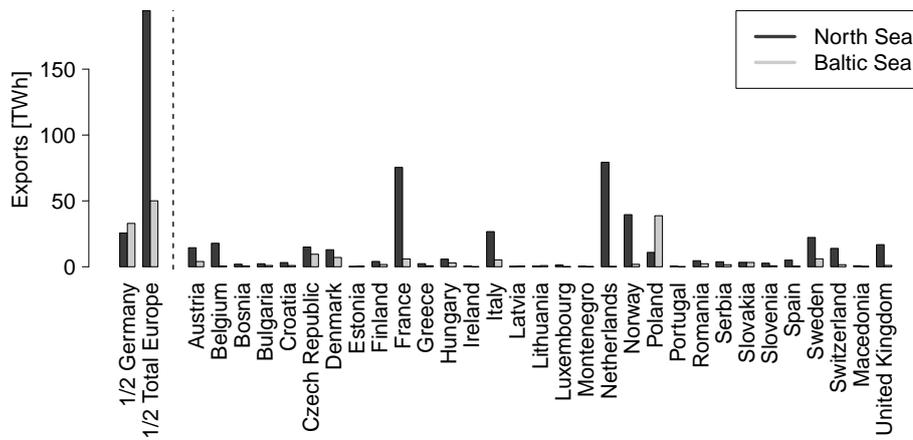


Figure 6.3: Electricity exports to European countries from the North Sea (left) and Baltic Sea (right) [TWh]; idealised German electricity grid; embedded scenario.

## 6.3 Parametric Optimisation

This section has appeared as:

*Schyska, Bruno U., Pinson, Pierre, Kies, Alexander and von Bremen, Lueder (2017): Reducing Power System Expansion Problems via Variable Parameterization, proceedings of the Wind Integration Workshop, Berlin, 2017*

Countries worldwide decide to transform their power systems – from fossil fuels to renewable sources. The question about the optimal pathway along which this transformation shall take place becomes increasingly important. Power system expansion models used to give answers are complex and computationally expensive. To overcome this issue, so called scenario reduction techniques are applied. The number of scenarios, i.e. the number of time steps considered, shall be minimised. In contrast to these techniques, we introduce an approach, which addresses the size of the power system. Installed capacities are estimated as a linear regression of pre-chosen, non-linearly weighted basis functions.

### 6.3.1 Introduction

As explained in Section 2, power system expansion models are optimisation problems aiming on minimising investment costs plus some operation costs. While the investment costs are determined by the investment costs of the different generator types and the total amount of capacity, which shall be installed, operation costs strongly depend on the temporal fluctuation of (i) the generation from volatile sources and (ii) the fluctuation of the demand. Since the temporal fluctuation of the generation is not only determined by meteorology (the fluctuation of wind and solar irradiance) but also by the installed capacities, minimising investment costs and minimising operation costs are coupled by this very capacity (compare Section 2.2: The expansion problem needs to find the optimal trade-off between investment costs and operation costs. The fact that investment and operation levels are coupled, however, yields complex and computationally expensive problems – especially when the power system under consideration is large and when many different generator types shall be considered. Furthermore, the operation part of the expansion problem would ideally be solved for a very high number of scenarios that covers the full range of variability, as described in Section 2.5. To overcome the issue of high computational costs, so called scenario reduction techniques are applied (e.g. Liu et al. [2017], Baringo and Conejo [2013], Pina et al. [2011], compare Section 5.6.2). These techniques aim on minimising the number of scenarios to a feasible number without losing to much information about the variability.

The aim of this section is to present an approach, which – in contrast to scenario reduction techniques – addresses the size of the power system. We developed an approach to reduce the number of decision variables in a power system expansion problem via spatial models: The installed capacities are estimated as a linear regression of pre-chosen, spatially weighted (using Gaussian weighting kernels) basis functions. We show results from applying this approach to a simplistic expansion problem and a model of an idealized German power system consisting of more than 400 nodes and more than 1000 links.

### 6.3.2 Methodology

For this study, we consider a heuristic version of an power system expansion problem as defined in Section 2.1. Instead of assigning costs to the expansion of transmission capacities, we minimise the quadratic flows along the branches. The objective is of the form

$$\min_{\bar{g}, g, f} \sum_{n,s} C_{n,s} \bar{g}_{n,s} + \sum_{n,s} O_{n,s} g_{n,s,t} + a \sum_{l,t} f_{l,t}^2 \quad (6.11)$$

where  $a$  is a tuning parameter to control the influence of the flow term on the optimisation result. Transmission capacities are set to infinity ( $\bar{f}_l = \infty \forall l$ ). This approach can be interpreted as a heuristic way to assign costs to transmission.

In order to reduce the computational costs of this model, we introduce the following approach: Assume the nominal power of renewable power generation technologies can be expressed as a model of the form

$$\bar{g}_{n,s} = S(\mathbf{x}_n, \boldsymbol{\beta}) + \varepsilon_{n,s} \quad (6.12)$$

where

$$S(\mathbf{x}_{n,s}, \boldsymbol{\beta}) = \sum_{i=1}^{|\mathcal{B}|} \sum_{j=1}^J \beta_{ij} \pi_j(\mathbf{x}_{n,s}) b_i(\mathbf{x}_{n,s}) \quad (6.13)$$

and  $b_i(\mathbf{x}_{n,s}) \in \mathcal{B}$  being the  $i$ th basis function evaluated at location  $\mathbf{x}_{n,s}$ ,  $\pi_j(\mathbf{x}_{n,s})$  being a non-negative weighting kernel centered at position  $\boldsymbol{\mu}_j$  evaluated at location  $\mathbf{x}_{n,s}$  and  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_{|\mathcal{B}|J})^T$  being a vector of unknown random coefficients. This approach is similar to the dynamic models for spatio-temporal data introduced by Stroud et al. [2002]. The expansion model can then be reduced by taking into account the following constraint

$$\bar{g}_{n,s} = \sum_{i=1}^{|\mathcal{B}|} \sum_{j=1}^J \beta_{ij} \pi_j(\mathbf{x}_{n,s}) b_i(\mathbf{x}_{n,s}) + \varepsilon_{n,s}, \forall n, s \quad (6.14)$$

Following the idea of least-squares ('choose regression coefficients such that the sum of quadratic residuals is minimised'), we furthermore adopt the objective function as

$$\min_{\bar{g}, g, f} \sum_{n,s} (C_{n,s} \bar{g}_{n,s} + \varepsilon_{n,s}^2) + \sum_{n,s} O_{n,s} g_{n,s,t} + a \sum_{l,t} f_{l,t}^2 \quad (6.15)$$

(added  $\varepsilon_{n,s}^2$  to the first term). Note, that in Equation 6.14 the variables for the nominal generation capacities  $\bar{g}_{n,s}$  are no longer independent but coupled via the  $\beta_{ij}$  – the new decision variables. In this case, the number of *free* decision variables reduces from  $|\mathcal{N}||\mathcal{S}|$  to  $|\mathcal{B}|J$  which reduces the complexity of the linear program.

For this study, we choose linear surfaces as basis functions:

$$\begin{aligned} \mathcal{B} &= \{b_1(\mathbf{x}_{n,s}), b_2(\mathbf{x}_{n,s}), b_3(\mathbf{x}_{n,s})\} \\ \text{with} \\ b_1(\mathbf{x}_{n,s}) &= 1 = \text{const.} \\ b_2(\mathbf{x}_{n,s}) &= x_{n_1,s}, \forall s \\ b_3(\mathbf{x}_{n,s}) &= x_{n_2,s}, \forall s \end{aligned} \quad (6.16)$$

evaluated at locations

$$\mathbf{x}_{n,s} = (\langle D_n \rangle_t, \langle \tilde{G}_{n,s} \rangle_t)^T, \forall s \in \mathcal{S}_n^{RE} \quad (6.17)$$

where  $\langle \dots \rangle_t$  denotes the arithmetic temporal mean and  $\mathcal{S}_n^{RE}$  is the set of volatile renewable resources at node  $n$ . Furthermore, we use spherical Gaussian weighting kernels with uniform variance  $\sigma = 1$

$$\pi_j(\mathbf{x}) = \frac{1}{\sqrt{2\pi}} \exp(-0.5(\mathbf{x} - \boldsymbol{\mu}_j)^2) \quad (6.18)$$

Kernel centers  $\boldsymbol{\mu}_j$  are determined by clustering the  $\mathbf{x}_{n,s}$  using the kmeans [MacQueen et al., 1967] algorithm with  $J$  cluster centroids. Here, we set  $J = 50$  and compare the results to a reference with uniform weights  $\pi_1 = \pi_2 = \pi_3 = 1$ . In the reference setup, the  $\bar{g}_{n,s}$  will be the z-value of a tilted plane at locations  $\mathbf{x}_{n,s}$ .  $\beta_1$  then determines the overall offset of that plane in z-direction and  $\beta_1$  and  $\beta_2$  determine the tilt in the direction of  $\langle D_n \rangle_t$  and  $\langle \tilde{G}_{n,s} \rangle_t$  respectively.

For illustration purposes we, furthermore, consider one time step only using averaged values of  $D_{n,t}$  and  $\tilde{G}_{n,s,t}$  and only one conventional generation technology (OCGT) as well as only one renewable generation technology (wind). Variable operation costs for wind power generation are set to zero, the parameter  $a$  to one. OCGT generation capacities are assumed to be available as needed, i.e. only the deployment of wind power capacity is optimised. Eq. 6.15, hence, further simplifies to

$$\begin{aligned} \min_{\beta_{i,j}, f_{l,t}, \epsilon_{n,s}} \sum_l f_l^2 + \\ \sum_n (C_{wind}^V \bar{g}_{wind,n} + \epsilon_{wind,n}^2 + \\ C_{OCGT}^O (\langle D_n \rangle_t - \bar{g}_{wind,n} \langle \tilde{G}_{wind,n} \rangle_t)) \end{aligned} \quad (6.19)$$

This model is applied to the idealistic German power system used by Kies et al. [2016e] consisting of 402 buses and 1024 branches.  $\langle D_n \rangle_t$  and  $\langle \tilde{G}_{n,s} \rangle_t$  are computed from weather data obtained from reanalysis products spanning three years with hourly resolution (same data as Kies et al. [2016e]). Costs assumptions for capital costs of wind power (1182 €/kW overnight) and variable operation costs for OCGT (58.4 €/MWh) were taken from Schlachtberger et al. [2017]. The expansion model was solved using the *R-Gurobi* interface.

### 6.3.3 Results

Figure 6.4 illustrate the results. For the left column both solutions obtained from the full and reduced model respectively have been interpolated onto a grid in the  $(\langle D \rangle_t, \langle \tilde{G}_{n,s} \rangle_t)$  domain. Each of the three sub-figures in the left column of Figure 6.4 will can be explained as follows:

Figure 6.4 top, left shows the full, meaning non-parameterised, solution. Obviously, the highest covariance can be found between the installed capacity and the average demand. The higher the load, the higher the installed wind power capacity. This positive correlation is less pronounced with the capacity factor. There are, however, a few buses characterized by relatively low demand and relatively high capacity factors exhibiting relatively high wind power capacities while others characterized by average demand and low capacity factors exhibit zero installed capacity. The latter is visible from the sharp decrease in installed capacities at the left border of the plotted surface. The reduced problem with uniform weights (Figure 6.4 center, left) is able to capture the main dependency on the average demand. Since we chose linear surfaces as basis functions, the plotted surface is a (tilted) plane. Any non-linearities (as the sharp decrease at the left border of the surface mentioned above) cannot be captured, which results in a more uniform distribution of wind power capacities. Note that the lowest point of the plane is the left corner in the front. Installed capacity increases towards the right corner (buses with very low demand but high capacity factors mainly located at the North Sea coast, compare Fig. 6.4 center, right). In contrast to the solution using uniform weights, the solution with 50 Gaussian weighting kernels is able to produce non-linear surfaces (Figure 6.4 bottom, left). The variance illustrated in Figure 6.4 top, left is captured well. Note that the grid is only evaluated at locations  $\mathbf{x}_{n,s}$ . Any effects at the boundary of the plotted surface or at any locations in between the  $\mathbf{x}_{n,s}$  are of no consequence. Hence, the spatial distribution of installed wind power capacity compares well with the full solution (Figure 6.4 top, right

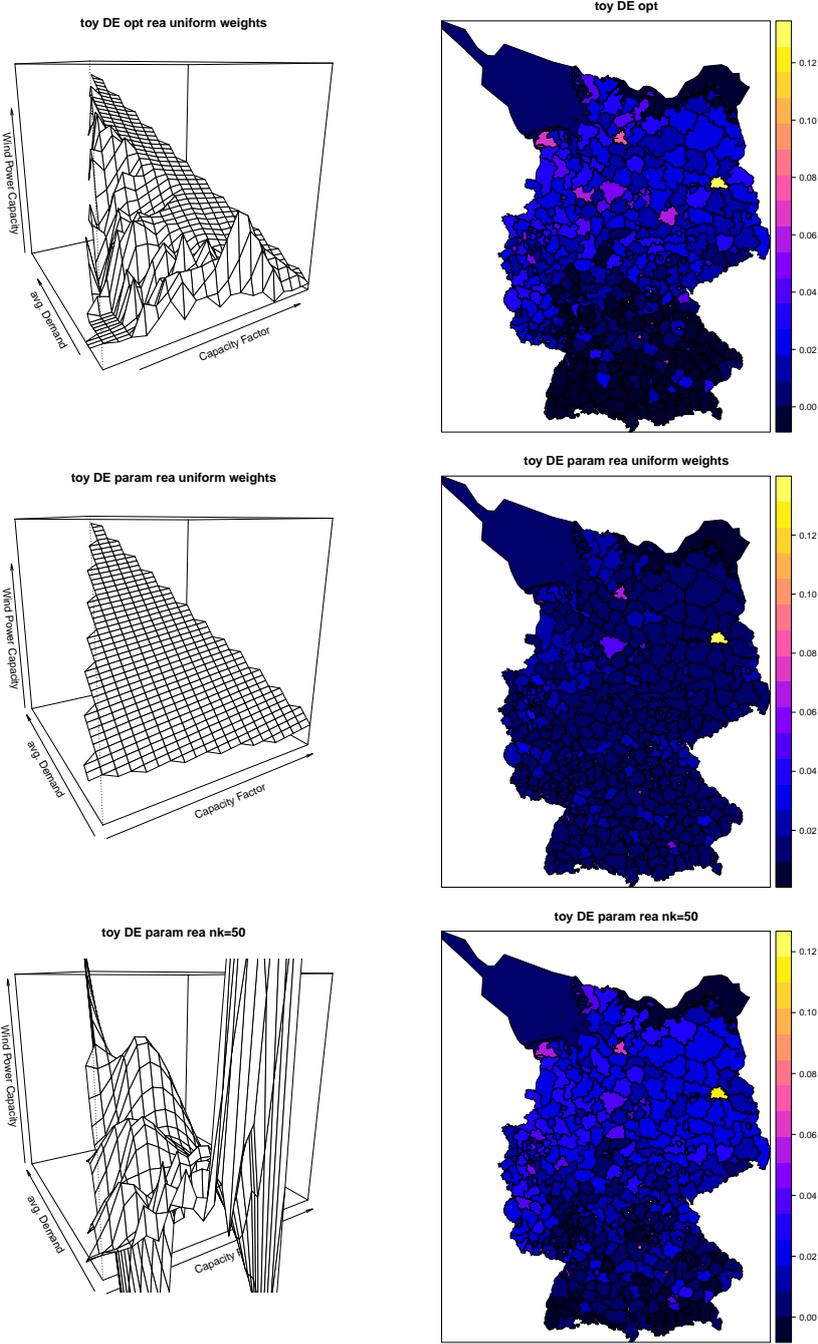


Figure 6.4: Left column: Top: non-parameterised solution, Center: reduced solution with uniform weights, Bottom: reduced solution with  $J = 50$  Gaussian weighting kernels, for details see text; Right column: Same as left column but plotted as color coded polygons, unit of color bar is GW.

and bottom, right).

### 6.3.4 Conclusion

We showed that variable parameterisation is a valuable tool to reduce the complexity of power system expansion models and the computation costs resulting from this complexity. Our model was proven to be able to reproduce the full solution well – even with a rather small number of basis functions. We plan to apply the presented approach on a detailed model of the European power system.

## 6.4 Stochastic Optimisation

Until now, we considered a linear program, which can be used to derive the cost-optimal expansion of power systems. It derives the optimal system design based on the investment cost for generation, storage and transmission assets, the marginal cost for generation and time series of electricity demand and resource availability. As described in Section 2.2, the linear program consists of an upper level for the investment decisions and a lower level for the operational decisions. More precisely, the lower level is an implementation of an idealised network-constraint electricity market. Based on the marginal cost of the generation assets it derives the least-cost generation schedule which ensures that the demand is met at any place in space and time and that the network equations are fulfilled. In the formulation used here (Equations (2.1)-(2.15)), we assume a *perfect foresight* for the availability of renewable resources: We assume, that the market is cleared at *exactly* that time the electricity is needed and that we know the *exact* availability of all generation assets in the system at that time. This perfect foresight is implemented by deriving the availability time series from reanalysis products and by optimising the market clearing for at least a whole year at once. This of course is a keen simplification. In reality, the electricity market is cleared at several stages *prior* to the actual delivery: from day-ahead to intra-day. To be able to do so, providers of renewable electricity and market operators rely on forecasts. As wind speed, solar irradiance (the sources of wind and solar power) and electricity demand are stochastic variables, forecasts of these quantities are subject to uncertainties [Pinson, 2006]. In numerical weather prediction, this uncertainty is commonly accounted for by not only deriving one forecast of a specific quantity, but an ensemble of several ideally independent forecasts (see for instance Molteni et al. [1996]). These forecasts are called *ensemble* forecasts or *stochastic* forecasts. While a single (point) forecast would be called *deterministic* forecast.

A simple network-constrained market model can be formulated as:

$$\min_{g_{n,s,t}, z_a} \sum_{n,s} O_{n,s} g_{n,s,t} \quad (6.20)$$

subject to constraints (2.2), (2.3) and (2.15). Where the resource availability  $\tilde{G}_{n,s,t}$  in (2.3) is derived from forecasts and the nominal power of the generation assets is no longer a decision variable but a parameter to the optimisation problem, i.e.  $\bar{g}_{n,s} = \bar{G}_{n,s}$ . For now, let us assume that the market (6.20) is cleared *day-ahead*, meaning at noon of the day prior to

the day of delivery. Hence, we consider forecasts with a forecast horizon of 12 - 24 hours. Let us, furthermore, assume perfect foresight of the electricity demand, i.e.  $D_{n,t}$  is known.

Now, we additionally consider the stage of delivery. At that stage, adaptations to the day-ahead schedule might be necessary, due to errors in the power forecasts. The actual observed availability might be different from the forecasted. Similarly to Morales et al. [2014], we assume, that this discrepancy can be balanced by some flexible producers. These flexible producers are able to reduce or ramp up their generation with very short lead times. However, the willingness to do so comes at relatively high costs. Flexible providers receive premiums for modifications to their day-ahead schedule. Hence, keeping the balancing actions small is beneficial in the sense of maximizing welfare (compare Section 2.3). Which generator optimally balances the potential forecasting errors, again, is decided on a market. According to Morales et al. [2014] such a *balancing market* can mathematically be defined as:

$$\min_{g_{n,s,t}^{+/-}, z_a} \sum_{n,s} (O_{n,s}^+ g_{n,s,t}^+ + O_{n,s}^- g_{n,s,t}^-) \quad (6.21)$$

where  $g_{n,s,t}^{+/-}$  denote the deviation from the day-ahead schedule  $G_{n,s,t}$  in either direction and  $O_n^{+/-}$  the corresponding cost, respectively. Equation (6.21) is subject to the following constraints:

$$\sum_s (g_{n,s,t}^+ - g_{n,s,t}^-) - \sum_l K_{n,l} f_{l,t} + s_{n,t} = D_{n,t} - \sum_s G_{n,s,t} \quad , \forall n, t \quad (6.22)$$

$$g_{n,s,t}^+ - g_{n,s,t}^- \leq \tilde{O}_{n,s,t} \tilde{G}_{n,s} - G_{n,s,t} \quad (6.23)$$

$$g_{n,s,t}^- \leq G_{n,s,t} \quad (6.24)$$

(6.22) ensures the nodal balances after modifying the original schedule. (6.23) limits the modified generation schedule for each generation asset to its actual observed availability and (6.24) limits the reduction from the original schedule to the original day-ahead schedule. In (6.22), an auxiliary load shedding variable  $s_{n,t}$  is introduced in order to ensure the feasibility of the constraint. Note that in (6.23) the forecasted availability  $\tilde{G}_{n,s,t}$  has been replaced by the observed availability  $\tilde{O}_{n,s,t}$ .

As shown by Morales et al. [2014], this setup of two separate market stages potentially leads to non-optimal market clearings on the day-ahead market because the day-ahead market does not account for the uncertainty in power forecasts and, hence, cannot anticipate potential additional costs originating from necessary balancing actions. Both market stages are considered independent. But in fact, they are not. This can be clearly seen from the fact that the solution of the day-ahead market enters the balancing market as a parameter, i.e.  $G_{n,s,t} = g_{n,s,t}^*$ . Consequently,  $g_{n,s,t}$  can be considered a complicating variable (compare Section 2.2) and both market stages can be combined to form a bi-level optimisation problem, which derives the optimal day-ahead schedule while implicitly taking into account the uncertainty in the power forecasts [Morales et al., 2014]:

$$\min_{g, g_{n,s,t}^{+/-}, z_a} \sum_{n,s} O_{n,s} g_{n,s,t} + E[C^B]_t \quad (6.25)$$

subject to (2.2), (2.3), (2.15) and (6.22) - (6.24). Here,  $E[C^B]$  denotes an estimate of the expected balancing costs. This estimate is derived from a set of scenarios. Each of these scenarios computes the expected balancing cost from one member of an ensemble forecast and the day-ahead decision  $g_{n,s,t}$ . Let us, in the following, call (6.20) a *deterministic* and (6.25) a *stochastic* day-ahead market. In both cases, the total cost for electricity generation can be computed as the sum of the optimal solutions from the day-ahead and the balancing market. It has been shown by Morales et al. [2014], that a stochastic market clearing leads to overall reduced cost. Overall, this setup can be considered as *decision making under uncertainty* (refer to Conejo et al. [2010, e.g] for an introduction).

Here, we use this setup to quantify the economical value of improved probabilistic power forecasts. The hypothesis is as follows: When the forecast skill is increased, the expected balancing costs can be estimated more precisely and the total cost for electricity generation can be reduced further. If this hypothesis holds, each improvement in forecast skill could be assigned an economical value in the sense of a reduction of costs.

For a first proof of concept let us consider the following setup: The power system is defined as an optimised version of the PyPSA-Eur data set with 256 nodes. The nominal power of generation, storage and transmission assets is derived by solving (2.1) - (2.14) with a global CO<sub>2</sub> emission limit of 5 % of the level of 1990 and a global transmission capacity limit of 1.5 times today's value over one year. For this idealised power system, the deterministic market is implemented and solved first. Then, the stochastic market is implemented and solved for two different ensemble forecasts. The first ensemble forecast considered is the raw ensemble of the *ensemble prediction system* (EPS) of the *European Center for Medium-range Weather Forecasts* [Molteni et al., 1996]. This EPS provides 50 ensemble members for each forecast horizon considered here. The second ensemble forecast considered is a bias-corrected version of the raw ensemble. In order to do this bias correction, yearly (for wind speed) and monthly (for solar irradiance) biases of the ensemble forecast compared to a reference data set have been computed and subtracted from the raw ensemble. Here, the ERA5 reanalysis [Copernicus Climate Change Service (C3S), 2017] serves as both, the reference for the bias correction and the observations used for the balancing market. I measure the increase in forecast skill via the *Continuous Ranked Probability Skill Score* (CRPSS) [Gneiting et al., 2007]. This skill score evaluates the forecast skill by the quadratic area between the cumulative probability function of the probabilistic forecast and the observation.

As shown in Figure 6.5 top, the bias correction leads to an increased forecast skill for wind power mainly in regions of pronounced orography (the Alps, the Balkan and Norway). For the whole system, the forecast skill could be improved by approximately 5 %.

In order to investigate whether this increase in forecast skill has an economical effect, we solve the stochastic market model two times for each day in January through March 2017: first with the raw ensemble forecast and second with the bias-corrected forecasts. The difference in these two solutions is depicted in Figure 6.5 center. Obviously, the largest differences occur in Italy, the South of Spain and France, Denmark and in England. Note, that the sum of all changes sum up to zero, since the same demand for electricity must be met in both setups. Hence, we only observe shifts of electricity generation from one generator to another. Nevertheless, these changes in the day-ahead schedule clearly influence the required balancing actions. The difference in nodal balancing cost shows a quite

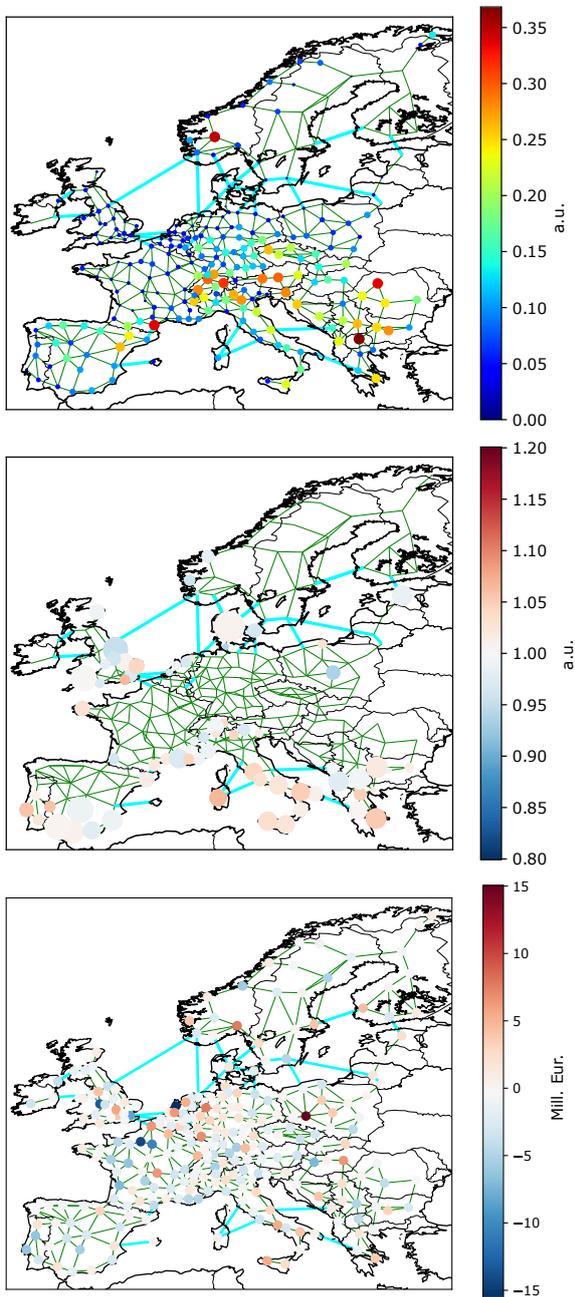


Figure 6.5: Top: Improvement in forecast skill compared to the raw ensemble as measured by the CRPSS [a.u.]; Center: Relative difference in the day-ahead schedule of the renewable generators between the stochastic market clearing with and without the bias-corrected forecasts [a.u.]; Bottom: Difference in nodal balancing costs compared between the stochastic market clearing with and without the bias-corrected forecasts [Mill. Eur].

heterogeneous picture: In France, Northern England, the Iberian Peninsula, Northern Italy, Southern Germany, Austria and the Czech Republic balancing costs decrease. In contrast, they increase in Poland, Northern Germany, the Benelux and Southern Italy. Overall, the day-ahead schedule derived from the stochastic market clearing with improved forecasts leads to a reduction in balancing cost of 20 %.

In summary, we can – for the chosen power system – relate an increase in forecast skill of 5 % to a decrease in balancing cost of 20 %. Whether this is a lucky circumstance or systematic still has to be proven. Therefore, this work is currently extended to cover a longer time frame, to use more differently calibrated forecasts and system designs and with additional intra-day market stages between the day-ahead and the balancing market.

## 6.5 Outlook

As described above, power system models can mathematically be formulated as linear or mixed-integer programs of varying complexity. Addressing all parameter coactions and sensitivities of these models would clearly go beyond the scope of this thesis. Therefore, I focused my analysis on the coaction of meteorological and economical parameters. Coactions between technical and economical parameters and sensitivities on technical parameters, as e.g. the efficiencies of storage units, have not been considered. Furthermore, I kept the model setup comparably simple. On the one hand, this obviously neglects some aspects of real-world power systems and might, hence, appear arbitrary and unrealistic. On the other hand, the simplifications made reduce the complexity of the model, which allows to perform the simulations within reasonable time and at reasonable computational cost. The model's reduced complexity makes the model easier understandable for non-experts, allows to easily reproduce the simulations and, consequently, contributes to the transparency of the results (compare Pfenninger [2017b], DeCarolis et al. [2017]). Additionally, the results obtained from the simulations should, from my point of view, be as generally applicable as possible and not characteristic to only one specific model design. Although modifying the model design and/or complexity clearly may change the simulation results, the general conclusions drawn from them should not significantly be affected.

Nevertheless, it is clear, that this thesis only addresses some aspects of power systems. Indeed, power systems are much more complex as power system expansion models can picture. This gives rise to the question how the ecological and/or societal scale could be incorporated, for example by linking power system optimisation to life cycle assessment and by investigating the resource needs and *energy returned on energy invested* (EROI) [Cleveland et al., 1984] for different system designs and shares of renewables.

In the future, this thesis might be extended by performing a dynamical fixed-point analysis of the availability time series as Mücke et al. [2015] did for power curves and Rinn et al. [2015] for market stages at the stock exchange. These fixed-points might be linked to weather situations or used as a weather classification mechanism itself. Furthermore, the results from research question 3 could be used to set up a framework for the quantification and comparison of model sensitivities in the context of the activities of the open modeling initiative (see <http://openmod-initiative.org/>) and/or to derive representative data sets for power system expansion studies.

# Chapter 7

## Summary and Conclusions

In this thesis, I compiled an overview of parameter coaction in numerical power system modeling and of the sensitivities of a common power system expansion model on the specific parameter choices and model resolutions. The concept of parameter coaction has been introduced via a simple toy-model and some theoretical derivations of the shadow prices in power system optimisation programs and of the numerical complexity of these programs. In particular, I explained how complicating variables and constraints, which couple the investment problem with the operational problems in power system expansion models, prevent the efficient distributed solution of these models, how the concept of Locational Marginal Prices derived from the corresponding dual can be used to allocate costs within a power system and how spatial smoothing, storage units and sector coupling influence the optimal mix of generation resources and the characteristics of the residual load.

By addressing the three research questions (i) How is the relation between weather classes and wind power production? (ii) How do regional differences in cost of capital influence the optimal design of power systems? and (iii) How large is the sensitivity of power system expansion models? I exemplarily investigated some of the aspects of parameter coaction in more detail, focusing on the coaction of meteorological and economical parameters: For research question I I bridged the gap between traditional meteorology, i.e. the definition of weather classes, and the issue of operating a power system with high shares of wind power by showing that the availability of the wind resource can clearly be linked to the prevailing weather situation and that more accurate estimates of the actual wind power production can be derived when the information about the weather situation is explicitly taken into account. Research question II focused on the importance of making adequate assumptions for the model parameters, here the investment cost for generation assets. I explained that whether regional differences in these cost are considered or not significantly influences the optimal design of a highly renewable European power system and the way how costs are allocated within the system via the Locational Marginal Prices of Electricity. The sensitivity of power system expansion models has been investigated in research question III by introducing a novel misallocation metric for quantifying the sensitivity. Again, I stressed the importance of making adequate parameter choices. Furthermore I concluded, that sensitivities should be reported along with the model results in order to increase the transparency in power system modeling and I explained why power system expansion models in hourly resolution are relatively insensitive to the spatial resolution by considering one basic meteorological aspect of wind and solar power: their correlation lengths. In fact, the sensitivity of power system expansion models to a specific parameter choice can be interpreted as the strength of the coaction of that specific parameter with the other parameters: The higher the sensitivity, the

stronger the parameter co-acts with the other parameters towards the optimisation result.

Possible extensions of the presented works have been given in Section 6. Here, I showed how probabilistic power forecasts can be incorporated in electricity market models, how these improved market models can be used to quantify the economical effect of increased forecast skill, how the concept of flow tracing can be used as an alternative way to allocate costs and how the concept of variable parameterisation can be used to decrease the numerical complexity of power system expansion models.

In summary, this work can contribute to a deeper understanding of the specific characteristics of power system expansion models and their sensitivities. Power system models are a valuable tool to address many of the open research questions in the context of the transformation of power and energy systems. I believe that power system modeling can make significant contributions to the societal decision-making process and, hence, support the technical, societal and political progress which is needed to cope with global environmental change. In fact, getting decisions societal and politically accepted might be more challenging as making the decision itself but this is beyond the scope of this thesis. In this context, the reliability of the models and the results obtained from them is crucial. Hence, power system modeling should clearly dedicate itself to transparency and openness in research. Furthermore, models should on the one hand be as close to what people experience in everyday life as possible and on the other hand explainable also to non-experts. Simplifications where needed should be conscientiously motivated and discussed.

## References

# Bibliography

- L. Baringo and A. Conejo. Correlated wind-power production and electric load scenarios for investment decisions. *Applied energy*, 101:475–482, 2013.
- M. S. Bazaraa, H. D. Sherali, and C. M. Shetty. *Nonlinear programming: theory and algorithms*, chapter 6, pages 257 – 314. John Wiley & Sons, 2013.
- S. Becker, R. A. Rodriguez, G. B. Andresen, S. Schramm, and M. Greiner. Transmission grid extensions during the build-up of a fully renewable pan-european electricity supply. *Energy*, 64:404–418, 2014.
- J. Bialek. Tracing the flow of electricity. *IEE Proceedings-Generation, Transmission and Distribution*, 143(4): 313–320, 1996.
- D. Birant and A. Kut. St-dbscan: An algorithm for clustering spatial–temporal data. *Data & knowledge engineering*, 60(1):208–221, 2007.
- H. Bloomfield, D. J. Brayshaw, L. Shaffrey, P. J. Coker, and H. E. Thornton. The changing sensitivity of power systems to meteorological drivers: a case study of great britain. *Environmental Research Letters*, 13(5):054028, 2018.
- H. C. Bloomfield, D. J. Brayshaw, L. C. Shaffrey, P. J. Coker, and H. Thornton. Quantifying the increasing sensitivity of power systems to climate variability. *Environmental Research Letters*, 11(12):124025, 2016.
- M. Bosilovich et al. Nasa’s modern era retrospective-analysis for research and applications: Integrating earth observations. *IEEE Earthzine*, 1(1-4):82367, 2008.
- S. Boyd and L. Vandenberghe. Interior-point methods. In S. Boyd and L. Vandenberghe, editors, *Convex optimization*. Cambridge university press, 2004a.
- S. Boyd and L. Vandenberghe. Duality. In S. Boyd and L. Vandenberghe, editors, *Convex optimization*. Cambridge university press, 2004b.
- E. D. Brill Jr, S.-Y. Chang, and L. D. Hopkins. Modeling to generate alternatives: The hsj approach and an illustration using a problem in land use planning. *Management Science*, 28(3):221–235, 1982.
- T. Brown, J. Hörsch, and D. Schlachtberger. Pypsa: Python for power system analysis. *arXiv preprint arXiv:1707.09913*, 2017.
- T. Brown, D. Schlachtberger, A. Kies, S. Schramm, and M. Greiner. Synergies of sector coupling and transmission reinforcement in a cost-optimised, highly renewable european energy system. *Energy*, 160: 720–739, 2018.
- R. Brückmann. What is the development of WACC for wind power in the 28 EU Member States -and why? [https://www.strommarkttreffen.org/2018-03\\_Brueckmann\\_Development\\_of\\_WACC\\_for\\_wind\\_in\\_EU28.pdf](https://www.strommarkttreffen.org/2018-03_Brueckmann_Development_of_WACC_for_wind_in_EU28.pdf), 2018. Accessed: 2018-11-30.
- C. Budischak, D. Sewell, H. Thomson, L. Mach, D. E. Veron, and W. Kempton. Cost-minimized combinations of wind power, solar power and electrochemical storage, powering the grid up to 99.9 *Journal of Power Sources*, 225:60–74, 2013. doi: 10.1016/j.jpowsour.2012.09.054.
- D. Cano, J.-M. Monget, M. Albuissou, H. Guillard, N. Regas, and L. Wald. A method for the determination of the global solar radiation from meteorological satellite data. *Solar Energy*, 37(1):31–39, 1986.

- K. Chattopadhyay, A. Kies, E. Lorenz, L. von Bremen, and D. Heinemann. The impact of different pv module configurations on storage and additional balancing needs for a fully renewable european power system. *Renewable Energy*, 113:176–189, 2017.
- C. J. Cleveland, R. Costanza, C. A. Hall, and R. Kaufmann. Energy and the us economy: a biophysical perspective. *Science*, 225(4665):890–897, 1984.
- A. J. Conejo, M. Carrión, J. M. Morales, et al. *Decision making under uncertainty in electricity markets*, volume 1. Springer, 2010.
- A. J. Conejo, L. Baringo, S. J. Kazempour, and A. S. Siddiqui. *Investment in Production Capacity*, chapter 5, pages 169 – 228. Springer International Publishing Switzerland, 2016.
- D. Connolly, H. Lund, B. V. Mathiesen, and M. Leahy. A review of computer tools for analysing the integration of renewable energy into various energy systems. *Applied Energy*, 87(4):1059–1082, 2010.
- Copernicus Climate Change Service (C3S). Era5: Fifth generation of ecmwf atmospheric reanalyses of the global climate, 2017. URL <https://cds.climate.copernicus.eu/cdsapp#!/home>. accessed on 2017 via Copernicus Climate Change Service Climate Data Store (CDS).
- Corine Land Cover. online <https://land.copernicus.eu/pan-european/corine-land-cover/clc-2012>, 2017.
- P. Costa, A. Estanqueiro, and P. Miranda. Building a wind atlas for mainland portugal using a weather type classification. In *EWEC'2006-European Wind Energy Conference*, pages 1–9, 2006.
- A. Couto, P. Costa, L. Rodrigues, V. V. Lopes, and A. Estanqueiro. Impact of weather regimes on the wind power ramp forecast in portugal. *IEEE Transactions on Sustainable Energy*, 6(3):934–942, 2014.
- J. DeCarolis, H. Daly, P. Dodds, I. Keppo, F. Li, W. McDowall, S. Pye, N. Strachan, E. Trutnevyte, W. Usher, M. Winning, S. Yeh, and M. Zeyringer. Formalizing best practice for energy system optimization modelling. *Applied energy*, 194:184–198, 2017.
- J. F. DeCarolis, S. Babaee, B. Li, and S. Kanungo. Modelling to generate alternatives with an energy system optimization model. *Environmental Modelling & Software*, 79:300–310, 2016.
- D. Dee, S. Uppala, A. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. Balmaseda, G. Balsamo, P. Bauer, et al. The era-interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656):553–597, 2011.
- G. Doms and M. Baldauf. A description of the nonhydrostatic regional cosmo model. part i: Dynamics and numerics. *Deutscher Wetterdienst, Offenbach*, 2011.
- F. Egli, B. Steffen, and T. S. Schmidt. A dynamic analysis of financing conditions for renewable energy technologies. *Nature Energy*, 2018.
- ENTSO-E. Country-specific hourly load data. online <https://www.entsoe.eu/data/data-portal/>, 2012. URL <https://www.entsoe.eu/data/data-portal/>.
- B. Ernst, B. Oakleaf, M. L. Ahlstrom, M. Lange, C. Moehrlen, B. Lange, U. Focken, and K. Rohrig. Predicting the wind. *IEEE power and energy magazine*, 5(6):78–89, 2007.
- A. Estanqueiro. Impact of wind generation fluctuations in the design and operation of power systems. In *7th international Workshop on Large Scale Integration of Wind Power and on Transmission Networks for Offshore Wind Farms*, 2008.
- European Commission. Clean Energy for all Europeans, Accessed: 2018-09-30. URL <https://ec.europa.eu/energy/en/topics/energy-strategy-and-energy-union/clean-energy-all-europeans>.

- European Commission. A fully-integrated internal energy market, accessed 2016-09-22. URL [http://ec.europa.eu/priorities/energy-union-and-climate/fully-integrated-internal-energy-market\\_en](http://ec.europa.eu/priorities/energy-union-and-climate/fully-integrated-internal-energy-market_en).
- P. Gabrielli, M. Gazzani, E. Martelli, and M. Mazzotti. Optimal design of multi-energy systems with seasonal storage. *Applied Energy*, 219:408–424, 2017. doi: 10.1016/j.apenergy.2017.07.142.
- G. Giebel, R. Brownsword, G. Kariniotakis, M. Denhard, and C. Draxl. The state-of-the-art in short-term prediction of wind power: A literature overview. *ANEMOS. plus*, 2011.
- T. Gneiting, F. Balabdaoui, and A. E. Raftery. Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 69(2):243–268, 2007.
- C. M. Grams, R. Beerli, S. Pfenninger, I. Staffell, and H. Wernli. Balancing europe’s wind-power output through spatial deployment informed by weather regimes. *Nature Climate Change*, 7(8):557–562, 2017.
- J. Grue, B. Gjevik, and J. E. Weber. *Waves and nonlinear processes in hydrodynamics*, volume 34. Springer Science & Business Media, 2012.
- H. Haehne, K. Schmietendorf, S. Tamrakar, J. Peinke, and S. Kettemann. Propagation of wind-power-induced fluctuations in power grids. *Physical Review E*, 99(5):050301, 2019.
- A. Hammer, D. Heinemann, A. Westerhellweg, P. Ineichen, J. Olseth, A. Skartveit, D. Dumortier, M. Fontoynt, L. Wald, H. Beyer, et al. Derivation of daylight and solar irradiance data from satellite observations. In *Proceedings 9th Conference on Satellite Meteorology and Oceanography*, pages 747–750, 1998.
- M. S. Hanni, T. van Giffen, R. Krüger, and H. Mirza. Foreign direct investment in renewable energy: Trends, drivers and determinants. *Transnational corporations*, 20(2):29–65, 2011.
- D. Heide, L. Von Bremen, M. Greiner, C. Hoffmann, M. Speckmann, and S. Bofinger. Seasonal optimal mix of wind and solar power in a future, highly renewable europe. *Renewable Energy*, 35(11):2483–2489, 2010.
- D. Heide, M. Greiner, L. Von Bremen, and C. Hoffmann. Reduced storage and balancing needs in a fully renewable european power system with excess wind and solar power generation. *Renewable Energy*, 36(9):2515–2523, 2011.
- L. Hirth and S. Müller. System-friendly wind power: How advanced wind turbine design can increase the economic value of electricity generated through wind power. *Energy Economics*, 56:51–63, 2016.
- L. Hirth and J. C. Steckel. The role of capital costs in decarbonizing the electricity sector. *Environ. Res. Lett.*, 11(11):114010, 2016.
- H. Holttinen, P. Meibom, A. Orths, B. Lange, M. O’Malley, J. O. Tande, A. Estanqueiro, E. Gomez, L. Söder, G. Strbac, et al. Impacts of large amounts of wind power on design and operation of power systems, results of iea collaboration. *Wind Energy*, 14(2):179–192, 2011.
- T. Huld, R. Gottschalg, H. G. Beyer, and M. Topič. Mapping the performance of pv modules, effects of module type and data averaging. *Solar Energy*, 84(2):324–338, 2010.
- R. Huth, C. Beck, A. Philipp, M. Demuzere, Z. Ustrnul, M. Cahynová, J. Kyselý, and O. E. Tveito. Classifications of atmospheric circulation patterns: recent advances and applications. *Annals of the New York Academy of Sciences*, 1146(1):105–152, 2008.
- J. Hörsch and T. Brown. The role of spatial scale in joint optimisations of generation and transmission for european highly renewable scenarios. In *14th International Conference on the European Energy Market*, 2017.

- J. Hörsch, F. Hofmann, D. Schlachtberger, and T. Brown. Pypsa-eur: An open optimisation model of the european transmission system. *arXiv preprint arXiv:1806.01613*, 2018a.
- J. Hörsch, H. Ronellenfitsch, D. Witthaut, and T. Brown. Linear optimal power flow using cycle flows. *Electric Power Systems Research*, 158, 2018b. doi: 10.1016/j.epsr.2017.12.034.
- T. Ishihara, A. Yamaguchi, T. Ogawa, K. Sakai, and Y. Fujino. An upscaling approach for the regional wind power forecasting. *Ratio*, 243:100–0, 2007.
- D. Jacob, J. Petersen, B. Eggert, A. Alias, O. B. Christensen, L. M. Bouwer, A. Braun, A. Colette, M. Déqué, G. Georgievski, et al. Euro-cordex: new high-resolution climate change projections for european impact research. *Regional environmental change*, 14(2):563–578, 2014.
- A. Jenkinson and F. Collison. An initial climatology of gales over the north sea. *Synoptic climatology branch memorandum*, 62:18, 1977.
- T. V. Jensen and M. Greiner. Emergence of a phase transition for the required amount of storage in highly renewable electricity systems. *The European Physical Journal Special Topics*, 223(12):2475–2481, 2014.
- P. A. Jiménez, J. F. González-Rouco, J. P. Montávez, E. García-Bustamante, and J. Navarro. Climatology of wind patterns in the northeast of the iberian peninsula. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 29(4):501–525, 2009.
- J. Jurasz and B. Ciapała. Integrating photovoltaics into energy systems by using a run-off-river power plant with pondage to smooth energy exchange with the power grid. *Applied energy*, 198:21–35, 2017.
- J. Jurasz, F. Canales, A. Kies, M. Guezgouz, and A. Beluco. A review on the complementarity of renewable energy sources: Concept, metrics, application and future research directions. *Solar Energy*, 195:703–724, 2020.
- A. Kies. *Modelling Highly Renewable Power Systems – The Importance of Backup, Demand Side Management and Curtailment*. PhD thesis, Carl von Ossietzky Universität Oldenburg, 2017.
- A. Kies, K. Chattopadhyay, E. Lorenz, L. v. Bremen, and D. Heinemann. Restore 2050 work package report d12: Simulation of renewable feed-in for power system studies. Technical report, University of Oldenburg, Insitute of Physics, ForWind, 2016a.
- A. Kies, K. Chattopadhyay, L. von Bremen, E. Lorenz, and D. Heinemann. Simulation of renewable feed-in for power system studies. Work Package Report D12, Restore 2050, 2016b.
- A. Kies, K. Chattopadhyay, L. von Bremen, E. Lorenz, and D. Heinemann. Restore 2050: Simulation of renewable feed-in for power system studies. Technical report, University of Oldenburg, 2016c.
- A. Kies, B. U. Schyska, and L. von Bremen. The optimal share of wave power in a highly renewable power system on the iberian peninsula. *Energy reports*, 2:221–228, 2016d.
- A. Kies, B. U. Schyska, and L. von Bremen. Curtailment in a highly renewable power system and its effect on capacity factors. *Energies*, 9(7):510, 2016e.
- A. Kies, B. U. Schyska, and L. von Bremen. The demand side management potential to balance a highly renewable european power system. *Energies*, 9(11):955, 2016f.
- A. Kies, B. U. Schyska, and L. von Bremen. The effect of hydro power on the optimal distribution of wind and solar generation facilities in a simplified highly renewable european power system. *Energy Procedia*, 97:149–155, 2016g.
- A. Kies, L. von Bremen, and D. Heinemann. Hydro energy inflow for power system studies, June 2017. URL <https://doi.org/10.5281/zenodo.803880>.

- L. Kitzing, C. Mitchell, and P. E. Morthorst. Renewable energy policies in europe: Converging or diverging? *Energy Policy*, 51:192–201, 2012.
- C. Klessmann, M. Rathmann, D. de Jager, A. Gazzo, G. Resch, S. Busch, and M. Ragwitz. Policy options for reducing the costs of reaching the european renewables target. *Renewable Energy*, 57:390–403, 2013.
- T. Klucher. Evaluation of models to predict insolation on tilted surfaces. *Solar Energy*, 23(2):111–114, 1979.
- L. Kotzur, P. Markewitz, M. Robinius, and D. Stolten. Impact of different time series aggregation methods on optimal energy system design. *Renewable Energy*, 117:474–487, 2018.
- M. Kovacevic, J. Assa, A. Bonini, C. Calderon, Y.-C. Hsu, C. Lengfelder, T. Mukhopadhyay, S. Nayyar, C. Rivera, and H. Tapia. Human development indices and indicators – 2018 statistical update. Technical report, United Nations Development Programme, 2018.
- M. Krutova, A. Kies, B. U. Schyska, and L. v. Bremen. The smoothing effect for renewable resources in an afro-urasian power grid. *Advances in science and research*, 14:253–260, 2017.
- M. Lange and D. Heinemann. Relating the uncertainty of short-term wind speed predictions to meteorological situations with methods from synoptic climatology. In *Proceedings of the European Wind Energy Conference & Exhibition*, 2003.
- Lazard. Lazard's levelized cost of energy analysis, version 11.0, 2019. URL <https://www.lazard.com/perspective/levelized-cost-of-energy-2017/>.
- P. Li, X. Guan, and J. Wu. Aggregated wind power generation probabilistic forecasting based on particle filter. *Energy conversion and management*, 96:579–587, 2015.
- Y. Liu, X. Gao, J. Yan, S. Han, and D. G. Infield. Clustering methods of wind turbines and its application in short-term wind power forecasts. *Journal of Renewable and Sustainable Energy*, 6(5):053119, 2014.
- Y. Liu, R. Sioshansi, and A. J. Conejo. Hierarchical clustering to find representative operating periods for capacity-expansion modeling. *IEEE Transactions on Power Systems*, 33(3):3029–3039, 2017.
- M. G. Lobo and I. Sanchez. Regional wind power forecasting based on smoothing techniques, with application to the spanish peninsular system. *IEEE Transactions on Power Systems*, 27(4):1990–1997, 2012.
- H. Lund. Renewable energy strategies for sustainable development. *Energy*, 32(6):912–919, 2007.
- H. Lund and W. Kempton. Integration of renewable energy into the transport and electricity sectors through v2g. *Energy policy*, 36(9):3578–3587, 2008.
- J. MacQueen et al. Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, volume 1 (14), pages 281–297. Oakland, CA, USA, 1967.
- C. Q. Marrone, K. Knorr, B. Lange, and H. Beyer. Characterization and modeling of the variability of the power output of aggregated wind farms. In *Proceedings of the 7th International Workshop on Large Scale Integration of Wind Power and on Transmission Networks for Offshore Wind Farms Characterization*, 2008.
- C. M. S. Martin, J. K. Lundquist, and M. A. Handschy. Variability of interconnected wind plants: correlation length and its dependence on variability time scale. *Environmental Research Letters*, 10(4):044004, 2015.
- G. Mavromatidis, K. Orehounig, and J. Carmeliet. Uncertainty and global sensitivity analysis for the optimal design of distributed energy systems. *Applied Energy*, 214:219–238, 2018.
- J. Miettinen, H. Holttinen, and G. Giebel. Nordic wind power forecast errors: benefits of aggregation and impact to balancing market volumes. In *Proceedings of the 13th International Workshop on Large-Scale Integration of Wind Power into Power Systems as well as on Transmission Networks for Offshore Wind Plants*, 2014.

- P. Milan, M. Wächter, and J. Peinke. Turbulent character of wind energy. *Physical review letters*, 110(13): 138701, 2013.
- B. Ming, P. Liu, S. Guo, X. Zhang, M. Feng, and X. Wang. Optimizing utility-scale photovoltaic power generation for integration into a hydropower reservoir by incorporating long-and short-term operational decisions. *Applied Energy*, 204:432–445, 2017.
- F. Modigliani and M. H. Miller. The cost of capital, corporation finance and the theory of investment. *The American economic review*, 48(3):261–297, 1958.
- F. Modigliani and M. H. Miller. Corporate income taxes and the cost of capital: a correction. *The American economic review*, 53(3):433–443, 1963.
- F. Molteni, R. Buizza, T. N. Palmer, and T. Petroliaigis. The ecmwf ensemble prediction system: Methodology and validation. *Quarterly journal of the royal meteorological society*, 122(529):73–119, 1996.
- J. M. Morales, A. J. Conejo, H. Madsen, P. Pinson, and M. Zugno. *Integrating renewables in electricity markets: operational problems*, volume 205. Springer Science & Business Media, 2013.
- J. M. Morales, M. Zugno, S. Pineda, and P. Pinson. Electricity market clearing with improved scheduling of stochastic production. *European Journal of Operational Research*, 235(3):765–774, 2014.
- S. Moret, V. C. Gironès, M. Bierlaire, and F. Maréchal. Characterization of input uncertainties in strategic energy planning models. *Applied Energy*, 202:597 – 617, 2017. doi: <https://doi.org/10.1016/j.apenergy.2017.05.106>.
- T. A. Mücke, M. Wächter, P. Milan, and J. Peinke. Langevin power curve analysis for numerical wind energy converter models with new insights on high frequency power performance. *Wind Energy*, 18(11): 1953–1971, 2015.
- L. Nacken, F. Krebs, T. Fischer, and C. Hoffmann. Integrated renewable energy systems for germany—a model-based exploration of the decision space. In *2019 16th International Conference on the European Energy Market (EEM)*, pages 1–8. IEEE, 2019.
- Natura 2000. Natura 2000 data, the european network of protected sites. online <https://www.eea.europa.eu/data-and-maps/data/natura-7>, 2016.
- F. Neumann and T. Brown. The near-optimal feasible space of a renewable power system model. *arXiv preprint arXiv:1910.01891*, 2019.
- P. Noothout, D. d. Jager, L. Tesnière, S. van Rooijen, N. Karypidis, R. Brückmann, et al. Diacore. the impact of risks in renewable energy investments and the role of smart policies. final report. *Ecofys. Utrecht*, 2016.
- S. Pacala and R. Socolow. Stabilization wedges: solving the climate problem for the next 50 years with current technologies. *Science*, 305(5686):968–972, 2004.
- I. Pawel. The cost of storage—how to calculate the levelized cost of stored energy (lcoe) and applications to renewable energy generation. *Energy Procedia*, 46:68–77, 2014.
- J. Peña, M. Aran, J. Cunillera, and J. Amaro. Atmospheric circulation patterns associated with strong wind events in catalonia. *Natural Hazards and Earth System Sciences*, 11(1):145, 2011.
- I. J. Perez-Arriaga and C. Battle. Impacts of intermittent renewables on electricity generation system operation. *Economics of Energy & Environmental Policy*, 1(2):3–18, 2012.
- U. Pfeifroth, S. Kothe, R. Müller, J. Trentmann, R. Hollmann, P. Fuchs, and M. Werscheck. Surface radiation data set - heliosat (sarah) - edition 2. Technical report, Satellite Application Facility on Climate Monitoring (CM SAF), 2017.

- S. Pfenninger. Dealing with multiple decades of hourly wind and pv time series in energy models: A comparison of methods to reduce time resolution and the planning implications of inter-annual variability. *Applied Energy*, 197:1–13, 2017a.
- S. Pfenninger. Energy scientists must show their workings. *Nature*, 542(7642):393–393, 2017b.
- S. Pfenninger, P. Gauché, J. Lilliestam, K. Damerau, F. Wagner, and A. Patt. Potential for concentrating solar power to provide baseload and dispatchable power. *Nature Climate Change*, 4(8):689–692, 2014.
- A. Philipp, J. Bartholy, C. Beck, M. Ericum, P. Esteban, X. Fettweis, R. Huth, P. James, S. Jourdain, F. Kreienkamp, et al. Cost733cat—a database of weather and circulation type classifications. *Physics and Chemistry of the Earth, Parts A/B/C*, 35(9-12):360–373, 2010.
- A. Pina, C. Silva, and P. Ferrão. Modeling hourly electricity dynamics for policy making in long-term scenarios. *Energy Policy*, 39(9):4692–4702, 2011.
- S. Pineda and J. M. Morales. Chronological time-period clustering for optimal capacity expansion planning with storage. *IEEE Transactions on Power Systems*, 2018.
- S. Pineda, J. M. Morales, and T. K. Boomsma. Impact of forecast errors on expansion planning of power systems with a renewables target. *European Journal of Operational Research*, 248(3):1113–1122, 2016.
- P. Pinson. *Estimation of the uncertainty in wind power forecasting*. PhD thesis, École Nationale Supérieure des Mines de Paris, 2006.
- P. Pinson, N. Siebert, and G. Kariniotakis. Forecasting of regional wind generation by a dynamic fuzzy-neural networks based upscaling approach. In *proceedings of EWEC 2003*, Madrid, Spain, 2003.
- P. Pinson, C. Chevallier, and G. N. Kariniotakis. Trading wind generation from short-term probabilistic forecasts of wind power. *IEEE Transactions on Power Systems*, 22(3):1148–1156, 2007.
- J. Price and I. Keppo. Modelling to generate alternatives: A technique to explore uncertainty in energy-environment-economy models. *Applied Energy*, 195:356–369, 2017.
- M. M. Rienecker, M. J. Suarez, R. Gelaro, R. Todling, J. Bacmeister, E. Liu, M. G. Bosilovich, S. D. Schubert, L. Takacs, G.-K. Kim, et al. Merra: Nasa’s modern-era retrospective analysis for research and applications. *Journal of Climate*, 24(14):3624–3648, 2011.
- P. Rinn, Y. Stepanov, J. Peinke, T. Guhr, and R. Schäfer. Dynamics of quasi-stationary systems: Finance as an example. *EPL (Europhysics Letters)*, 110(6):68003, 2015.
- L. Rockach. A survey of clustering algorithms. In O. Maimon and L. Rockach, editors, *Data mining and knowledge discovery handbook*, pages 269–298. Springer, New York, 2 edition, 2010.
- R. A. Rodriguez, S. Becker, G. B. Andresen, D. Heide, and M. Greiner. Transmission needs across a fully renewable european power system. *Renewable Energy*, 63:467–476, 2014.
- M. N. S. Kisilevich, F. Mansmann and S. Rinzivillo. Spatio-temporal clustering. In O. Maimon and L. Rockach, editors, *Data mining and knowledge discovery handbook*, pages 855–874. Springer, New York, 2 edition, 2010.
- F. Santos-Alamillos, D. Pozo-Vázquez, J. Ruiz-Arias, L. Von Bremen, and J. Tovar-Pescador. Combining wind farms with concentrating solar plants to provide stable renewable power. *Renewable Energy*, 76: 539–550, 2015.
- W.-P. Schill. Electricity storage and the renewable energy transition. *Joule*, 2020.
- D. P. Schlachtberger, T. Brown, S. Schramm, and M. Greiner. The benefits of cooperation in a highly renewable european electricity network. *Energy*, 134:469–481, 2017.

- D. P. Schlachtberger, T. Brown, M. Schäfer, S. Schramm, and M. Greiner. Cost optimal scenarios of a future highly renewable european electricity system: Exploring the influence of weather data, cost parameters and policy constraints. *Energy*, 163:100–114, 2018.
- M. Schlott, A. Kies, T. Brown, S. Schramm, and M. Greiner. The impact of climate change on a cost-optimal highly renewable european electricity network. *Applied Energy*, 230:1645–1659, 2018.
- A. Schröder, F. Kunz, J. Meiss, R. Mendelevitch, and C. von Hirschhausen. Current and prospective costs of electricity generation until 2050. Data Documentation 68, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin, 2013.
- B. U. Schyska and A. Kies. How regional differences in cost of capital influence the optimal design of power systems. *Applied Energy*, 262:114523, 2020.
- B. Shirzadeh, Q. Perrier, and P. Quirion. How sensitive are optimal fully renewable power systems to technology cost uncertainty? Policy Papers 2019.04, International Research Center on Environment & Development (CIRED), 2019.
- S. Späth, L. von Bremen, C. Junk, and D. Heinemann. Time-consistent calibration of short-term regional wind power ensemble forecasts. *Meteorol. Z.*, 24:381–392, 2015.
- F. Steinke, P. Wolfrum, and C. Hoffmann. Grid vs. storage in a 100% renewable europe. *Renewable Energy*, 50:826–832, 2013.
- J. R. Stroud, P. Müller, and B. Sansó. Dynamic models for spatio-temporal data. *Journal of the Royal Statistical Society*, 63(4), 2002.
- I. Temperton. Reducing the cost of financing renewables in europe. *Agora Energiewende*, 2016.
- B. Tranberg, A. B. Thomsen, R. A. Rodriguez, G. B. Andresen, M. Schäfer, and M. Greiner. Power flow tracing in a simplified highly renewable european electricity network. *New Journal of Physics*, 17(10): 105002, 2015.
- B. Tranberg, M. Schäfer, T. Brown, J. Hörsch, and M. Greiner. Flow-based analysis of storage usage in a low-carbon european electricity scenario. In *2018 15th International Conference on the European Energy Market (EEM)*, pages 1–5. IEEE, 2018.
- A. Trancoso. *Operational Modelling as a Tool in Wind Power Forecast and Meteorological Warnings*. PhD thesis, Instituto Superior Técnico, Technical University of Lisbon, 2012.
- R. M. Trigo and C. C. DaCamara. Circulation weather types and their influence on the precipitation regime in portugal. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 20(13): 1559–1581, 2000.
- E. Trutnevte. Does cost optimization approximate the real-world energy transition? *Energy*, 106:182–193, 2016.
- F. Ueckerdt, R. Brecha, G. Luderer, P. Sullivan, E. Schmid, N. Bauer, D. Böttger, and R. Pietzcker. Representing power sector variability and the integration of variable renewables in long-term energy-economy models using residual load duration curves. *Energy*, 90:1799–1814, 2015.
- C. L. Vincent. *Mesoscale wind fluctuations over Danish waters*. PhD thesis, Risø National Laboratory for Sustainable Energy, Technical University of Denmark, 2010.
- A. Voldoire, E. Sanchez-Gomez, D. S. y Méliá, B. Decharme, C. Cassou, S. Sénési, S. Valcke, I. Beau, A. Alias, M. Chevallier, et al. The cnrm-cm5. 1 global climate model: description and basic evaluation. *Climate Dynamics*, 40(9-10):2091–2121, 2013.

## Bibliography

---

- R. Wall, S. Grafakos, A. Gianoli, and S. Stavropoulos. Which policy instruments attract foreign direct investments in renewable energy? *Climate policy*, 19(1):59–72, 2019.
- B. Wealer, S. Bauer, L. Göke, C. von Hirschhausen, and C. Kemfert. High-priced and dangerous: nuclear power is not an option for the climate-friendly energy mix. Technical Report DIW Weekly Report 30/2019, DIW, 2019.
- J. Weber, H. U. Heinrichs, B. Gillessen, D. Schumann, J. Hörsch, T. Brown, and D. Witthaut. Counter-intuitive behaviour of energy system models under co2 caps and prices. *Energy*, 170:22–30, 2019.
- S. Weitemeyer, D. Kleinhans, L. Wienholt, T. Vogt, and C. Agert. A european perspective: potential of grid and storage for balancing renewable power systems. *Energy Technology*, 4(1):114–122, 2016.
- X. Yue, S. Pye, J. DeCarolis, F. G. Li, F. Rogan, and B. O. Gallachóir. A review of approaches to uncertainty assessment in energy system optimization models. *Energy Strategy Reviews*, 21:204 – 217, 2018. doi: <https://doi.org/10.1016/j.esr.2018.06.003>.
- W. Zappa, M. Junginger, and M. van den Broek. Is a 100% renewable european power system feasible by 2050? *Applied Energy*, 233:1027–1050, 2019.

# Appendix

## Load Flow Calculation

Starting point of the theoretical derivation of load flows are the line equation for the active power  $\Phi_l$  and reactive power  $Q_l$  over a transmission line  $l$  connecting a bus  $n$  with another bus  $m$  in an alternating current (AC) network:

$$\Phi_l = \Phi_l(\delta_{n,m}, V_{n,m}) = |V_n|^2 g_l + |V_n||V_m| (g_l \cos(\delta_n - \delta_m) + b_l \sin(\delta_n - \delta_m)) \quad (\text{A1})$$

$$Q_l = Q_l(\delta_{n,m}, V_{n,m}) = |V_n|^2 b_l + |V_n||V_m| (g_l \sin(\delta_n - \delta_m) - b_l \cos(\delta_n - \delta_m)) \quad (\text{A2})$$

with  $V_{n,m}$  being the voltage magnitude at buses  $n$  and  $m$  respectively,  $\delta_{n,m}$  being the voltage angles and  $b_l, g_l$  being the susceptance and conductance of the transmission line, respectively. In the following, this section follows the derivations of Kies [2017] on the nodal injection pattern. It introduces the so-called DC-approximation for load flows in AC networks, which can also be found in text books.

Aim of the DC approximation is the linearization of the above-mentioned non-linear equations, in order to be able to include the load flows, i.e. the active power, in a linear optimization problem. It bases on the following four assumptions:

1. Reactive power in an AC network is small and can consequently be neglected.
2. Voltage angle differences are also small, hence  $\sin(\delta_n - \delta_m) \approx \delta_n - \delta_m, \forall n, m \in \mathcal{N}$
3. The conductance is much smaller than the susceptance, such that the corresponding term can be neglected.
4. Voltage magnitudes are approximately one.

When these assumptions hold, equation A1 can finally be simplified to

$$\Phi_l = b_l(\delta_n - \delta_m) = f_l \quad (\text{A3})$$

This equations expresses the load flow along a transmission line  $l$  as a function of the voltage angles at the terminating buses  $n$  and  $m$ . Due to its similarity to the load flow in DC networks, where voltage angles are replaced by the voltage magnitudes, this equation is called the DC approximation.

Physicality of the flows  $f_l$  is ensured by invoking Kirchoff's current (KCL) and voltage law (KVL), which state that (i) the power reaching each bus must equal the power withdrawn from the bus – either via attached lines or by consumption – and (ii) all partial voltages, i.e. differences in the electrical potential, along a closed cycle sum up to zero. For the following derivations we need to define the following three matrices:

1. the incidence matrix  $\mathbf{K}$  with

$$k_{nl} = \begin{cases} 1 & \text{if line } l \text{ begins at node } n \\ -1 & \text{if line } l \text{ ends at node } n \\ 0 & \text{otherwise.} \end{cases} \quad (\text{A4})$$

2. the diagonal susceptance matrix  $\mathbf{X}$  with  $x_{ll} = b_l$  and

3. the network Laplacian  $\Lambda = \mathbf{KXK}^T$

With the incidence matrix, the flows can be expressed as

$$f_l = b_l \sum_n k_{nl} \delta_n, \forall l = 1 \dots L \quad (\text{A5})$$

and KCL reads

$$p_n = \sum_l k_{nl} f_l \quad (\text{A6})$$

$$= \sum_m \lambda_{nm} \delta_m, \forall n = 1 \dots N \quad (\text{A7})$$

where  $p_n$  is the net active power at bus  $n$ , i.e. the difference between consumption and generation, and  $\lambda_{nm}$  is the element of the network Laplacian  $\Lambda$ .

From these considerations, several different methods to determine the flow of electricity in the framework of a power system model, i.e. to define constraint 2.15, can be derived, for instance the well known *Power Transfer Distribution Factors* (PTDF). For this thesis, I mainly used the following formulation (compare Hörsch et al. [2018b]): In order to determine the active power flow, the voltage angles are set as auxiliary variables to the linear program (2.1)-(2.15), i.e.  $z_a = \delta_n$ , and the following corresponding constraints are invoked:

$$\left| \sum_n (XK^T)_{ln} \delta_n \right| \leq f_l, \quad \forall l = 1 \dots L \quad (\text{A8})$$

$$p_n = \sum_m \lambda_{nm} \delta_m, \quad \forall n = 1 \dots N \quad (\text{A9})$$

$$\delta_0 = 0 \quad (\text{A10})$$

Here, (A8) prohibits line overloading, (A9) ensures the fulfillment of KCL and (A10) fixes the voltage angle at a reference bus (the *slack*) because (A9) is under-determined. Compared to the PTDF approach, this formulation increases the number of decision variables and equality constraints. However, PTDF lead to a significant increase in the solution time for a number of different test cases caused by dense matrices in the PTDF formulation and corresponding large file sizes [Hörsch et al., 2018b].

In a more simplified setup, the transmission lines are replaced by simplified HVDC links. In this case, the load flow along these links are introduced as additional decision variables and the only constraint ensures that these flows do not exceed the net transfer capacity of the

respective link (Eq. 2.14). This setup is used in chapter 4.

Note, that unless stated otherwise, a global transmission capacity limit constraint (equation 2.7) is invoked in all simulations limiting overall transmission capacity to a multiple of today's capacity. In this case, the price for transmission expansion is given by the dual variable corresponding to constraint 2.7 and transmission costs  $C_I$  are consequently set to zero.

## Research Question 2: Supplemental Figures

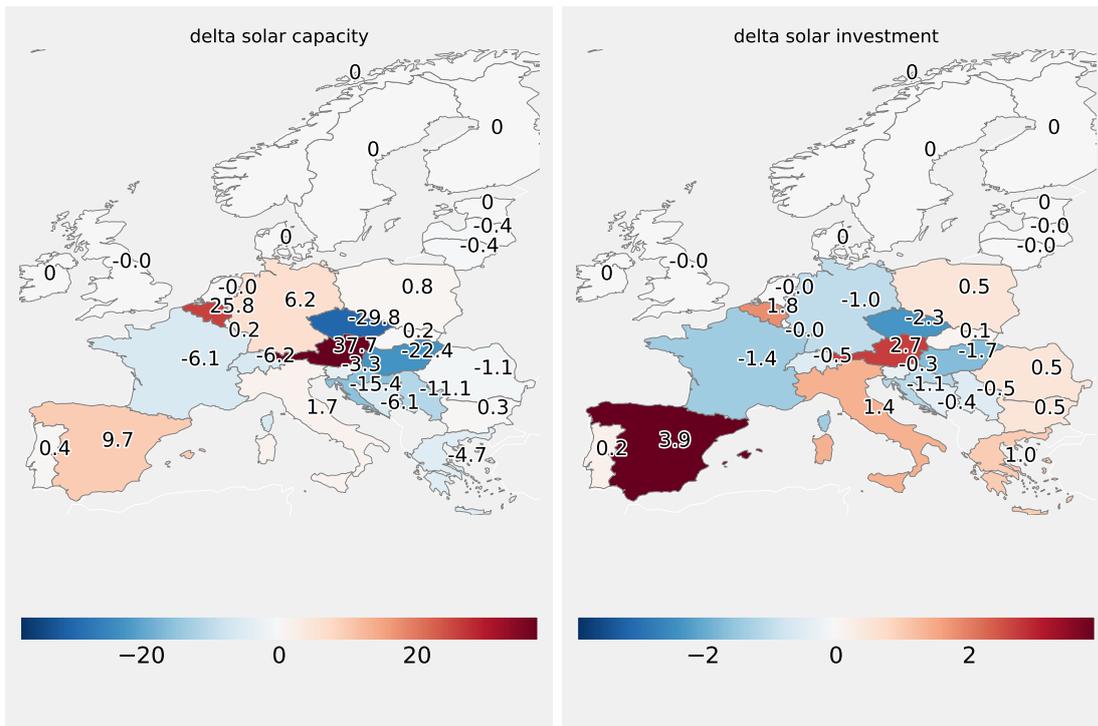


Figure A1: Regional differences in installed solar PV capacities [GW] (left) and investments [Bill. Euro] (right) between the homogeneous and the today scenario.





