Technical University of Munich
Department for Electrical and Computer Engineering
Chair for Renewable and Sustainable Energy Systems

Handling uncertainty in power system optimization

Magdalena Julia Ruth Regina Stüber

Vollständiger Abdruck der von der Fakultät für Elektrotechnik und Informationstechnik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktor-Ingenieurs (Dr.-Ing.)

genehmigten Dissertation.

Vorsitzender: Prof. Dr. Ing. Ulf Schlichtmann

Prüfer der Dissertation:

1. Prof. Dr. rer. nat. Thomas Hamacher
2. Prof. Dr. rer. soc. David Wozabal

Die Dissertation wurde am 21.11.2019 bei der Technischen Universität München eingereicht und durch die Fakultät für Elektrotechnik und Informationstechnik am 28.10.2020 angenommen.

Abstract

Power system modelers face uncertainty every day in various facets: Next to lacking data or unsatisfying data quality, methods on how to mathematically model uncertain behavior of system participants like renewable energy sources are typical cases. If results incorporating uncertainty are generated some way or another, the question on how valuable they are and what decisions to draw from these results arises.

These three problems are addressed in this thesis step-by-step. For the data part, the central questions of a modeler regarding data quality and licensing are answered. Two data sets to represent the German and Bavarian power system are presented and the underlying assumptions, aggregations and transformations can be easily reproduced with the published scripts.

As for the mathematical modeling, the focus is on one method in particular: Stochastic Dual Dynamic Programming (SDDP). The derivation of the method includes the deterministic version Dual Dynamic Programming (DDP) and basics of decomposition methods such as Benders decomposition. The close relation between Benders decomposition and SDDP is used to employ a unified cut generation procedure developed for Benders decomposition and adapt it for SDDP.

At last, the presented model for Germany is analyzed by a sensitivity analysis and deployed to the SDDP algorithm to perform a short term and long term case study. The results are compared to a classic perfect foresight approach.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

> Uncertainty is not an accident of the scientific method, but its substance.
>
> Saltelli et al. [1]

Facing uncertainty in the context of power systems is not a new phenomenon. Uncertainty itself, which is a fundamental part of science and its search for truth, has been discussed over centuries [2, 3]. Ayyub and Klir [3, p. 128] state that uncertainty can be understood as lack of information about a phenomenon or condition. Hence, they argue that reducing uncertainty can be used as unit to measure information. The exact amount of information generated by an experiment or simulation of a phenomenon is given by the difference of the *a priori* to the *a posteriori* uncertainty [3, p. 128] [4, p. 5]. This way to express and handle uncertainty was introduced by Klir [5] as Generalized Information Theory (GIT). Ayyub and Klir [3] build their methods on uncertainty upon the ideas of GIT.

The scientific procedure often requires a certain model to describe a complex system tractably. Klir and Wierman [4, p. 4] state the three main properties of each model: complexity, credibility and uncertainty. There exists a close relationship between all three properties, which is illustrated by an example: To reduce complexity of the model, one has to accept an increase of uncertainty while the credibility increases as well. A typical application would be a linearization of a nonlinear function, which leads to a higher uncertainty, but dealing with a linear function instead of a nonlinear one is both less complex and easier to understand, hence, more credible.

By thinking about the requirements science has on dealing with uncertainty and research in general, one comes close to conducts [6], ethical questions including transparency [7] and a general criticism in the structure and working of the sciences itself [8, 122ff].

In the spirit of transparency and reproducibility, it is tried throughout this thesis not only to provide the underlying data and assumptions but also the model code and the scripts for aggregation and analysis of the data. Before going into more detail, uncertainty in the context of power system analysis is highlighted and examples in the literature are provided.

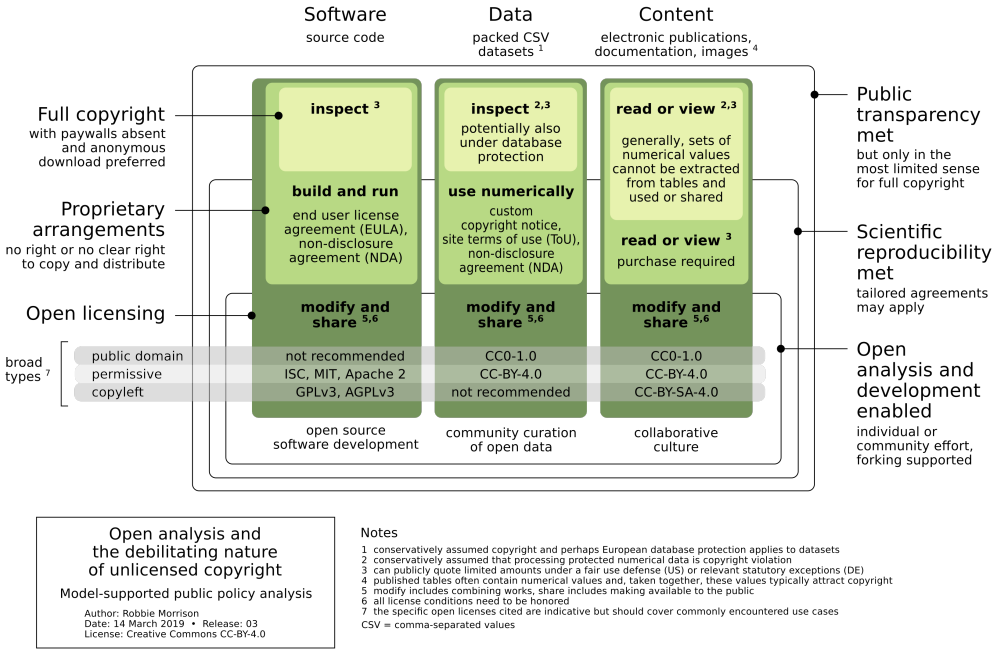


Figure 1.1: Overview of licensing in the context of software, data and content (Robbie Morrison, CC BY 4.0)

1.1 Role of uncertainty in power system optimization

Dealing with uncertainty in power system optimization consists of three parts: First, the uncertainty regarding raw data and assumptions needed for an analysis. Second, the underlying abstraction of the power system in a mathematical optimization problem and third, the way of interpreting results obtained by such an optimization model.

The current situation in each of these parts and their main problems are presented in the following.

1.1.1 Modeling, data and transparency

At the beginning of each energy system analysis, the question on how to model and where to get the data arises. Both questions come hand in hand with transparency, as results and decisions drawn from these models have a significant impact in not only economic but also social context.

The matter of transparency in case of power system modeling has been tackled by various authors. Figure 1.1 summarizes the three main parts which have to be taken into account for open modeling: the software, data and content. The three layers of licensing include the bandwidth from full copyright to open licensing leading to three levels of transparency: public, scientific and completely open. While public means that inspection of code or raw data is possible, scientific transparency includes the usage of software, data and content. The complete open level enables not only the utilization but also the further development of these.

Furthermore, Morrison [9] highlights different licenses in the context of energy system analysis. It is important to distinguish between licenses for software, data and content. Typical

representatives include the GPLv3 license for software or the Creative Commons (CC) licenses for data and content. The later are also a topic of the recent study on the reuse policy of the European Union, conducted by the European Commission [10]. After their decision in 2011 to enable reuse of data and content of their institutions, a Reuse Notice was applied in 2013. The study explains how the Notice does not fulfill their targets of guaranteeing an easy reusability, hence, it reviews three alternatives: using CC licenses, using Open Data Commons licensing or deriving their own. They conclude that they are going to use the CC BY 4.0, which will be explained in more detail in chapter 2.

On the matter of publishing under open licenses, Pfenninger et al. [11] present their experience in opening up software and data in the context of energy modeling. They mention the difficulties in publishing data and software. The paper can be seen as a general introduction into working in energy system analysis and what main points an analyst has to consider. They conclude by emphasizing the importance of openness in energy research and the missing acknowledgment of software and data development. Next to their experience, the checklist of Cao et al. [12] might be helpful for an energy modeler starting to investigate open publishing. They base their list on several inquiries of experts and literature reviews with the aim to tackle two main problems regarding studies of power systems, namely knowledge transfer and transparency of the study. Both of these problems are also addressed within this thesis.

When working in energy system analysis, one often encounters models. Pfenninger, Hawkes, and Keirstead [13] review many models in the context of energy systems and sort them by challenges or use cases like simulation model or electricity markets. Another challenge they mention is uncertainty and transparency, which has to be tackled by today's energy system modelers. Savvidis et al. [14] focus more on energy models in context to how fit they are to meet policy decisions. They present an elaborate categorization system to group models and again handling uncertainty and the transparency criterion are mentioned.

A more recent review [15] focuses on 75 different models which include renewable energy sources. One of the main conclusions is that transparency and openness should be supported, especially for policy driven decisions. They see the value of models – even if they cannot foresee all possible technological advances – mainly in the insight they provide about the general workings and connections of these complex systems.

Groissböck [16] reviews solely open energy system models in his paper and tries to give an overview about their functionality. He concludes that while some models still lack certain important physical relations, the open-source models can already be used for various applications. So, the question arises why not more models and data are open in the context of energy system analysis.

Pfenninger et al. [17] summarize four main arguments on why models and data are not open (yet). Next to ethical and safety issues concerning open data, exposure and failure are feared by the scientists. These arguments are closely connected with the introductory paragraphs about the criticism in the scientific sector by Labini [8], which mentions the current tendency to more publicly presentable results while sacrificing scientific accurateness. This also relates to the third argument Pfenninger et al. [17] state: Proper documentation and labeling of model and data costs time a researcher usually does not have. The last point notes the general inertia of scientists and institutes avoiding struggling with cumbersome processes to retrieve republishable data.

These struggles were one of the main drivers for the Open Power System Data (OPSD) project [18]. The authors aimed at providing one central starting point for energy modelers and analysts regarding data. The available data includes generation capacities and time series for demand or renewable energy sources from different sources, including official state agencies. They try to make the data easily accessible and provide a comprehensive labeling over all sources. This project is one of those mentioned in the ongoing of chapter 2 to ease the accessibility and which are the base for the presented data processing.

1.1.2 Optimization and uncertainty

The next step in working with uncertainty in power system optimization concerns the underlying optimization problem. As soon as the input data is chosen and prepared for the energy system model, it is important to understand the mathematical formulation of the model, which is in most cases an optimization problem. There are different techniques on how to incorporate uncertainty in optimization problems. At first, some general approaches will be mentioned and then applications of three different methods – robust and stochastic optimization and hybrid techniques – will be explained.

The motivation and application of optimization problems for energy system analysis is reviewed by Banos et al. [19]. The authors categorize different optimization approaches by their application on renewable energy sources like wind, hydro or solar. Next to simplifications like linearizations, the complexity of the models is tackled with nonlinear approaches which include heuristics. The authors remark the increase in publications on these topics and the potential of research in the field of multi-objective optimization and parallel processing. Especially the later is of interest in the ongoing of the thesis.

Jordehi [20] reviews several methods on how to work with uncertainties in the case of energy system analysis. He distinguishes between two main fields, namely probabilistic and possibilistic approaches. Probabilistic methods rely mainly on the representation of the uncertainty by probability density functions. Hence, it involves the field of sampling and scenarios for the calculation of the uncertainties. Possibilistic approaches do not express the uncertainty of a parameter by probabilities but by fuzzy membership functions. The functions are defined over a fuzzy set, which describes the possible outcomes of the parameter [4, 10f]. A membership function is equal to one in the mid region of the set and has a linear ramp down to zero at the boundaries of the set. Hybrid approaches combine both methods, e.g. in the case of several uncertain parameters, where some have a probability density function and others are described by the membership functions. As the fuzzy representation of uncertainty is not considered for this thesis, the interested reader is referred to the book by Lodwick and Thipwiwatpotjana [21], which provides a good introduction to the topic.

Similar reviews have been conducted by Soroudi and Amraee [22] and Aien, Hajebrahimi, and Fotuhi-Firuzabad [23], which identify two main drivers of uncertainty in energy systems: uncertain parameters and uncertainty based on the modeling technique. Both reviews mention next to probabilistic and possibilitic methods the so-called information gap decision theory, interval analysis and robust optimization. All three methods have in common that there is no underlying function which describes the uncertainty (like density or membership function). For this thesis, mainly probabilistic, namely stochastic approaches are further used. Hence,

applications of that field are highlighted a bit more.

Before going into detail about stochastic optimization approaches, the concept of stages and decisions is introduced. Diwekar [24, 25] describes both terms *here-and-now* and *wait-and-see* clearly:

here and now The decision has to be made before the uncertain parameter takes its value.

wait and see A recourse action can be taken while/as soon as the uncertainty turns certain.

A stage of an optimization problem describes a defined time step or interval which is considered in the optimization. Optimization problems which incorporate uncertainty often consist of at least two stages: in the first stage, the *here and now* decision takes place, e.g. an investment decision, in the second-stage (*wait and see*) a recourse action (or observation and reaction [25, p. 151]) can be taken, e.g. generation scheduling of the power plant. In case of a multi-stage problem, each stage has its own *here and now* and *wait and see* decision, which could not only include investments or scheduling, but also storage content planning or considerations regarding emissions.

There have been many advances and applications in the field of probabilistic uncertainty modeling in the field of power system analysis. Typical uncertainties used with stochastic approaches include uncertain prices [26] and renewable feed-in as for both of them historical data is used to analyze frequency of certain events and generate probability density functions. When stages are combined with probabilistic approaches, it is often denoted as so-called *scenario trees*. These trees represent different realizations of the uncertain parameter per stage and can evolve quickly. Hence, methods to deal with the exponentially growing tree have been developed.

One of the main techniques to handle the exponential growth was developed based on dynamic programming by Pereira and Pinto [27]. They call their method Stochastic Dual Dynamic Programming (SDDP) because it involves the dual problem of the formulation. Each state is calculated separately, while costs for the subsequent stages are approximated by the so-called future-cost-function. The authors explain the difficulties of dynamic programming arising from the fact that the future costs are discretized for certain values of the decision variable of each stage. Their solution is to build up a piecewise linear function based on so-called "trial decisions" $\hat{x_1}$ and the dual problem of the future cost problem. The algorithm to solve the deterministic program is presented and called Dual Dynamic Programming (DDP). The algorithm has a forward and a backward simulation to calculate the trial decisions and the dual variables. Furthermore, they introduce the SDDP algorithm, as they are applying their DDP algorithm on a stochastic problem. The forward simulation to calculate the trial decisions is replaced by a monte carlo simulation, as the calculation would grow exponentially.

The general idea behind SDDP can be found under various terms in the literature. Another common term is nested Benders decomposition or the L-shaped method [28]. Rebennack [28] gives a good introduction into SDDP and the underlying similarities to Benders decomposition, which will be exploited further in the thesis.

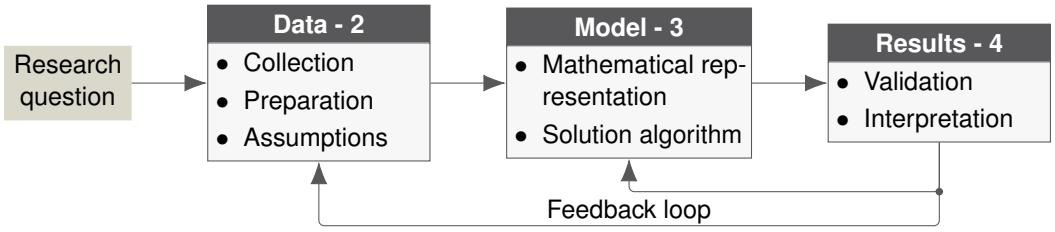


Figure 1.2: Structure of the thesis

1.1.3 Sensitivity analysis and interpretation of results

After dealing with uncertainty in the input and the underlying optimization problem, the model results have to be evaluated. The classical approach on how to deal with these uncertainties and especially on how to interpret values derived from simulation or optimization with such models is the so-called *sensitivity analysis* [1].

There are many ways in literature on how to analyze the sensitivity of any model (not necessarily uncertain). Typical approaches include parameter variations for the input and comparing the output [29], [30, p. 123] or analyzing local derivatives of the optimal solution by the dual variables [31, p. 177]. For both methods, it is important to understand which parameters might have a big impact on the solution. As Labini [8, p. 16] stresses out that especially for complex problems it is important to pick just a small amount of parameters and understand their impact on the system and, hence, understanding the system itself better.

Next to a sensitivity analysis, it is important to understand how to interpret the results based on a stochastic approach, which includes several possible paths in the future. What does each of these paths give as information? Which questions (like investment or operating decisions) can be answered by that approach? Diwekar [25, 150f] explains measures on how to compare deterministic models and problems with recourse, which will be used in the thesis.

1.2 Structure of thesis

Aim of the thesis is to motivate analysts to tackle uncertainties in their model with one of the provided ideas to get a better understanding of their model. Similar to the introduction, the thesis covers three main parts of energy system analysis, visualized in figure 1.2: data processing, model optimization and output analysis.

Chapter 2 highlights the relation of data, uncertainty and knowledge in more detail. Next to defining open data, processing techniques on how to openly work with data is presented.

The focus of chapter 3 lies within the underlying optimization problem and the formulation of uncertainty. It starts with an introduction into decomposition techniques and SDDP. The parallels of both methods are highlighted and a further development of Benders decomposition is applied to SDDP.

In chapter 4, a short analysis of the volatility of renewable energy sources is conducted. The analysis is the base for choosing one of these sources as uncertain parameter for the stochastic method and the impact of having an uncertain approach for this particular renewable energy source will be examined.

Finally, the main parts are summarized and an outlook for further development is given.

Chapter 2

Open data and open data processing

Missing or low quality data constitutes one of the main causes for uncertainty in power system analysis. This uncertainty is not only a problem in developing countries but also an issue in European data sets, which is illustrated in Figure 2.1. It compares the reported installed power plant capacities in Europe among different data sources. While deviations would be expected in case of photovoltaics, given their highly decentralized amount of small units, it might surprise that they also exist in case of conventional, large unit power plants. Deviations of 100GW in case of hard-coal-fired plants, which is more than twice as much installed power comparing two sources (CARMA and ENTSO-E), make a large difference in power system analysis.

As the missing data accuracy not only affects specific technical parameters but also key elements of the power system, it is a clear factor for uncertainty in power system analysis. This thesis does not solve this problem, but rather tries to enable the reader and potential analyst to understand assumptions and to reproduce the same aggregated input and, with help of an open-source framework, even the output of the presented models.

This chapter, in particular, deals with motivating the necessity for open data and its licensing. The chapter starts by introducing the concept of the *knowledge society* and its connection to data. It is followed by how business with data, especially in power systems, takes place and highlights the need for licenses. Additionally, an exemplary processing of input data is shown for a model of Germany and Bavaria for the utilized framework based on urbs. The developed framework is subsequently tested on basis of these data models.

2.1 Terminology

The basis of this chapter lies within the clarification on the terms of open data, their importance and processing. The next section deals with the concept of knowledge and data. In the literature, the most common link between those is given by the Data-Information-Knowledge-Wisdom (DIKW) hierarchy, illustrated in figure 2.2. Rowley [32] introduces the concept of the DIKW hierarchy and reviews various books and articles about the concept. As represented in the pyramid, wisdom is based on knowledge, which itself is based on information and so on. While there have been discussions about the terms which belong in the hierarchy and even the hierarchical structure itself, the underlying link between data and knowledge through information is always visible and agreed on.

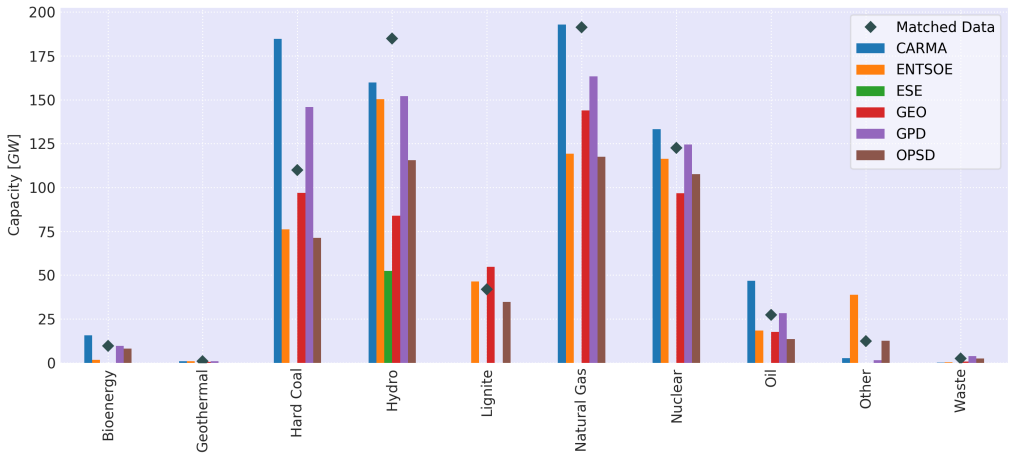


Figure 2.1: Comparison of installed capacities for different sources by data set (Frankfurt Institute for Advanced Studies, CC BY 4.0)

Zins [33] conducted a study on the definition of the terms data, information and knowledge by collecting the answers of 45 scholars in different fields of research about these terms. It becomes clear that the definitions diverge between the fields, although the connection itself holds.

The transformation process of how data becomes information and then knowledge diverges between the authors of both the review [32] and the study [33]. It becomes evident that one has to clarify the meaning of the terms before going into more detail. The next section 2.2 introduces the ideas about how this transformation process of data is seen in a theoretical concept in the context of the thesis. As many other authors, the top of the pyramid – wisdom – is not elaborated further on.

Section 2.3 focuses on the economic part of data, especially of governmental data (e.g. geospatial data or statistics), as well as of power system data. This discussion directly links to the concept of open data (when does data qualify as open?) and the licensing thereof in section 2.4.

The last section 2.5 comes back to the link from data to knowledge in the DIKW hierarchy with a practical example on how raw energy system data is processed to become the input for a power system modeling framework.

2.2 Knowledge and data

> Open knowledge is what open data becomes when it's useful, usable and used.
>
> Open Knowledge International [34], CC BY 4.0

Considering the above statement, a valid connection between knowledge and data can be drawn. It is important to understand the value of knowledge and data to appreciate the efforts

Figure 2.2: DIKW pyramid: The relation between data, information, knowledge and wisdom can be illustrated by a pyramid. While information is created out of data, knowledge is based on information. The highest form is gained by wisdom.

done to make them both open and available. This section is devoted to build the relation of the two and highlight its value.

The statement by Open Knowledge International [34] is also cited in the book of Wessels et al. [35]. The book deals with the effects of open data for society and introduces the term *knowledge society* in [35, Chapter 2]. The term *knowledge society* is defined with help of Castelfranchi [36], meaning that in contrast to an *information society*, the *knowledge society* does not only use information as the basis of production, consumption and innovation but rather that information leads to an improvement of the human society in general.

The term *knowledge society* is also the topic of [37]. The author provides an introduction about the concept of *knowledge societies*, their theory and background, and discusses in length the effects of knowledge, its production and value. As the economic effect of knowledge and its production is difficult to quantify, no explicit number can be given on how it affects the growth of an economy. In which parts of society or economy this growth comes into play exactly, and whether it leads to social welfare or is just an increase in specific sectors like the military sector, for example, is not certain for Stehr [37, 157f]. In his differentiated discussion about the consequences of knowledge, he mentions on the one hand the negative effects of knowledge such as repression and distortion in context of power. On the other hand, the positively perceived features like productivity and development are taken into account [37, 96f].

When Stehr [37, 99ff] talks about knowledge produced from science, especially the natural sciences, he sees a direct impact of increase of productivity by those sciences. He directly mentions “production of data and systems” [37, p. 103], which leads to an increase of knowledge productivity. An important argument lies within the fact that knowledge is not by itself productive, but needs an actor to control any kind of work or production within [37, p. 120].

Wessels et al. [35] take up the concept of Stehr’s knowledge societies and the role of science therein and link it with data. They state that not only the creation of data is done by science, but also the validation and analysis of the value of specific data [35, p. 27].

Coming back to the introductory statement, the term *knowledge* and its beneficial impact has been highlighted in the above paragraphs. However, the argument for openness is lacking. The paper by Cummings et al. [38] focuses on knowledge, the knowledge society and ways to implement those in two kinds of categories: techno-scientific-economic discourse and pluralist-

participatory discourse. The techno-scientific-economic view is held by national governments, whereas the pluralist-participatory view is located mostly in international organizations like the UNESCO [39] or the academic environment. The review by Cummings et al. [38] focuses on how these two discourses incorporate a successful strategy in fulfilling the United Nations Sustainable Development Goals [40]. The authors conclude that in case of the techno-scientific-economic view, the implementation ideas lack innovation and tend to be not changing the current situation right now and therefore not succeeding in reaching the goals. In case of the pluralist-participatory view, the authors see potential in creating a transformative environment for new collaborations and innovative projects to fulfill the goals. The discussion furthermore includes directives in how to proceed with knowledge: Knowledge should be freely and openly accessible, as it is seen as a public good. The value of knowledge is not only seen in terms of economic value, but additionally in a cultural and social context.

2.3 Business with data

> Data is the new oil.
>
> Clive Humby, 2006

This often quoted statement, probably originally stated by mathematician Clive Humby, shows the importance and the business possibilities behind data.

Recently, the statement is under discussion on whether it is still accurate. Taherivand [41] argued in his talk at the German “Das ist Netzpolitik”-conference that the metaphor is not correct in its current form, as data behaves more like ground water, as they describe it as a limitless resource, which reshapes itself constantly. However, data is seen as a valued commodity of economic interest. These statements are often used in debates about privacy and personal data in the digital age but it is also true in the context of energy systems for many years.

In 1909, Warren Cumming Platts started publishing the “National Petroleum News”, which reported prices in the oil sector [42]. The nowadays known company S&P Global Platts was founded by him and still sells information about oil, gas, coal and electricity. The provided data ranges from prices up to worldwide installed power plant capacities.

Similarly, the International Energy Agency, founded in 1974, started to publish its “World Energy Outlook” [43] in 1977. In 1978, the first “Basic Energy Statistics and Energy Balances of Developing Countries” was published. Nowadays, energy statistics are still available at their online shop, including free data, like cost assumptions for their World Energy Outlook, and paid data, like the “Electricity information” or the “World energy statistics and balances” with prices starting from 550€ per publication [44].

The question arises: If there are already offers to obtain data from data scientist experts (meaning that they have dealt with data for a long time, not assessing the provided data quality itself), why bother with open data? One obvious answer is cost-related concerns. A company might have a budget to buy data from commercial providers, but public institutions like universities or small companies might struggle with the acquisition.

Yet, there is also a public interest in data. To pick up the reasoning in section 2.2, data shapes a society by being a certain type of knowledge. A typical example for the public

interest in data is the Open Government Data (OGD) movement, which requests to open up governmental data. Ubalbi [45] gives an extensive introduction into the topic including benefits and risks regarding opening up data in government including the possibility to enable economic value as well as challenges in technical or organizational matters.

The inclusion of the community by opening up data and the general impact of open data in case of cities is the main topic of the book edited by Goldstein and Dyson [46]. It consists of various cases in which cities decided to open up their government data and the effect it had on several parts of city life, including transparency and new business opportunities.

Another example for OGD and its business potential is given in [47]. Dobusch et al. [47] analyzed the open geodata concept of the city of Linz with regard to financial risks and opportunities. The general open commons project of Linz comprised opening up various data sets of the government as part of the OGD initiative. One part of the data was geodata and included 3D renderings of buildings. The study came to the conclusion that while the business case of selling the respective data was lost, new opportunities arose in selling services which utilized the openly licensed data and that new projects within the open data community can have a similar impact.

Regarding national governments, the view on how to work with knowledge or data in specific is changing as well. As highlighted, the OGD initiative is on the rise and more and more governments affiliate themselves with this trend. The European Commission even summarizes possible categories for open publishing like geospatial data (postcodes, maps), energy consumption and emission levels or statistics (GDP, demographic data) [48]. But the initiative on opening up governmental data is not without question:

Kucera and Chlapek [49] discuss the advantages and disadvantages which come with OGD. Benefits include transparency, improvement of government services and increasing exchange between public and government organizations. Mentioned risks include privacy concerns, contraventions and security concerns. Measures to overcome these risks were also highlighted in the study and they cover anonymization techniques for data sets, compliance assessment by legislation and division between internal and external data sets.

Nevertheless, the European Parliament is planning to open up data like geospatial data and statistics and create a common data space [50]. The published briefing shows that the process in providing access to government data is still under development, but actively worked on. In parliament, the distribution of data is seen to potentially increase economic growth in Europe. In 2015, the business with data in the European Unions has already been more than €285 billion, equaling 1.94 % of the European Gross Domestic Product (GDP). But opening up data is not only beneficial in creating business cases regarding to the development of new services, but in increasing the efficiency of governmental services and an improved decision-making process.

The political aspect of open data can also be observed in the power sector. Transparency and politic decision play an important role in investment decisions in the system. Regarding the electricity infrastructure, public interest outranks the risks of opening up data, as the public is directly affected by investment decisions. It should be noted that the open data should not include exact blueprints of power plant structures, but rather the main data which is important for modeling and the decision process. Taking a step back, the question arises for what purpose data about the power system is needed. In general, it influences decisions on expansion,

operation and political trends concerning the power infrastructure from power plants to the distribution system.

A similar argument motivated the development of the Global Power Plant Database [51]. The authors argue that information of power plants affects not only investment decisions, but operation and reliability of the power system. Additionally, the impact on the climate due to emissions and fuel and water consumption is highlighted. This motivation resulted in a recent start of the openly licensed Global Power Plant Database, a project which summarizes not only power plant capacities, but also generation and other technical parameters. The project is open to the community, so everyone can improve the quality of the data and make suggestions.

To summarize, data in general is a business case. Opening the data to the public is not necessarily in conflict with business interest and the provision of data related services. Additionally, the benefits of transparency and participation of users usually outweighs the risks.

2.4 Open data licensing

> Especially on the Internet, more freedoms for the users and less control will often lead to higher revenues than all rights reserved paradigms.
>
> Till Kreutzer, 2016 [52, p. 14]

The current situation for most published content is that the author holds all the rights. This is expressed by the term *all rights reserved*. While this rule is generally appropriate, it might be outdated for content which is published on the internet. *All rights reserved*, which is still the standard of all published works, has its valid backgrounds, but can be seen as problematic in the context of the world wide web. Copyright infringements occur on a daily basis, as users mostly are unaware of the fact that visibility on the internet is not equal to having all rights in the public domain.

These problems are the basic motivation for why organizations like Creative Commons (CC) and Open Data Commons have decided to provide easily understandable licenses for content. In the following, the term *open data* is defined and licenses which guarantee this openness are presented. While the licenses are mostly defined for content in general, meaning not only data but pictures and texts as well, the section focuses entirely on the data part.

A general acknowledge definition of *open content* is not yet defined [52, p. 13], but as the focus in this thesis lies on open data, the definition of the open data handbook [53] is used. It can be summarized by the following key terms:

Availability and Access of the data must be granted easily, e.g. via the internet, for at most a reasonable cost in a preferably machine-readable form.

Re-use and Redistribution have to be ensured by granting users rights to work with the data, e.g. by providing a license to the data.

Organization	License	Description
Open Data Commons	Public Domain Dedication and License (PDDL)	Grants all rights to users
	Attribution License (ODC BY)	Naming the author
	Open Database License (ODbL)	Naming the author and redistribution of modified data with the same license
Creative Commons (CC)	CCZero (CC0)	Grants all rights to users
	Attribution 4.0 (CC BY 4.0)	Naming the author; also available in older versions (e.g. 3.0)
	Attribution Share-Alike 4.0 (CC BY-SA 4.0)	Naming the author and redistribution of modified data with the same license; also available in older versions (e.g. 3.0)

Table 2.1: Overview of different open content/data licenses conform to the open data definition based on [55]

Universial Participation means the possibility for everyone to contribute to and use the data without discrimination. This includes not restricting the data to non-commercial usages, as it would exclude commercial re-use of the data.

Data which is published under such terms is referred to as *open*. The European commission uses the terms findable, accessible, interoperable and reusable (FAIR) for their strategy on opening up data [54]. Additionally, they specifically name CC licenses as recommendation for open licenses.

Coming back to the introductory statement of the section, it becomes evident that only publishing the data online does not suffice as being open because of copyright laws and the standard *all rights reserved*. This is where licenses from organizations like CC and Open Data Commons come into play. The idea behind open content licenses was to make it easier for amateurs of the legal system to apply a suitable legal statement [52, p. 12]. The main reasons were adopted from the Free and Open Source Software (FOSS) movement for which licenses like the GNU General Public License is one of the most used licenses for open-source software.

Table 2.1 shows the open licenses which fulfill the definition of *open data* according to the above mentioned open data handbook. Both license systems of the organizations include several licenses. A common term, which is often applied to CC licenses, is the non-commercial (NC) term. It excludes the use of the content for commercial purposes. A probable reason for why it is excluded from the list of open data licences in Table 2.1 might be that the success of open data – be it in government or be it in science – is linked to economic possibilities which can be generated out of it. In this case, a NC license is a reason why interesting ideas for start-ups might not be realized. However, it has to be mentioned that even a data set, which is licensed under a NC license, might be used for commercial projects, if an agreement between the right holder and the respective project team is drawn [52, 19f].

Open licenses are not without criticism. Myška et al. [56] evaluate the effects of CC licenses. The authors explain the difference and close connection between legal language and law. They argue that the easification of legal language does not lead to higher legal certainty. They studied various webpages with online content released under a CC license and found a high amount of copyright infringements. A common problem is the wrong attribution in case of the CC BY license, as users do not know if and how to correctly attribute the author. They conclude that, while CC licenses have made a great impact in trying to make law more understandable, the authors are not sure whether the goal of CC can ever be accomplished.

From the perspective of this thesis, the concerns highlighted by [56] are valid, but miss some points: On the one hand, material that has been uploaded to websites without an open license encounters copyright infringement as well. On the other hand, CC licenses offer at least a guideline and starting point for non-legal experts to understand the issues behind copyright and uploaded material.

The same argumentation is used by Lupascu [57]. Nowadays, increased accessibility of material, which is due to publication in the world wide web, requires easy applicable licenses written in an understandable language. The access is not restricted to people with legal background, but to everyone. Everyone needs to be enabled to understand the valid license structure. In case of the specific license additions of CC, Lupascu [57] sees in the share alike (SA) license a guarantee to provide the follow-up work also to the same license, while Kreutzer [52, 22f] mentions the threat of license incompatibilities in case of SA terms. If a work is a collection of several different works licensed under the SA conditions, the author cannot publish and properly licenses his work without infringing at least one of those licenses. This is a common problem working with data from different sources.

In addition to license incompatibilities in case of data combinations, there is the distinction between *data* and *database*. Therefore, the license of a database can be a different one than the license of its content, meaning the data which the database contains. This might be beneficial, especially when contents are differently licensed, but this distinction might lead to more uncertainty and the confusion of users.

This problem is still not thoroughly solved, although there are different propositions on how to deal with it. There has been development in the open software community regarding the creation of easy copyright incompatibility tests [58]. Regarding data, the problems in the legal context inflicted by the open licenses are summarized in [59]. The authors state that the open data movement with its licenses is two-fold: on the one side an awareness of the unclear legal status regarding data and databases. On the other side, it led to forbidden limitations of actually open content by applying attribution restrictions or the already mentioned incompatibilities. The way to overcome these issues lies, according to the authors, in overcoming the need for licenses for data in the first place. They conclude with the following statement:

> We conclude, therefore, by claiming that the Open Data movement will only succeed after it has completely disappeared. [59, p. 22]

2.5 Open data processing

Coming back to the argumentation of Stehr [37, p. 120] that knowledge needs an actor to produce, the motivation for creating and publishing the following data processing steps reads as follows.

In this thesis, data is seen as a direct result or output of knowledge generation, which means data itself is a specific type of knowledge. Therefore, data is necessary, but not self-sufficient to enable anyone to use it productively. The gap can be filled by introducing a tool, which enables the target group to act on the data. This is where the publishing of the processing scripts comes into play. They are designed not only to enable a specific kind of energy modeler with the same background as the author, but also a wider spread range of scientists and people with basic programming skills to reproduce and, hence, act on the same basis as the author.

The processing scripts are written in a programming language with an openly available interpreter, namely Python. As parts of the code are specific to the respective data used for the example studies, not every part of the script is completely and immediately reusable to any case of power system. In general, the main goal of the script is to ensure a high reusability and reproducibility.

For this thesis, for two case studies (Bavaria and Germany) two Jupyter Notebooks for data processing (pre- and post-) have been written. The pre-processing scripts allow automatic aggregating of different data sources and prepare it for the framework which is used in this thesis. Additionally, visualization code blocks are included to create many of the following figures.

The framework is based on urbs [60]. In the following, the different parts of the notebook are explained with more detail. Post-processing features, such as model validation and visualization, are highlighted in chapter 4.

2.5.1 Modeled regions

As mentioned before, during the thesis, two models have been developed. The model for Bavaria is structured via its seven administrative regions, the German model by its 16 states plus one offshore region, accordingly. A list of the respective regions with its corresponding abbreviations can be found in table 2.2. Maps of both modeled regions are given in figure 2.3.

The decision on structuring the models by states or administrative regions does not result from a power system point of view. In fact, the structure might have been more advantageously set via the four Transmission System Operators (TSOs) regions in Germany and their mid-voltage correspondents. Instead, the reason for choosing the political structure is due to the fact of data availability. Statistics for validation are mostly given in an aggregated format for administrative, not technical regions.

2.5.2 Power plant capacities

As it has been visualized in the introduction to the chapter by figure 2.1, the data for power plant capacities diverges from source to source. This is not only the case for Europe, but also for Germany and Bavaria. Fortunately, the major differences can be explained.

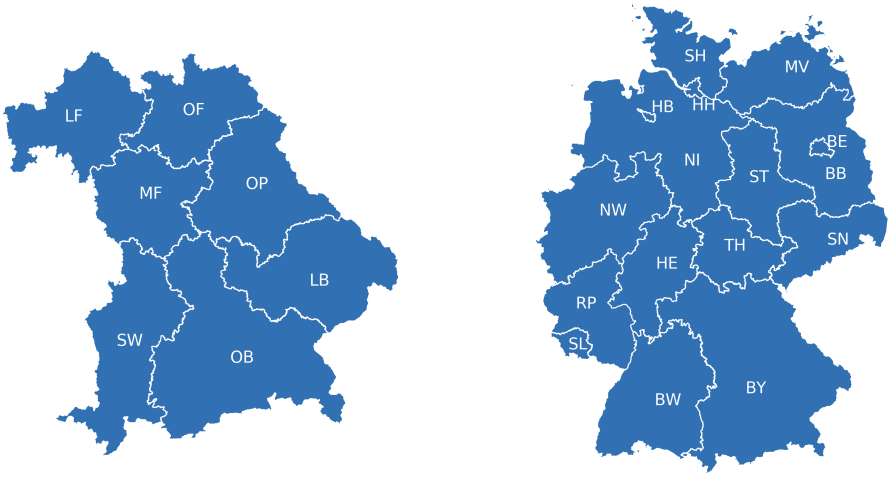


Figure 2.3: Bavaria with its seven administrative regions and Germany with is sixteen states

Bavaria	Germany (ISO 3166-2:DE)	Germany (English)
Lower Bavaria (LB)	Baden-Württemberg (BW)	Baden-Württemberg
Lower Franconia (LF)	Bayern (BY)	Bavaria
Middle Franconia (MF)	Berlin (BE)	Berlin
Swabia (SW)	Brandenburg (BB)	Brandenburg
Upper Bavaria (OB)	Bremen (HB)	Bremen
Upper Franconia (OF)	Hamburg (HH)	Hamburg
Upper Palatinate (OP)	Hessen (HE)	Hesse
	Niedersachsen (NI)	Lower Saxony
	Mecklenburg-Vorpommern (MV)	Mecklenburg-Vorpommern
	Nordrhein-Westfalen (NW)	North-Rhine Westphalia
	Rheinland-Pfalz (RP)	Rhineland-Palatinate
	Saarland (SL)	Saarland
	Sachsen (SN)	Saxony
	Sachsen-Anhalt (ST)	Saxony-Anhalt
	Schleswig-Holstein (SH)	Schleswig-Holstein
	Thüringen (TH)	Thuringia

Table 2.2: Used sites in model for Bavaria and Germany with the ISO 3166 abbreviations for the German states

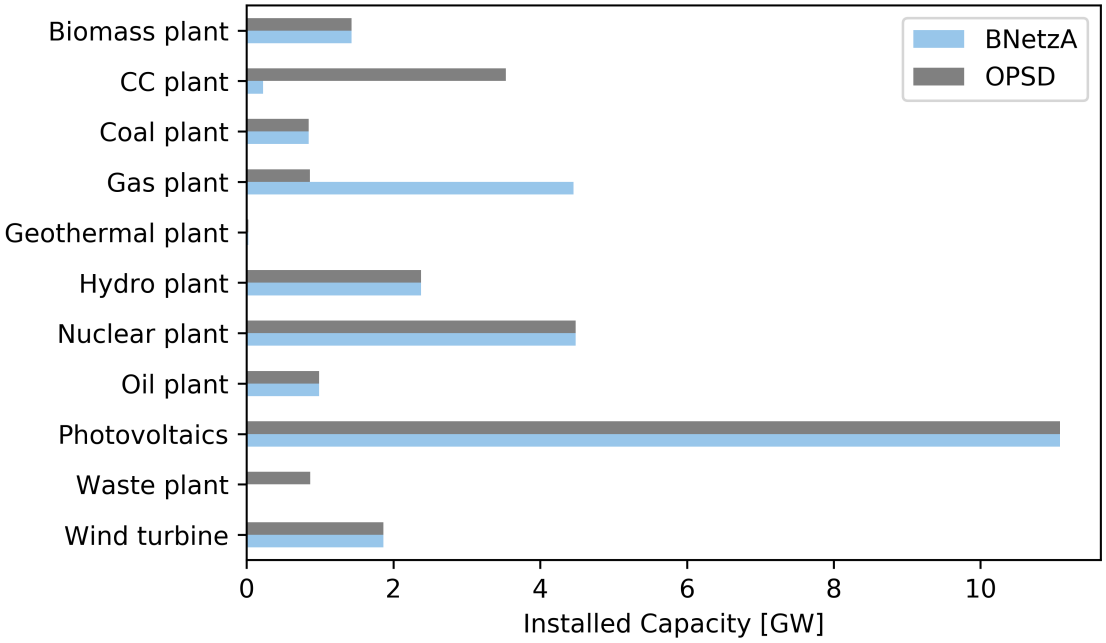


Figure 2.4: Installed capacities per commodity taken from Bundesnetzagentur (German Federal Network Agency) (BNetzA) [61] and Open Power System Data (OPSD) [62, 63] in Bavaria

Figure 2.4 shows the deviations of two sources for Bavarian power plant capacities. The main differences are for the commodities waste and gas. The problem lies within inaccurate tagging of plants, especially for gas. Some gas plants in the BNetzA list [61] are not correctly tagged as combined cycle plants. The OPSD [64] list is actually based on the BNetzA list, but more effort was put into correct tagging. The BNetzA list is licensed under the data license Germany – attribution – version 2.0 dl-de/by-2-0. This license has not been mentioned in the license comparison beforehand. It is similar to a CC BY license, as it may be used for both commercial and non-commercial purposes as long as it is correctly attributed.

Table A.3 and figure 2.5 show the installed capacities for every administrative region of Bavaria. The capacities of the states of Germany are gathered in table A.4 and visualized in figure 2.6.

2.5.3 Storage capacities

The difficulty in choosing an appropriate source for storage data arises in the mapping of storages from different sources: Some pumped-storages are geographically located outside of Germany, but as they only feed in their power production into the German grid, some data sources count them as German. As this thesis and the scenarios conducted later on are interested in the German power system, these specific storages are also counted as German and added to the respective state in which they feed in.

The sources, which are compared for the thesis, are represented in figure 2.7 and 2.8. BNetzA represents the German “Power plant list” from the Bundesnetzagentur [61]. ELMOD-DE is an open-source optimization model for the German electricity sector, written in GAMS.

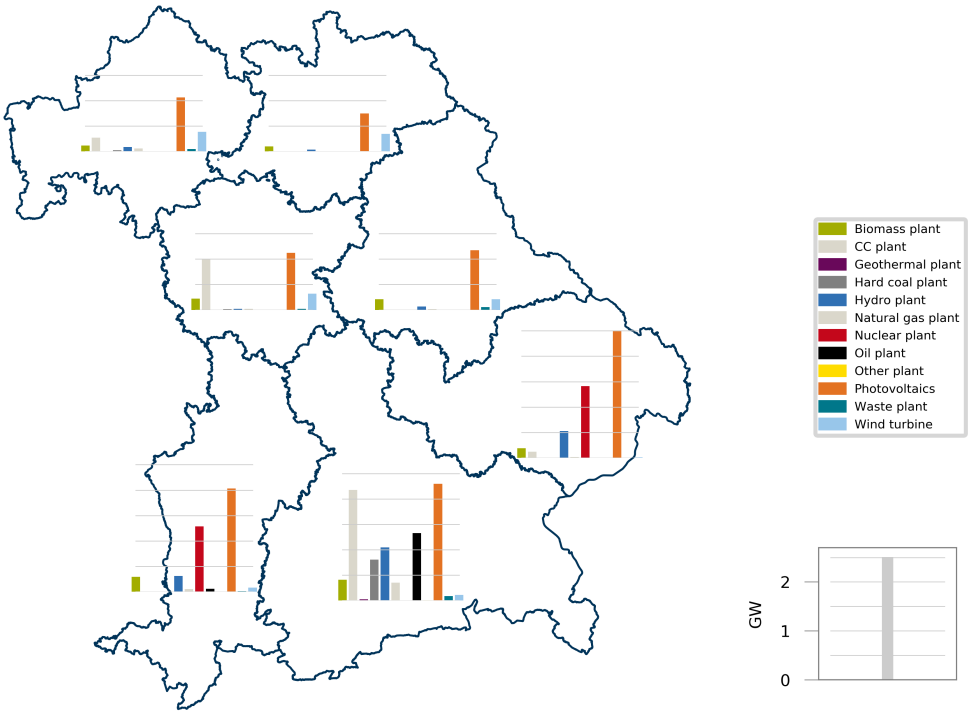


Figure 2.5: Installed process capacities in Bavaria by administrative region in GW [61, 62, 63]

Since GAMS can only be used by purchasing a license, therefore, ELMOD-DE has only been of interest regarding their input data [66]. Due to unclear licenses on the data, it is only used for comparison reasons to ensure reliable data.

FRESNA represents the accumulated data from the powerplantmatching project of the Frankfurt Renewable Energy Systems & Network Analysis Chair of FIAS university [67]. From this project, the introductory figure 2.1 has been used as well. FRESNA yields mostly smaller amounts of installed capacities as the other resources and is therefore not used.

For Germany, the same source as for the power plant capacities itself, data from OPSD [62] based on “Power plant list”, is used. As mentioned earlier, the original list of Bundesnetzagentur is licensed under dl-de/by-2-0. In case of Bavaria, however, storage capacity, which is located in Middle Franconia, is missing in OPSD compared to the BNetzA list, visualized in figure 2.8. All other capacities are equal. Therefore, the original BNetzA list is taken for the Bavarian model.

In table A.8 the technical parameters for the modeled pumped-storages are given. The parameters are used for both models (Bavaria and Germany). The respective installed capacities of each site are stated in table A.9 and A.10.

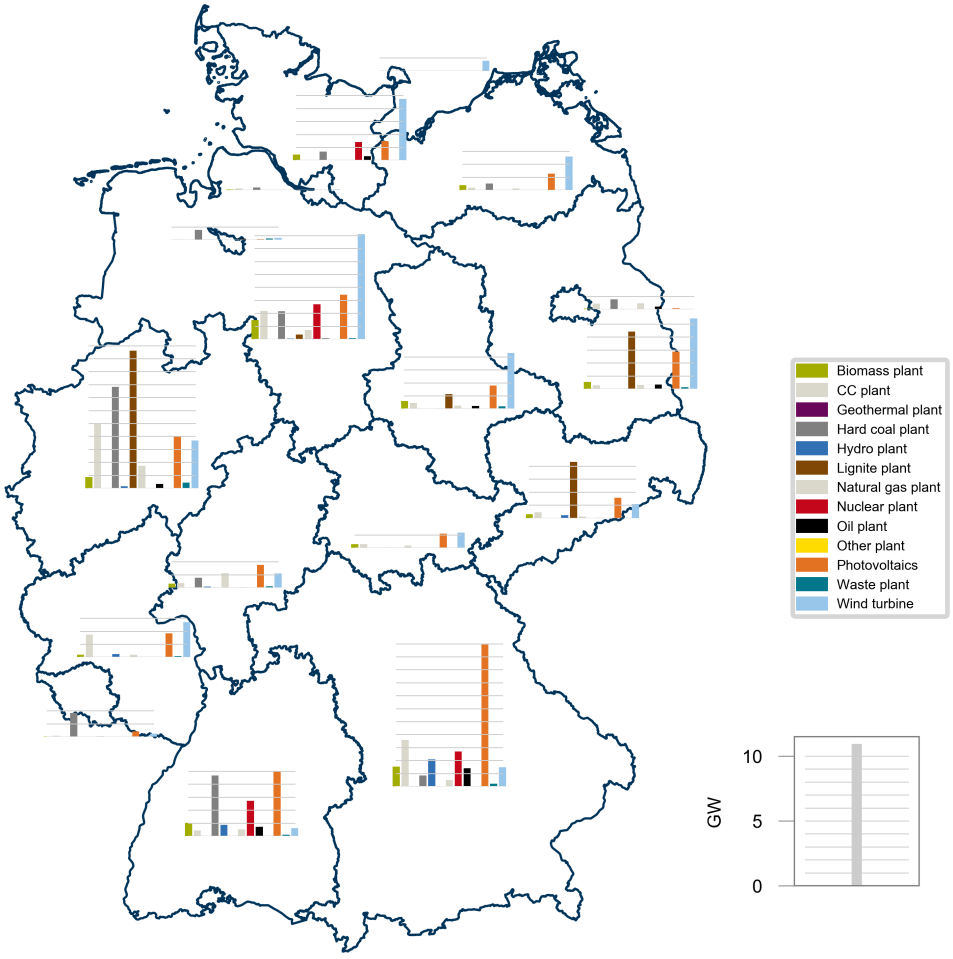


Figure 2.6: Installed process capacities in Germany by state in GW [61, 62, 63, 65]

2.5.4 Transmission grid

The capacity of transmission lines are calculated using the specific factors of the St. Clair curve [69] and the surge impedance load of the respective lines. The parameters used for the calculation are taken from the SciGrid project [68], which is based on OpenStreetMap data [70]. Hence, the database is licensed via ODbL v1.0 and its content via DbCL v1.0. Figure 2.9 shows the raw SciGrid data for Germany and the aggregated version. The line thickness represents the calculated capacities for the aggregated case.

2.5.5 Demand time series

The time series for this thesis for demand is modeled in an hourly resolution. The data for both models – Germany and Bavaria – is taken from the same sources and the methodology for

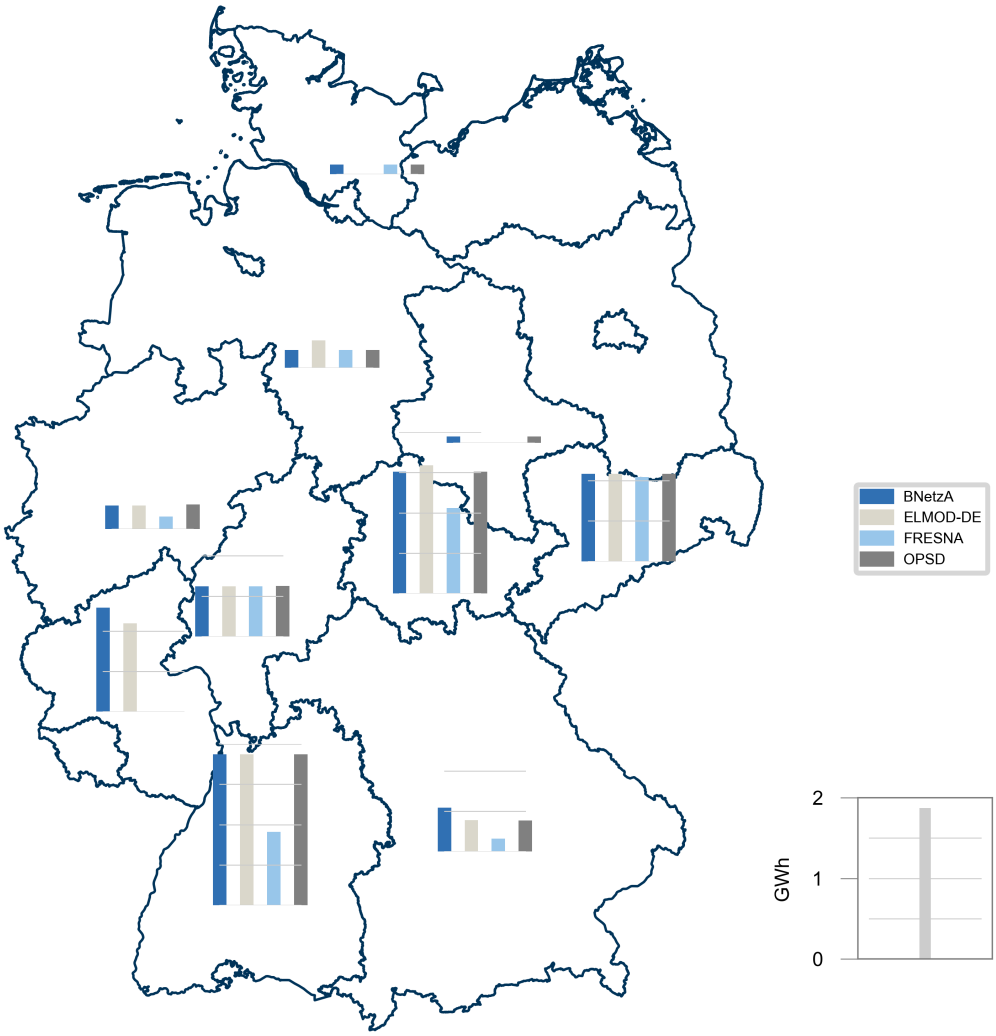


Figure 2.7: Comparison of different sources for pumped-storage power capacities [GWh] in Germany

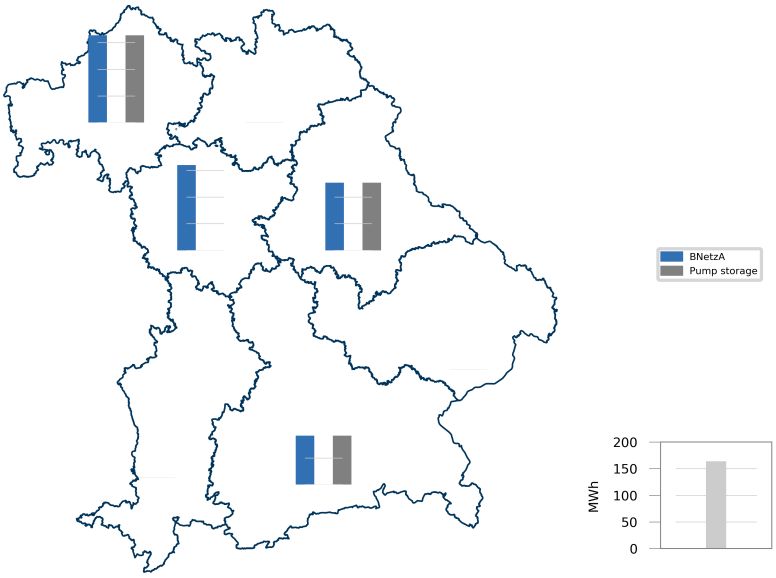


Figure 2.8: Comparison of different sources for pumped-storage power capacities [MWh] in Bavaria

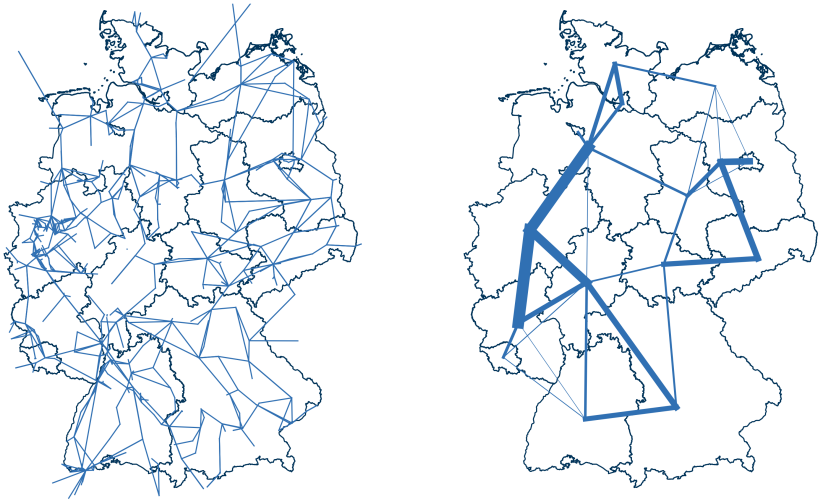


Figure 2.9: German transmission grid based on SciGrid [68] and its aggregated version: Thicker lines in the aggregated version indicate higher capacities of the lines.

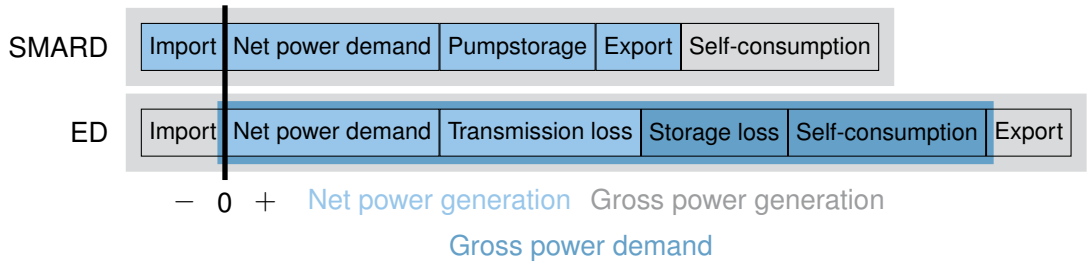


Figure 2.10: Definitions of power generation and demand terms from SMARD [71] and Energy Data [72]

deriving the time series is equivalent.

The difficulties in deriving consistent demand time series result from inconsistent definitions on the topic. The German Bundesnetzagentur started publishing time series of predicted and resulting demand in a fifteen-minute resolution under an open license on their SMARD platform [71].

The German Federal Ministry for Economic Affairs and Energy (FMEAE) compiles their data in the so-called Energy Data [72] report. It consists of a data sheet with historic data on energy (divided by sectors, primary energy, etc.) and graphs drawn from it. The data is given as yearly sums and can therefore not be used alone. In case of demand, they distinguish between net and gross power demand.

For the validation of the model, one resource, which would give not only the demand but also the corresponding generation to compare the results with, would be beneficial. As none such resource exists in the required temporal resolution, the sources have to be mixed. This would not be a problem, if both sources agree on the definitions of the terms. In case of SMARD and Energy Data, this is not the case.

Figure 2.10 summarizes the issues regarding the combination of using the temporal resolution of SMARD and the more detailed data sums of Energy Data and their counterparts in the state governments. SMARD publishes the time series of the net power demand and generation on their platform under the CC BY 4.0 license. While Energy Data (ED) describes the *net power generation* as sum of net power demand and transmission loss, SMARD includes import (negative) and export (positive) as well. SMARD does not explicitly mention the transmission losses in their description and does also not define the term *gross power demand*. ED uses it as sum of the net power generation, storage losses and power plant self-consumption. It can be seen that as both definitions vary highly, data from one source cannot be easily transferred into the other.

As input to the urbs-model, data on gross power demand would be useful, because self-consumption of power plants is not considered in the case study. On the other hand, storage and transmission losses are included in the optimization and can therefore be drawn from the results. The resulting modeled time series now consist of the four different time series for each transmission network operator in Germany [64]. The time series are weighted by the statistics of [73] (which in sum nearly correspond to the Energy-Data [72]). If two network operators are responsible for one state, both time series are used with a weighting factor. Figure 2.11 shows the resulting time series per state for one week in July.

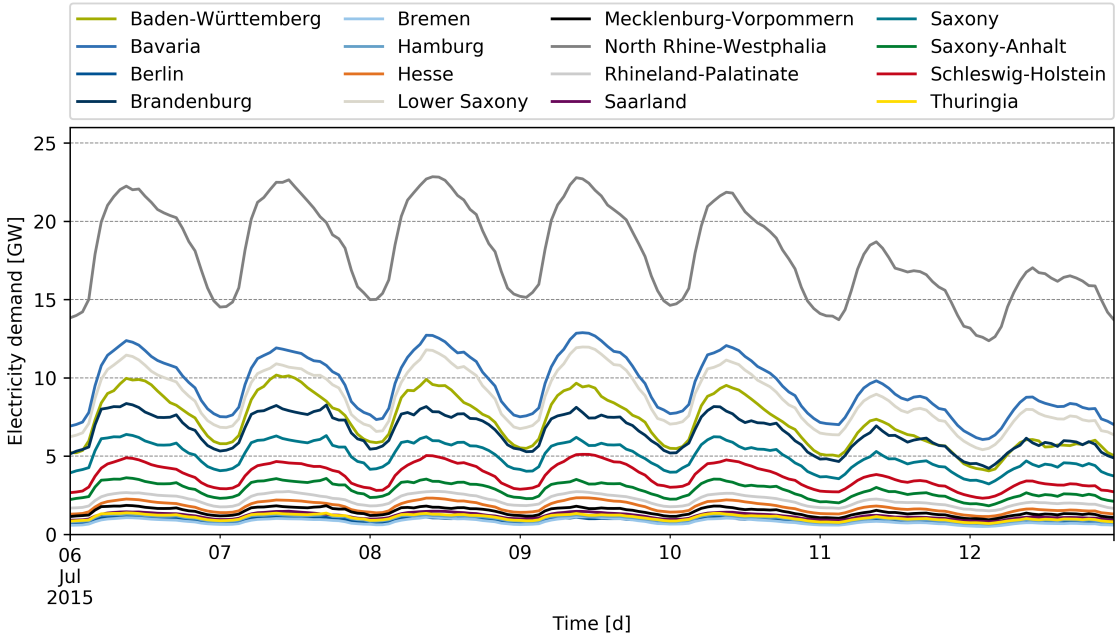


Figure 2.11: Resulting demand times series of Germany per state for one week

As data from the network operator is provided by ENTSO-E, no clear license regarding the republication for this kind of model could have been acquired. Instead, the method on how time series have been created based on the given input data is given in a Jupyter Notebook [74].

2.5.6 Solar and wind time series

The intermittent supply of these specific commodities is given in the respective sheet of the input spreadsheet. The percentage of available capacity per commodity is stated for each time step and site. Power output of one time step in one region due to solar and wind is therefore given by the installed capacity in that specific region multiplied with the capacity factor at that time.

The time series for the Bavarian and the German input file are collected from renewables.ninja, which are licensed under a creative commons license CC BY-NC 4.0: Pfenninger and Staffell [75] present their method on how the solar capacity factors are calculated based on solar radiation data from sources like MERRA-2 [76]. Similarly, this is done for wind based on measured wind speed [77].

For both time series, several parameters can be specified on what to retrieve. Starting from an exact location, a year can be chosen. Additional parameters include the capacity, the system loss and the ability to track the sunlight, i.e. the respective photovoltaic panel is installed in such a way that it follows the sun. Both tilt and azimuth of the installed panel can be chosen as well. As Germany has not only southern oriented panels, but also east and west orientation, the resulting time series for each region consists of a weighted sum of all three orientations.

In case of wind time series, the main parameters next to the capacity are the hub height and the turbine model. The turbine model includes information about cut-off wind speed and

the spectrum in which the turbine runs. An analysis of a distribution of wind turbines and heights over the modeled regions might be a starting point on how to improve both models as only height and type is used for all regions of one model.

2.5.7 Hydro time series

Similar to solar and wind, hydro power is taken as a predefined source. Therefore, a time series for the available power plant capacity in each time step and region is necessary. The time series, which have been used in this thesis, are based on measured water levels. Oeding and Oswald [78, 104ff] describe the characteristics of river hydro power which is dependent on the volume of river water. As a replacement, the water levels are taken to factor in seasonality like high water in the southern regions of Germany in summer. For each site, one main river is chosen to represent the hydro power output in that region.

In case of Germany, water time series have been requested from the German administration of waterways and shipping [79]. Data from that site is not specifically licensed with an open license, but redistribution of the adapted data has been granted for this thesis. In case of Bavaria, water levels are collected and can be downloaded from [80]. Fortunately, the data is published under a creative commons license CC BY 4.0.

The water level time series are available in a 15-minute resolution, which have to be prepared first: Data is provided with a time stamp either in daylight saving time or standard time. Before further processing the time series, it has to be transformed to one timezone. Afterwards, adaptations which result in a capacity factor can be made.

Firstly, the mean per hour is taken to acquire an hourly resolution. Oeding and Oswald [78] state that hydro plants mostly use only a part of the existing water in the river. Additionally, it is assumed that due to small dams at the hydro plant, a sudden change of the water level has not an immediate effect on the power output. Therefore, in the first preparatory steps, the time series is smoothed by bounding the slope in the time series, i.e. steps up and down cannot be indefinitely steep.

For each region, the full load hours of the hydro power plants are calculated via the installed capacities and the yearly power output of the statistics (Bavaria [81], Germany [73]). The bounded time series is hence normalized and weighted with the respective full load hours. Additionally, upon a certain point, more water in the river does not lead to higher production anymore. Therefore, the time series is transformed with a scaled tangens hyperbolicus as such that yielding values between $[0, 1]$.

Figure 2.12 shows the steps taken to generate the water time series. The first plot shows the water level of one week in June as raw input. In the next plot, the time series is reduced to the hourly resolution and with bounded slopes. The normalization with the full load hours is shown in the third plot. The transformation by the tangens hyperbolicus is shown in the last plot.

2.5.8 Technical parameters

The sheet *Global* contains data for constraints which affect all sites at all time. A typical example is a CO_2 limit for the whole model. It could for example represent a country's climate target.

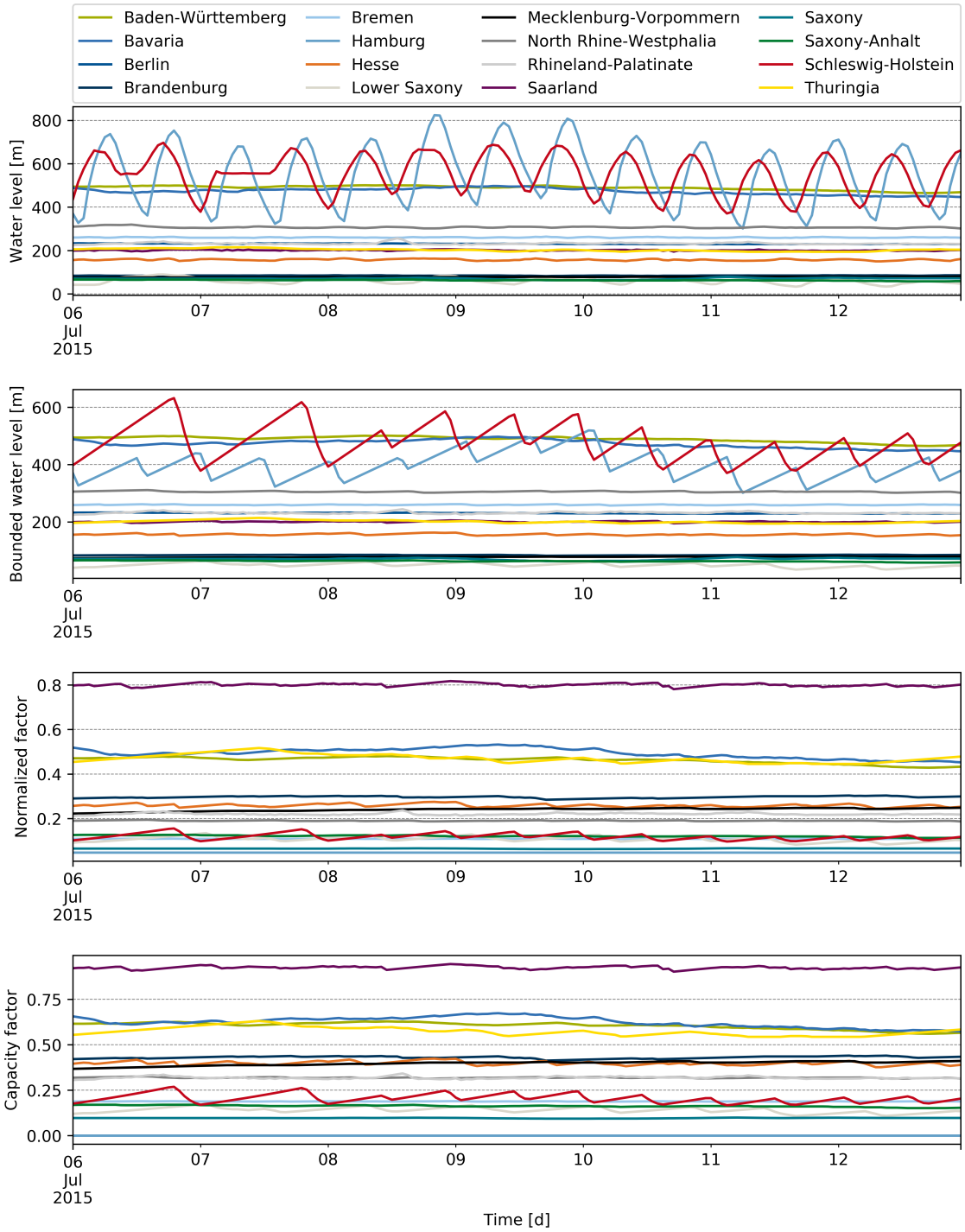


Figure 2.12: Generation of hydro time series shown exemplary for one week in Germany

For the models which represent the current status, the CO_2 values are taken from official statistics. Future scenarios represent climate goals set by the governments.

Table A.1 shows the global input data which was used for the model of Bavaria and Germany, respectively. The values are taken from statistics of 2015 [73].

2.6 Summary

It has been shown that it is – to some extent – possible, to gather data from open sources to model the electricity sector of Bavaria and Germany. Still, not all sources are licensed by one of the presented open licenses, hence, effort has to be made to contact right holders and agree on usage terms.

Major difficulties, like inconsistent data, gaps, incompatible sources have been highlighted and ideas given to overcome them. The resulting datasets [82] and openly available Jupyter notebooks [74], written in Python, therefore, provide methods and knowledge on how to handle the raw open data and therefore uncertainty provided by different sources.

One of the main aims of this chapter was to enable other energy modelers to work with similar data. The notebooks which were created for this thesis shall be seen as a tutorial on how to face common problems and prepare data for other models. In contrast to similar, readily available data collections and code snippets, the provided guide aims to be comprehensive, emphasizing the usage of available open data and illustrating methods to overcome problems associated with open data acquisition and handling. This means to show that it is not only the techniques for collecting data and the code that processes the data, which are important for gathering knowledge. The data itself holds within it other important lectures on issues a data scientist faces every day.

Chapter 3

Uncertainty in optimization problems

As one of the main causes for uncertainty — namely data — has been dealt with in the previous chapter, the question arises on how to deal with uncertainties in optimization problems.

In this chapter, the mathematical background of the used methods is introduced. The chapter starts with an introduction to decomposition techniques to split up the optimization problem into several smaller ones, which are not necessarily related to uncertainty modeling, but will play a vital role. It is followed by a general introduction and overview of uncertainty modeling techniques. The theory to Stochastic Dual Dynamic Programming (SDDP) is derived step-by-step and a new cut strategy for Benders decomposition is transferred to the stochastic case.

3.1 Decomposition techniques

There are various reasons for why optimization problems might be decomposed before being solved. One reason might be that the optimization variables have different properties, e.g. some variables might be integer while others are real numbers. A decomposition of the optimization across these differences can be beneficial in terms of solving options and time.

Other reasons for decomposition include saving time and resources: Large problems might not even be solvable on conventional computers anymore due to memory restrictions. In that case, it might be an option to decompose the large problem into several smaller subproblems and solve them separately [83].

There are several techniques for decomposing an optimization problem. The decision on which technique to choose depends mostly on the problem type and structure. In the case of linear problems, the most common approaches include the Dantzig-Wolfe algorithm [84] and Benders decomposition [85]. Nonlinear problems can be decomposed with adapted methods of the before mentioned algorithms or methods including Lagrangian relaxation, augumented Lagrangian decomposition or the auxiliary problem principle [86].

This thesis focuses on a linear expansion and dispatch optimization problem, therefore, only the main techniques dealing with linear problems are highlighted in the following.

For such linear problems, the objective function is easily decomposable. Hence, to analyze the structure of the constraints and the distribution of the variables over the constraints is

especially important for choosing a suitable method. Linear constraints can be coupled either through linked rows or columns of the constraint matrix. Both cases will now be explained.

3.1.1 Complicating constraints

Given is a general decomposable linear program in the following form:

$$\min_{\boldsymbol{\chi}} \boldsymbol{c}^{\mathsf{T}}\boldsymbol{\chi} \tag{3.1}$$

$$\text{s.t. } \boldsymbol{A}\boldsymbol{\chi} \geq \boldsymbol{b} \tag{3.2}$$

$$\boldsymbol{\chi} \geq \boldsymbol{0} \tag{3.3}$$

The objective function is a linear function with the real cost vector $\boldsymbol{c} \in \mathbb{R}^n$ and the vectors of variables $\boldsymbol{\chi} \in \mathbb{R}^n$. The m affine constraints consist of the matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ and the real vector $\boldsymbol{b} \in \mathbb{R}^m$. If the coupling in a linear problem is due to linked rows, it falls into the category of complicating constraints. This can easily be illustrated by looking at the constraint matrix.

In case the matrix $\boldsymbol{A}$ has a special structure as shown in (3.4), the optimization problem (3.1)-(3.3) is decomposable except for the lower constraints:

$$\boldsymbol{A} = \left[\begin{array}{c|c|c} \boldsymbol{A}^1 & & \\ \hline & \boldsymbol{A}^l & \\ \hline & & \boldsymbol{A}^L \\ \hline \multicolumn{3}{c}{\boldsymbol{E}} \end{array}\right] \tag{3.4}$$

There are several ways on how to decompose the presented structure. Dantzig and Wolfe [84] introduced their decomposition algorithm for linear programs with complicating constraints in 1960. The basic idea behind the algorithm, which is illustrated in detail by Conejo et al. [86], is to divide the optimization problems into several subproblems and solve the subproblems for different trial cost coefficients in the objective. The master problem searches for the cost optimal convex optimization of these trial solutions with respect to the complicating constraint.

Using the example structure (3.4) with L main blocks $\boldsymbol{A}^1$-$\boldsymbol{A}^L$ and one complicating block $\boldsymbol{E}$, the decomposed structure consists of L subproblems and one master problem. The algorithm starts with calculating the optimal solution of each subproblem for several trial cost vectors $\boldsymbol{c}$. Each resulting solution vector (corresponding to one cost vector) is used to calculate the respective value of the complicating constraint. The value of the trial objective function of the subproblems is used as weightning coefficient in the objective function of the master problem, while the values of the complicating constraint are used as weighting coefficients in the complicating constraint of the master problem. The master problem optimizes over the convex combination of all calculated trial values. The procedure is illustrated in algorithm 3.1.

Conejo et al. [86] mention that extensions of Dantzig-Wolfe decomposition can also be used for nonlinear problems. One section in their book is furthermore dedicated to deriving Dantzig-Wolfe decomposition from Lagrangian relaxation [86, Section 5.7.2, p. 236ff]. The relaxation method can be used for nonlinear optimization problems as well and uses relaxed primal and dual versions of the original problem for decomposition.

Algorithm 3.1: Dantzig-Wolfe algorithm

Data: $\boldsymbol{A}$, $\boldsymbol{b}$, $\boldsymbol{c}$

Initialization:

$\sigma \leftarrow \infty$

// Choose p random cost vectors $\boldsymbol{c}^p$

// Optimize all L subproblems with random cost vectors

$z_l^p \leftarrow \min_{\mathcal{X}_l} \boldsymbol{c}_l^p \boldsymbol{\chi}_l$ with $\mathcal{X}_l = \{\boldsymbol{A}_l \boldsymbol{\chi}_l \geq \boldsymbol{b}_l\}$

// Calculate z_p and $\boldsymbol{r}_p$ for all p random cost vectors

$z \leftarrow \sum_l z_l^p$

$\boldsymbol{r}_p \leftarrow \sum_l \boldsymbol{E} \boldsymbol{\chi}_l$

while $z < \sigma$ **do**

  // Solve master problem

  $\boldsymbol{\upsilon}, \boldsymbol{\lambda}, \sigma \leftarrow \arg\min_{\mathcal{U}} \sum_p z_p \upsilon_p$ with $\mathcal{U} = \left\{ \sum_p \boldsymbol{r}_p u_p = \boldsymbol{b} \ : \boldsymbol{\lambda},\ \sum_p u_p = 1 \ : \sigma,\ \boldsymbol{\upsilon} \geq \boldsymbol{0} \right\}$

  // Calculate new cost vector

  $\boldsymbol{c}^{p+1} \leftarrow \boldsymbol{c} - \boldsymbol{\lambda}^\mathsf{T} \boldsymbol{A}$

  // Optimize all l subproblems with random cost vectors

  $z_l^{p+1} \leftarrow \min_{\mathcal{X}_l} \boldsymbol{c}_l^{p+1} \boldsymbol{\chi}_l$ with $\mathcal{X}_l = \{\boldsymbol{A}_l \boldsymbol{\chi}_l \geq \boldsymbol{b}_l\}$

  // Calculate z_{p+1} and $\boldsymbol{r}_{p+1}$ for all $p+1$ random cost vectors

  $z \leftarrow \sum_l z_l^{p+1}$

  $\boldsymbol{r}_{p+1} \leftarrow \sum_l \boldsymbol{E} \boldsymbol{\chi}_l$

end

$\boldsymbol{\chi}^* \leftarrow \sum_p \upsilon_p \chi_p$

return $\boldsymbol{\chi}$, z

3.1.2 Complicating variables

In case of complicating variables, the structure of matrix $\boldsymbol{A}$ looks different: Instead of rows, which are linked to each other, columns lead to the coupling of the constraints:

$$\boldsymbol{A} = \left[\begin{array}{c|c|c|c} \boldsymbol{A}^1 & & & \\ \hline & \boldsymbol{A}^i & & \boldsymbol{E}_0 \\ \hline & & \boldsymbol{A}^I & \end{array}\right] \tag{3.5}$$

From duality theory, it can be derived that any problem with complicating variables can be transferred into a problem with complicating constraints and vice versa. The dual problem of (3.1) – (3.3) is given by:

$$\max_{\boldsymbol{\gamma}} \boldsymbol{b}^{\mathsf{T}}\boldsymbol{\gamma} \tag{3.6}$$

$$\text{s.t. } \boldsymbol{A}^{\mathsf{T}}\boldsymbol{\gamma} \leq \boldsymbol{c} \tag{3.7}$$

$$\boldsymbol{\gamma} \geq \mathbf{0} \tag{3.8}$$

The transpose $\boldsymbol{A}^{\mathsf{T}}$ of matrix $\boldsymbol{A}$ in case of complicating variables has a structure corresponding to the case for complicating constraints.

However, the classic approach to deal with complicating variables lies not within dualizing it and solving it by a Dantzig-Wolfe algorithm, but with a Benders decomposition approach. Benders [85] developed the algorithm to handle these kinds of problems. For an easier understanding, the problem is restated and reduced to one coupled block $\boldsymbol{A}_1$ instead of I and one uncoupled block $\boldsymbol{A}_0$ as follows:

$$\min_{\boldsymbol{\chi}_0, \boldsymbol{\chi}_1} \boldsymbol{c}_0^{\mathsf{T}}\boldsymbol{\chi}_0 + \boldsymbol{c}_1^{\mathsf{T}}\boldsymbol{\chi}_1 \tag{3.9}$$

$$\text{s.t. } \boldsymbol{A}_0\boldsymbol{\chi}_0 \geq \boldsymbol{b}_0 \tag{3.10}$$

$$\boldsymbol{E}_0\boldsymbol{\chi}_0 + \boldsymbol{A}_1\boldsymbol{\chi}_1 \geq \boldsymbol{b}_1 \tag{3.11}$$

$$\boldsymbol{\chi}_0, \boldsymbol{\chi}_1 \geq \mathbf{0} \tag{3.12}$$

The objective function is a linear function with rational cost vectors $\boldsymbol{c}_0 \in \mathbb{R}^{n_0}$, $\boldsymbol{c}_1 \in \mathbb{R}^{n_1}$ and vectors of variables $\boldsymbol{\chi}_0 \in \mathbb{R}^{n_0}$, $\boldsymbol{\chi}_1 \in \mathbb{R}^{n_1}$. Constraints (3.10) define the solution space of $\boldsymbol{\chi}_0$. Constraints (3.11) couple the complicating variables $\boldsymbol{\chi}_0$ with the decomposable variables $\boldsymbol{\chi}_1$.

In case of fixed complicating variables, the rest of the problem is easily solvable. Therefore, Benders [85] splits problem (3.1)-(3.3) into two (or more) problems. The problem corresponding to the complicating variable is called *master problem*, whereas the remaining parts are called *subproblems*.

In the master problem, the objective of the remaining cost function are substituted by a new variable α. The master problem is optimized over the constraint set belonging only to the complicating variable. The resulting value is then taken as input for the subproblems, which then can be easily solved. With help of the solution of the dual subproblems, a boundary or cut is then calculated to restrict the cost term α in the master problem. This process of solving master and subproblems and updating the variables and adding cuts is done until the algorithm converges.

The main problem in this approach lies within the cut generation. In general, two kinds of cuts are considered in most applications: a feasibility and an optimality cut. The former cut is needed in case one of the subproblems yields no solution as the constraint set is empty due to the choice of the master variables. An optimality cut is used if the choice of the master variables leads to a solution of the subproblems, but one that is still not good enough for the overall system. Another idea involves the unification of the cuts, meaning that a cut could not only be used as feasibility but as an optimality cut as well. The choice of the cuts and their generation is still under research. Rahmaniani et al. [87] reviewed recent development in Benders decomposition and mentioned cut selection as one of the four main development branches regarding Benders decomposition, the others being decomposition strategy, solution generation and solution procedure.

Classical applications for Benders decomposition in power systems include unit commitment problems. The complicating variables in this case are the integer variables, which are put into the master problem. Dependent on the status of the plants, decided in the master problem, the dispatch decisions are optimized in the subproblems.

As shown in the following sections, the structure and idea behind the algorithm is the same as for SDDP. Methods on refinement of Benders decomposition algorithm can therefore be used for the SDDP approach.

3.2 Modeling and handling of uncertainty

Uncertainty in power system analysis is encountered daily by scientists. Hence, many researchers focus on how to model and handle uncertain technical and economical parameters like load, generation and prices. One of the earliest advances in solving optimization problems with uncertainty was done by Dantzig [88] (reprint of the original paper of 1955). He derived a linear optimization problem with a stochastic approach. Before going into detail about stochastic programming and SDDP specifically, an introduction to uncertainty modeling is given.

3.2.1 Terminology and overview

The term *modeling* here refers to the mathematical representation of uncertain events or parameters. Whether the uncertainty should be dealt with in an optimization problem or a simulation approach is not necessarily defined by the modeling. Soroudi and Amraee [22] and the more recent review of Aien, Hajebrahimi, and Fotuhi-Firuzabad [23] provide an overview of many techniques to model uncertainty in the context of power systems.

They all agree on the characterization in six categories:

Probabilistic approach Probability Density Function (PDF) based

Possibilistic approach Membership functions

Hybrid possibilistic-probabilistic approach Combining PDF and membership functions

Information gap decision theory Deviation error measurement

Robust optimization Deterministic uncertainty sets

Interval analysis Determining bounds of interval

Soroudi and Amraee [22] state examples for the given methods, like operation and planning of renewable energy sources, transmission, and electricity markets. Probabilistic approaches can be found for every mentioned application, hybrid approaches or interval analysis are not yet that commonly used.

Whereas the description of methods mostly focuses on the determination of a suitable PDF or intervals, the main goal of this thesis is to explain how to solve optimization problems with included uncertainty. In case of robust optimization [89], for example, as the uncertainty is

modeled via a deterministic uncertainty set, the resulting optimization problem does not face the issue of exponentially increasing variables as it does for stochastic optimization.

Typical issues concerning robust optimization problems [90] do include the formulation of the uncertainty set in particular and the handling and solving of the minimax problem or the dualized resulting bilinear problem [91, 92, 93, 94, 95]. As one of the main points of criticism against robust optimization is the conservatism of the solution, advantages in combining robust optimization with a method based on a probabilistic modeling approach like stochastic optimization can be found in various studies [96, 97, 98, 99, 100].

3.2.2 Handling uncertainty with stochastic optimization

In case of stochastic optimization, uncertainty is measured with help of probabilities. The uncertain parameters are described by random variables on a given probability distribution.

There are many good introductions to stochastic optimization like the book by Birge and Louveaux [101] or by Shapiro, Dentcheva, and Ruszczyński [102]. In the following the main terms and solution ideas are highlighted.

3.2.2.1 Terminology

Given a multi-stage linear stochastic programming problem:

$$\min_{\boldsymbol{\chi}_0} \boldsymbol{c}_0^{\mathsf{T}}\boldsymbol{\chi}_0 + \mathbb{E}\left[\min_{\boldsymbol{\chi}_1} \boldsymbol{c}_1(\omega_1)^{\mathsf{T}}\boldsymbol{\chi}_1(\omega_1) + \mathbb{E}\left[\cdots + \mathbb{E}\left[\min_{\boldsymbol{\chi}_T} \boldsymbol{c}_T(\omega_T)^{\mathsf{T}}\boldsymbol{\chi}_T(\omega_T)\right]\right]\right] \tag{3.13}$$

$$\text{s.t. } \boldsymbol{A}_0\boldsymbol{\chi}_0 = \boldsymbol{b}_0 \tag{3.14}$$

$$\boldsymbol{E}_0(\omega_1)\boldsymbol{\chi}_0 + \boldsymbol{A}_1(\omega_1)\boldsymbol{\chi}_1(\omega_1) \geq \boldsymbol{b}_1(\omega_1)\ \forall \omega_1 \in \Omega_1 \tag{3.15}$$

$$\vdots$$

$$\boldsymbol{E}_{T-1}(\omega_T)\boldsymbol{\chi}_{T-1}(\omega_{T-1}) + \boldsymbol{A}_T(\omega_T)\boldsymbol{\chi}_T(\omega_T) \geq \boldsymbol{b}_T(\omega_T)\ \forall \omega_{T-1} \in \Omega_{T-1},\ \omega_T \in \Omega_T|\omega_{T-1} \tag{3.16}$$

$$\boldsymbol{\chi}_0, \boldsymbol{\chi}_t(\omega_t) \geq \boldsymbol{0}\ \forall t \in \{1, \ldots, T\} \tag{3.17}$$

A possible event or realization in time step or stage t is denoted by ω_t. The cost function has a nested form by including the expected or future costs for the following stages. The properties are summarized as with [28, p. 347]:

Fixed recourse $\boldsymbol{A}_t(\omega_t) = \boldsymbol{A}_t\ \forall t \in \{1, \ldots, T\}$

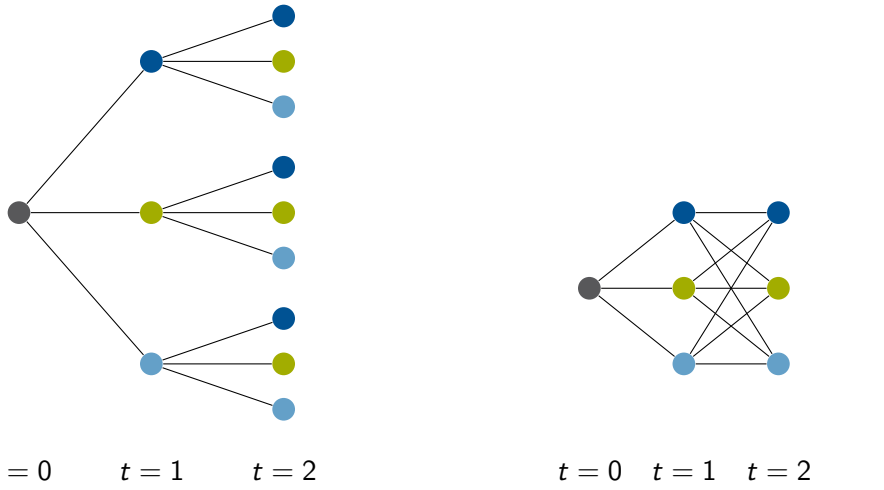
Fixed technology $\boldsymbol{E}_t(\omega_t) = \boldsymbol{E}_t\ \forall t \in \{1, \ldots, T\}$

Fixed right-hand-side $\boldsymbol{b}_t(\omega_t) = \boldsymbol{b}_t\ \forall t \in \{1, \ldots, T\}$

Fixed objective coefficients $\boldsymbol{c}_t(\omega_t) = \boldsymbol{c}_t\ \forall t \in \{1, \ldots, T\}$

Dependent on which parameters are uncertain random variables, different properties like convexity of the expectation functions hold with respect to the decisions of the respective stage.

For the stochastic optimization problems, the following terms are defined:



(a) Scenario tree for stochastic programming (b) Branched scenario tree for SDDP

Figure 3.1: Comparison of classic stochastic programming to SDDP: colors represent different realization per time step t

Relatively complete recourse All stages of the stochastic optimization problem are feasible, regardless of the feasible decisions of previous stages [103].

Complete recourse A special case for relatively complete recourse; easier to prove as only knowledge about the matrix $\boldsymbol{A}$ is necessary [101, p. 113] (c.f. fixed recourse). Even choices which lead to infeasible previous stages can be compensated by recourse action [104].

Nonanticipativity During a decision in time step t, future observations and informations cannot be taken into account. Only past decisions are known [101, p. 21].

Figure 3.1a represents the exponential growth of a scenario tree. Even if nodes in the tree in one stage have the same value (visualized by the same colors), the paths are not connected because the history of the respective value is kept.

Murphy [105] gives an easily understandable introduction to the topic of using Benders decomposition for stochastic programming. He also highlights the necessary steps to get from a continous uncertainty to the deterministic stochastic programming problem and explains the usage of Benders and the so-called nested Benders decomposition for the multi-stage case.

3.2.2.2 Stochastic dual dynamic programming

Instead of building a scenario tree with the same values on stages in different branches, each node in one stage is potentially reachable from the previous and the following. In figure 3.1, the comparison between the two concepts is illustrated with the scenario trees. With SDDP, the different branches stay connected in each stage, which leads to a branched tree like in figure 3.1b. The tree representation visualizes the distinction which can be made by solving the problem: For a general stochastic calculation, for every realization, i.e. point in the tree, all variables need to be defined separately. The objective function, hence, grows exponentially

with the number of time steps. In case of SDDP, the information which is calculated for one realization is used for more than one branch. The nested representation often used SDDP highlights these usages.

Pereira and Pinto [27] introduced the SDDP algorithm and derived it from dual dynamic programming. It is one of the central starting papers on the derivation of SDDP parallel to Dantzig and Infanger [106, 107]. In the literature, the algorithm cannot only be found by that term, but also as so-called nested Benders decomposition [28], [105]. On basis of these papers, the derivation of SDDP with help of the concept of dual dynamic programming is done in the following.

Dual dynamic programming As an introduction to the concept, a linear programming problem is given in the same structure as already encountered for the introduction of Benders decomposition:

$$\min_{\boldsymbol{\chi}_0, \boldsymbol{\chi}_1} \ \boldsymbol{c}_0^\mathsf{T}\boldsymbol{\chi}_0 + \boldsymbol{c}_1^\mathsf{T}\boldsymbol{\chi}_1 \tag{3.18}$$

$$\text{s.t. } \boldsymbol{A}_0\boldsymbol{\chi}_0 \geq \boldsymbol{b}_0 \tag{3.19}$$

$$\boldsymbol{E}_0\boldsymbol{\chi}_0 + \boldsymbol{A}_1\boldsymbol{\chi}_1 \geq \boldsymbol{b}_1 \tag{3.20}$$

$$\boldsymbol{\chi}_0, \boldsymbol{\chi}_1 \geq \boldsymbol{0} \tag{3.21}$$

The problem (3.18)-(3.21) can be seen as two stage problem. For an overall optimization, both variables $\boldsymbol{\chi}_0$ and $\boldsymbol{\chi}_1$ have to be optimized. Unfortunately, the realization of $\boldsymbol{b}_1$ and hence the optimal value of $\boldsymbol{\chi}_1$ is not known during the optimization of $\boldsymbol{\chi}_0$. On the other hand, the optimization of $\boldsymbol{\chi}_1$ itself depends on the realization of $\boldsymbol{\chi}_0$ because of constraint (3.20).

In the report [27], the second term of the objective function (3.18) is called *future cost* and abbreviated by the term $\alpha(\boldsymbol{\chi}_0)$:

$$\alpha(\boldsymbol{\chi}_0) = \min_{\boldsymbol{\chi}_1} \ \boldsymbol{c}_1^\mathsf{T}\boldsymbol{\chi}_1 \tag{3.22}$$

$$\text{s.t. } \boldsymbol{A}_1\boldsymbol{\chi}_1 \geq \boldsymbol{b}_1 - \boldsymbol{E}_0\boldsymbol{\chi}_0 \tag{3.23}$$

$$\boldsymbol{\chi}_1 \geq \boldsymbol{0} \tag{3.24}$$

The dependence of $\boldsymbol{\chi}_1$ on $\boldsymbol{\chi}_0$ becomes evident by looking at the optimization problem (3.22)-(3.24). With help of the future costs, the optimization problem (3.18)-(3.21) can be reformulated as follows:

$$\min_{\boldsymbol{\chi}_0} \ \boldsymbol{c}_0^\mathsf{T}\boldsymbol{\chi}_0 + \alpha(\boldsymbol{\chi}_0) \tag{3.25}$$

$$\text{s.t. } \boldsymbol{A}_0\boldsymbol{\chi}_0 \geq \boldsymbol{b}_0 \tag{3.26}$$

$$\boldsymbol{\chi}_0 \geq \boldsymbol{0} \tag{3.27}$$

If the future cost function $\alpha(\boldsymbol{\chi}_0)$ is known, the optimization problem (3.25)-(3.27) becomes a standard linear program which can easily be solved. As it is unknown, an approximation method is needed to handle the unknown term.

The idea of dual dynamic programming is to approximate the future costs $\alpha(\boldsymbol{\chi}_0)$ by calculating its value for a certain set of $\boldsymbol{\chi}_0$, the so-called *trial values* $\mathcal{T} = \{\hat{\boldsymbol{\chi}}_{1i},\ i = 1, \ldots n\}$.

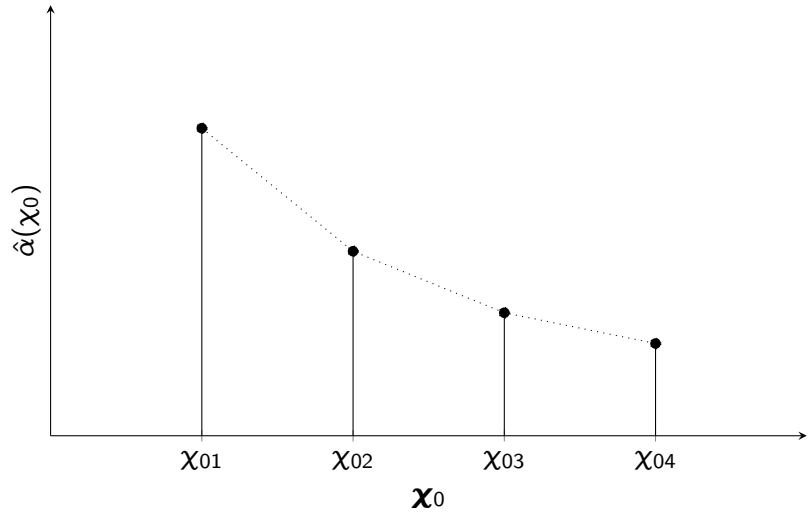


Figure 3.2: Piecewise linear future cost function $\hat{\alpha}(\boldsymbol{\chi}_0)$ with trial values $\hat{\boldsymbol{\chi}_{1i}}$ for example $n = 4$.

The values of $\alpha(\boldsymbol{\chi}_{1j})$ for $\boldsymbol{\chi}_{1j} \notin \mathcal{T}$ stay unknown. A continuous function $\alpha(\boldsymbol{\chi}_0)$ is defined by creating hyperplanes connecting the trial values $\hat{\boldsymbol{\chi}}_{1i}$. As $\alpha(\boldsymbol{\chi}_0)$ is not only continuous but also convex, the maximum of the hyperplanes represents the function. This results in a piecewise linear function $\hat{\alpha}(\boldsymbol{\chi}_0)$, which is illustrated in figure 3.2 for an example with $n = 4$ trial values.

A relaxation of the piecewise linear function is used to compensate the problem of dimensionality: The number of trial values needed to represent the future cost function increases exponentially by the dimensionality of the vector $\boldsymbol{\chi}_0$.

In the formulation of the future costs (3.25)-(3.27), the dependence lies within the definition of the constraint set. Deriving the dual problem the dependence will be shifted to the objective function. This means that the feasible region of the dual problem is not dependent on the first stage parameter and therefore, the constraint set can be completely represented by a finite number of cuts.

The derivation of the relaxation of the piecewise linear function uses the dual problem of the second-stage problem (3.22)-(3.24) for the mentioned reasons. This results in the following problem:

$$\alpha(\boldsymbol{\chi}_0) = \max_{\boldsymbol{\gamma}} \boldsymbol{\gamma}^{\mathsf{T}}(\boldsymbol{b}_1 - \boldsymbol{E}_0\boldsymbol{\chi}_0) \tag{3.28}$$

$$\text{s.t. } \boldsymbol{\gamma}^{\mathsf{T}}\boldsymbol{A}_1 \leq \boldsymbol{c}_1 \tag{3.29}$$

$$\boldsymbol{\gamma} \geq \mathbf{0} \tag{3.30}$$

Looking more closely at problem (3.28)-(3.30), it can be seen that the constraint set, and therefore the feasible set, stays the same for every realization of $\boldsymbol{\chi_0}$. Hence, this problem corresponds to find the vertice of $\boldsymbol{\gamma}^{\mathsf{T}}\boldsymbol{A}_1 \leq \boldsymbol{c}_1$ which results in the maximal value of the objective for a given value $\boldsymbol{\chi}_0$. As the possible optimal solutions, i.e. the vertices of the constraint set (3.29) are not dependent on $\boldsymbol{\chi}_0$, it gets clear why the approximation of the future costs $\alpha(\boldsymbol{\chi}_0)$ is achievable by calculating the hyperplanes $\boldsymbol{\gamma}^{\mathsf{T}}(\boldsymbol{b}_1 - \boldsymbol{E}_0\boldsymbol{\chi}_0)$ for different vertices of the constraint set. This concept is elaborated by the following example.

Plant/Storage	Costs per unit	Capacity
Wind	2	variable
Gas	5	5
Battery	1	10

Time step	Demand	Capacity Wind
1	1	3
2	3	2

Table 3.1: Data for Dual Dynamic Programming (DDP) example 3.2.1

Example 3.2.1 *A company has one wind and one gas turbine and a battery storage for its power demand. The cost of generating power with the wind turbine is lower than the cost of using the gas turbine for power generation. The wind turbine's available capacity for the second time step is not explicitly known in the first. Table 3.1 summarizes the relevant information.*

The resulting optimization problem reads as follows:

$$
\begin{aligned}
\min_{\boldsymbol{\chi}_W, \boldsymbol{\chi}_G, \boldsymbol{\chi}_B} \quad & 2\chi_{W_0} + 5\chi_{G_0} + \chi_{B_0} + 2\chi_{W_1} + 5\chi_{G_1} + \chi_{B_1} \\
s.t. \quad & \chi_{W_0} + \chi_{G_0} - \chi_{B_0} = 1 \\
& \chi_{B_0} + \chi_{W_1} + \chi_{G_1} - \chi_{B_1} = 3 \\
& \chi_{W_0} \leq 3 \\
& \chi_{W_1} \leq 2 \\
& \boldsymbol{\chi}_G \leq \mathbf{5} \\
& \boldsymbol{\chi}_B \leq \mathbf{10} \\
& \boldsymbol{\chi}_P \geq \mathbf{0} \; \forall P \in \{W, G, B\}
\end{aligned}
$$

The variables χ_{W_t} and χ_{G_t} represent the power output of the wind and gas turbine, respectively, for time step t. The storage content of the battery in time step t is given by variable χ_{B_t}. It is assumed that the storage is empty when the optimization starts.

The solution to this optimization problem can be drawn by inspection. As the wind power availability in the second time step does not suffice in order to meet the demand in the second time step, it is necessary to produce one power unit more from wind power in the first time step and store it. So the resulting optimal values are stated in table 3.2 with the resulting costs of 9. The future cost function in this case represents the second stage costs depending on the storage content χ_{B_0}. For understanding the approximation of the future costs functions, the problems per stage are stated and analyzed closely.

Time step	Wind	Gas	Battery
1	2	0	1
2	2	0	0

Table 3.2: Result for DDP example 3.2.1

The problem of the first stage reads as follows:

$$\begin{aligned} \min_{\chi_{W_0},\chi_{G_0},\chi_{B_0}} \; & 2\chi_{W_0} + 5\chi_{G_0} + \chi_{B_0} + \alpha(\chi_{B_0}) \\ s.t. \; & \chi_{W_0} + \chi_{G_0} - \chi_{B_0} = 1 \\ & \chi_{W_0} \leq 3 \\ & \chi_{G_0} \leq 5 \\ & \chi_{B_0} \leq 10 \\ & \chi_{P_0} \geq 0 \; \forall P \in \{W, G, B\} \end{aligned}$$

It is important to note that the future costs is not dependent on χ_{W_0}, but on the storage content χ_{B_0} as it is the complicating variable of the time steps. The future cost reads as follows:

$$\begin{aligned} \alpha(\chi_{B_0}) = \min_{\chi_{W_1},\chi_{G_1},\chi_{B_1}} \; & 2\chi_{W_1} + 5\chi_{G_1} + \chi_{B_1} \\ s.t. \; & \chi_{B_0} + \chi_{W_1} + \chi_{G_1} - \chi_{B_1} = 3 \\ & \chi_{W_1} \leq 2 \\ & \chi_{G_1} \leq 5 \\ & \chi_{B_1} \leq 10 \\ & \chi_{P_1} \geq 0 \; \forall P \in \{W, G, B\} \end{aligned}$$

It can be seen that the dependency of the future cost function on χ_{B_0} lies in the constraint set. For getting a fixed constraint set, the dual problem of the second stage has to be created. Then, the dependency will shift to the cost function. The dual problem of the second stage results in:

$$\begin{aligned} \alpha(\chi_{B_0}) = \max_{\gamma,\gamma_W,\gamma_G,\gamma_B} \; & \gamma(\chi_{B_0} - 3) - 2\gamma_W - 5\gamma_G - 10\gamma_B \\ s.t. \; & \gamma \geq -\gamma_W - 2 \\ & \gamma \geq -\gamma_G - 5 \\ & \gamma \leq \gamma_B + 1 \end{aligned}$$

As the dependency is now in the objective function of the second stage, the feasible region will always stay the same, regardless of the storage content χ_{B_0}. The optimization problem corresponds on evaluating the value of the objective function for all fixed vertices of the constraint set with changing storage contents.

To analyze the extreme points of the constraint set, the vertices have to be analyzed. The projection of the feasible region in the respective planes when two of the γ_P variables are zero

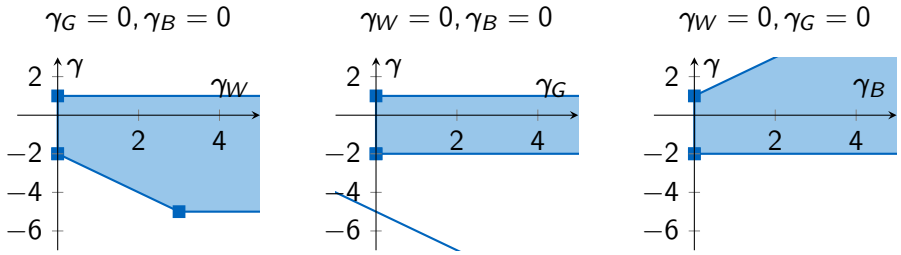


Figure 3.3: Projections of second stage dual problem of example 3.2.1 in respective two-dimensional space.

are illustrated in figure 3.3. These plots lead to three obvious choices for vertices of the dual problem:

$$P_1 = (1, 0, 0, 0)$$
$$P_2 = (-2, 0, 0, 0)$$
$$P_3 = (-5, 3, 0, 0)$$

The vertices are given in the order $P_i = (\gamma, \gamma_W, \gamma_G, \gamma_B)$. The points lead to three functions for the future cost function by plugging in the values in the objective function of the dual second stage problem:

$$f_1(\chi_{B_0}) = \chi_{B_0} - 3$$
$$f_2(\chi_{B_0}) = -2\chi_{B_0} + 6$$
$$f_3(\chi_{B_0}) = -5\chi_{B_0} + 9$$

The functions are plotted in figure 3.4. The future cost function is then given by the maximum of the three plotted functions and is therefore a step wise linear function.

The three resulting domains for χ_{B_0} have an interesting interpretation:

Domain $0 \leq \chi B_0 \leq 1$ *In this area, the use of the gas turbine is needed, so the marginal costs for producing enough power are proportional to the marginal costs of the gas turbine.*

Domain $1 \leq \chi B_0 \leq 3$ *Similar to the first domain, the costs of the wind turbine gives the slope of the cost function. As soon as at least one unit is stored in the first time step, the rest can be created by the wind turbine.*

Domain $3 \leq \chi B_0 \leq 10$ *In this domain, the cost function increases again, as everything which is more than the demand in the second time step (=3) has to be stored again which results in costs of one per unit. This is due to the fact that the demand constraint is not relaxed, but rather a strict bound.*

The general idea, on how to construct the future cost function, should now be clear. Still, evaluating the future cost problem (3.28)-(3.30) for every vertex is not computationally optimal. To develop a strategy on how to proceed with the evaluation and fining of vertices and therefore cuts for the future cost function, Pereira and Pinto [27] proposed the following method, presented in algorithm 3.2.

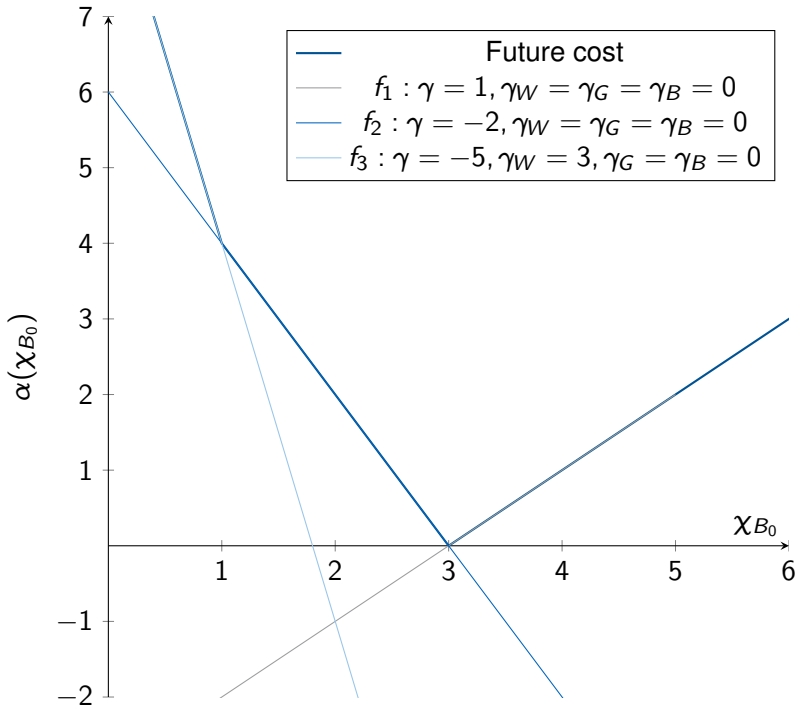


Figure 3.4: Future cost function for vertices of example 3.2.1: It results from taken the maximum over all three possible cuts (f_1, f_2, f_3).

At first, the upper bound and the variable, which represents the future costs per time step t, α_t^* is initialized. In the forward path, the decisions for every step are optimized and their result passed to the next time step. The upper bound is then calculated via the current decisions. As α_t^* has no restrictions, yet, and is a non-negative variable with a positive coefficient in the objective function, it will stay zero in all iterations of the first forward recursion.

In the following backward simulation, the algorithm starts from the last time step T and optimizes the objective function. The resulting optimal dual variables result in a cut depending on the decision of the previous step, which is then transferred in the previous step. It gives a lower bound on α_{T-1}^*. This is done for every time step, done in reverse order. The cut, which is added to the problem, results in higher values for α_t^*. It has to be noted that the dual variables of the cuts, which are generated for α_t^*, have to be taken into account for building the new cuts.

After the calculation of the backward simulation, the lower bound is updated by using the first time step decision with the corresponding future costs. As long as the difference between upper and lower bound is beyond a specified threshold, the forward simulation and, hence, the algorithm starts again.

How the presented algorithm is working is again shown via the introduced example in the following.

Example 3.2.1 (continuing from p. 43) *For the company with its gas plant, wind turbine and battery storage, the DDP algorithm would start by initializing the upper bound $\overline{z}$ with infinity, the future cost function $\alpha(\chi_{B_0})$ with zero and calculate the first stage optimization problem,*

Algorithm 3.2: DDP algorithm

Data: $\boldsymbol{A}_t$, $\boldsymbol{b}_t$, $\boldsymbol{c}_t$, $\boldsymbol{E}_t$, T

Initialization:

$\bar{z} \leftarrow \infty$

$\alpha_t^* \leftarrow 0 \; \forall t = 0, \dots T$

// Solve approximate first stage problem

$\boldsymbol{\chi}_0^* \leftarrow \arg\min_{\mathcal{X}_0} \boldsymbol{c}_0 \boldsymbol{\chi}_0 + \alpha_0^*$ with $\mathcal{X}_0 = \{\boldsymbol{A}_0 \boldsymbol{\chi}_0 \geq \boldsymbol{b}_0\}$

// Calculate lower bound

$\underline{z} \leftarrow \boldsymbol{c}_0 \boldsymbol{x}_0^*$

while $\bar{z} - \underline{z} > \epsilon$ **do**

  // Forward simulation

  for $t \leftarrow 2, \dots T$ **do**

    $\boldsymbol{\chi}_t^*, \alpha_t^* \leftarrow \arg\min_{\mathcal{X}_t} \boldsymbol{c}_t \boldsymbol{\chi}_t + \alpha_t$ with $\mathcal{X}_t = \left\{\boldsymbol{A}_t \boldsymbol{\chi}_t \geq \boldsymbol{b}_t - \boldsymbol{E}_{t-1} \boldsymbol{\chi}_{t-1}^*\right\}$

  end

  // Calculate upper bound

  $\bar{z} \leftarrow \sum_{t=1}^{T} \boldsymbol{c}_t \boldsymbol{x}_t^*$

  // Backward simulation

  for $t \leftarrow T, \dots 2$ **do**

    $\boldsymbol{\gamma}_t^* \leftarrow \arg\min_{\mathcal{X}_t} \boldsymbol{c}_t \boldsymbol{\chi}_t + \alpha_t^*$ with $\mathcal{X}_t = \left\{\boldsymbol{A}_t \boldsymbol{\chi}_t \geq \boldsymbol{b}_t - \boldsymbol{E}_{t-1} \boldsymbol{\chi}_{t-1}^* \; : \boldsymbol{\gamma}_t\right\}$

    $\alpha_{t-1}^* \geq \boldsymbol{\gamma}_{t-1}^{*,T} \boldsymbol{E}_{t-1} \boldsymbol{\chi}_{t-1} + \varsigma_{t-1}$

  end

  // Solve approximate first stage problem

  $\boldsymbol{x}_0^*, \alpha_0^* \leftarrow \arg\min_{\mathcal{X}_0} \boldsymbol{c}_0 \boldsymbol{\chi}_0 + \alpha_0$ with $\mathcal{X}_0 = \left\{\boldsymbol{A}_0 \boldsymbol{\chi}_0 \geq \boldsymbol{b}_0, \alpha_0 \geq \boldsymbol{\gamma}_1^{*,T} \left(\boldsymbol{b}_1 - \boldsymbol{E}_1 \boldsymbol{\chi}_0\right)\right\}$

  // Calculate lower bound

  $\underline{z} \leftarrow \boldsymbol{c}_0 \boldsymbol{\chi}_0^* + \alpha_0^*$

end

return $\boldsymbol{\chi}_t$, z

which is given by:

$$\begin{aligned} \min_{\chi_{W_0}, \chi_{G_0}, \chi_{B_0}} \quad & 2\chi_{W_0} + 5\chi_{G_0} + \chi_{B_0} \\ s.t. \quad & \chi_{W_0} + \chi_{G_0} - \chi_{B_0} = 1 \\ & \chi_{W_0} \leq 3 \\ & \chi_{G_0} \leq 5 \\ & \chi_{B_0} \leq 10 \\ & \chi_{P_0} \geq 0 \; \forall P \in \{W, G, B\} \end{aligned}$$

The optimal solution of this problem is given by $\boldsymbol{\chi}_{\boldsymbol{P_0}}^T = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$ *and objective value 2. The objective function* f_0 *with regard to the complicating variable* χ_{B_0} *is illustrated in figure 3.5. The objective consists of two linear functions with the slopes corresponding to the respective costs of the cheapest available plant plus the storage costs: For the domain* $[0, 2]$*, the power*

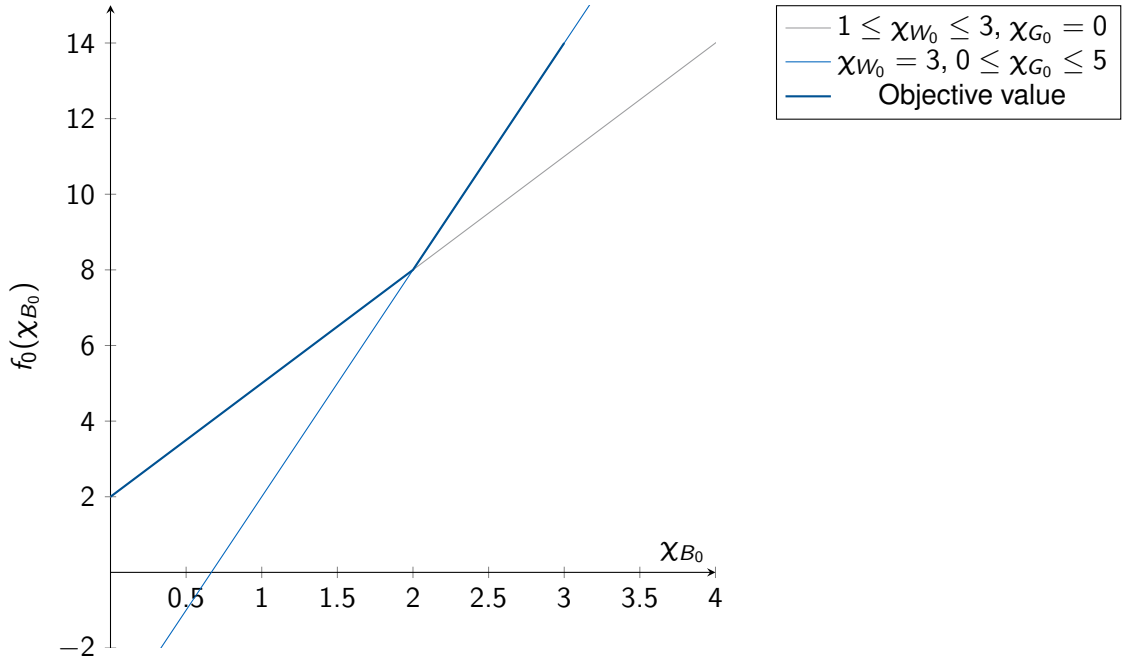


Figure 3.5: Objective function value dependent on outcome of χ_{B_0} of example 3.2.1: The objective value is the maximum of both cost driving variables: Either the wind turbine is completely used ($\chi_{W_0} = 3$) and hence, the storage has to be used ($\chi_{B_0} \geq 2$), or else the wind turbine is only partially used ($1 \leq \chi_{W_0} \leq 3$) and therefore less to nothing has to be stored at all ($0 \leq \chi_{B_0} \leq 2$). The objective value of the first stage does not take the future cost functions into account, yet.

to be stored can be created by the wind turbine, so the slope of the line is given by $2 + 1 = 3$*. As soon as more than two units are to be stored in the battery, the gas plant is needed additionally, which means that the slope results in* $5 + 1 = 6$ *for the domain* $[2, \infty[$*.*

Coming back to the algorithm, the next step is calculating the lower bound, which is given by the objective value, i.e. $\underline{z} = 2$*. As the upper bound is still to the preset value (*$\bar{z} = \infty$*), the algorithm will not stop and proceed with the forward simulation. In the given example, the forward simulation just consists of calculating the second stage problem with the result of the first stage* $\chi_{B_0} = 0$ *as constant:*

$$\begin{aligned} \alpha(\chi_{B_0} = 0) = \min_{\chi_{W_1}, \chi_{G_1}, \chi_{B_1}} \quad & 2\chi_{W_1} + 5\chi_{G_1} + \chi_{B_1} \\ \text{s.t.} \quad & \chi_{W_1} + \chi_{G_1} - \chi_{B_1} = 3 \\ & \chi_{W_1} \leq 2 \\ & \chi_{G_1} \leq 5 \\ & \chi_{B_1} \leq 10 \\ & \chi_{P_1} \geq 0 \; \forall P \in \{W, G, B\} \end{aligned}$$

Solving the primal problem of the second stage results in the following solution $\boldsymbol{\chi}_{\boldsymbol{P_1}}^T = \begin{bmatrix} 2 & 1 & 0 \end{bmatrix}$ *and objective value 9. The upper bound is updated by summing up all objectives of the forward simulation including the first time step* $\bar{z} = 2 + 9 = 11$*. Subsequently, the*

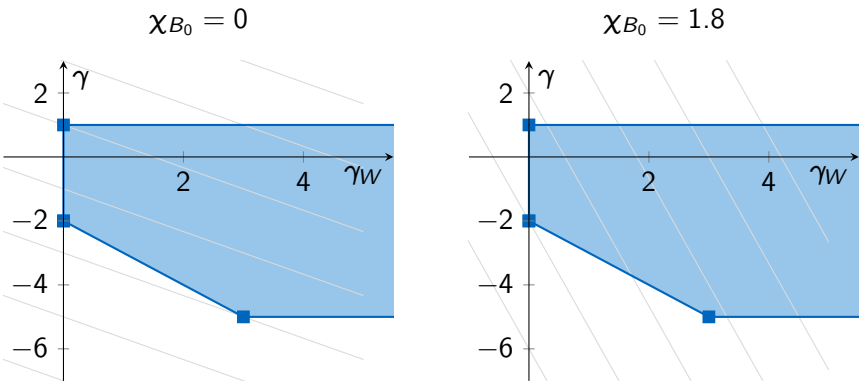


Figure 3.6: Projections of second stage dual problem of example 3.2.1 in respective two-dimensional of $\gamma_G = 0, \gamma_B = 0$ for different first stage decisions χ_{B_0}.

backward simulation starts. The dual variables of the second stage problem have to be evaluated.

The dual problems reads as follows (Side note: In algorithm 3.2, $\boldsymbol{\gamma}$ denotes the vector containing all dual variables, in this example, γ is taken for the equality constraint, the γ_Ps for the inequality constraints.):

$$\begin{aligned} \alpha(\chi_{B_0} = 0) = \max_{\gamma,\gamma_W,\gamma_G,\gamma_B} & -3\gamma - 2\gamma_W - 5\gamma_G - 10\gamma_B \\ s.t.\ & \gamma \geq -\gamma_W - 2 \\ & \gamma \geq -\gamma_G - 5 \\ & \gamma \leq \gamma_B + 1 \end{aligned}$$

The left subplot of figure 3.6 shows the dual problem in the projection space $\gamma_G = 0, \gamma_B = 0$. The gray lines represent the objective function. The optimal solution lies on the vertex $P_3 = (-5, 3, 0, 0)$. The resulting cut for the first stage problem is then given by:

$$\begin{aligned} \alpha_0 &\geq -5(\chi_{B_0} - 3) - 2 \cdot 3 \\ &\geq -5\chi_{B_0} + 9 \end{aligned}$$

As the backward iteration, starting with the last stage, creates a cut for each previous step, the problems are directly solved again with their new cuts. The backward simulation ends with solving the first stage with the newly calculated cut from the second stage (and in this case

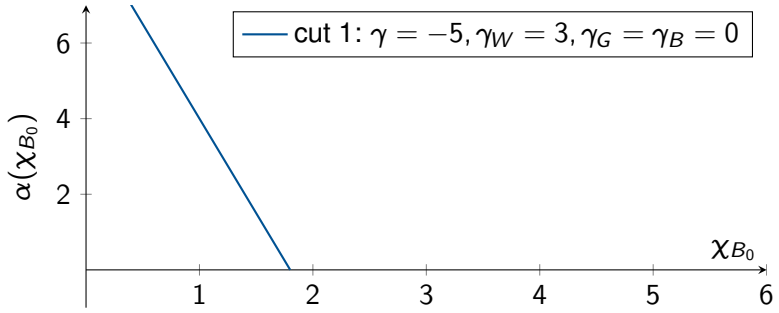


Figure 3.7: Future cost function in the first iteration after adding the first cut (example 3.2.1)

the last as well) with the new optimization variable α_0*:*

$$\begin{aligned}
\min_{\chi_{W_0},\chi_{G_0},\chi_{B_0},\alpha_0} \quad & 2\chi_{W_0} + 5\chi_{G_0} + \chi_{B_0} + \alpha_0 \\
s.t. \quad & \chi_{W_0} + \chi_{G_0} - \chi_{B_0} = 1 \\
& \chi_{W_0} \leq 3 \\
& \chi_{G_0} \leq 5 \\
& \chi_{B_0} \leq 10 \\
& \alpha_0 \geq -5\chi_{B_0} + 9 \\
& \chi_{P_0} \geq 0 \; \forall P \in \{W, G, B\} \\
& \alpha_0 \geq 0
\end{aligned}$$

The approximation of the future costs $\alpha_0(\chi_{B_0})$ *is visualized in figure 3.7. As there is just one cut added, the approximation consists only of that cut. Figure 3.8 shows the objective function of the first stage with the new cut included. The resulting objective function has its minimum at the point* $\chi_{B_0} = 1.8$*. This results in the solution* $\boldsymbol{\chi}_{\boldsymbol{P_0}}^T = \begin{bmatrix} 2.8 & 0 & 1.8 \end{bmatrix}$*,* $\alpha_0 = 0$ *and objective value 7.4. As the approximation of the future cost function gets zero for the storage value 1.8, it is the option the first stage optimum chooses.*

The algorithm proceeds by updating the lower bound to $\underline{z} = 7.4$*. While the difference between upper and lower bound is still not below the stopping criterion* $\bar{z} - \underline{z} = 11 - 7.4 = 3.6$*, the algorithm starts again with the forward iteration. The new resulting second stage with the battery storage content of* $\chi_{B_0} = 1.8$ *reads as follows:*

$$\begin{aligned}
\alpha(\chi_{B_0} = 1.8) = \min_{\chi_{W_1},\chi_{G_1},\chi_{B_1}} \quad & 2\chi_{W_1} + 5\chi_{G_1} + \chi_{B_1} \\
s.t. \quad & 1.8 + \chi_{W_1} + \chi_{G_1} - \chi_{B_1} = 3 \\
& \chi_{W_1} \leq 2 \\
& \chi_{G_1} \leq 5 \\
& \chi_{B_1} \leq 10 \\
& \chi_{P_1} \geq 0 \; \forall P \in \{W, G, B\}
\end{aligned}$$

The solution to the problem is given by $\boldsymbol{\chi}_{\boldsymbol{P_1}}^T = \begin{bmatrix} 1.2 & 0 & 0 \end{bmatrix}$ *and objective value 2.4. The upper bound is then given by* $\bar{z} = 7.4 + 2.4 = 9.8$*. The backward simulation continues.*

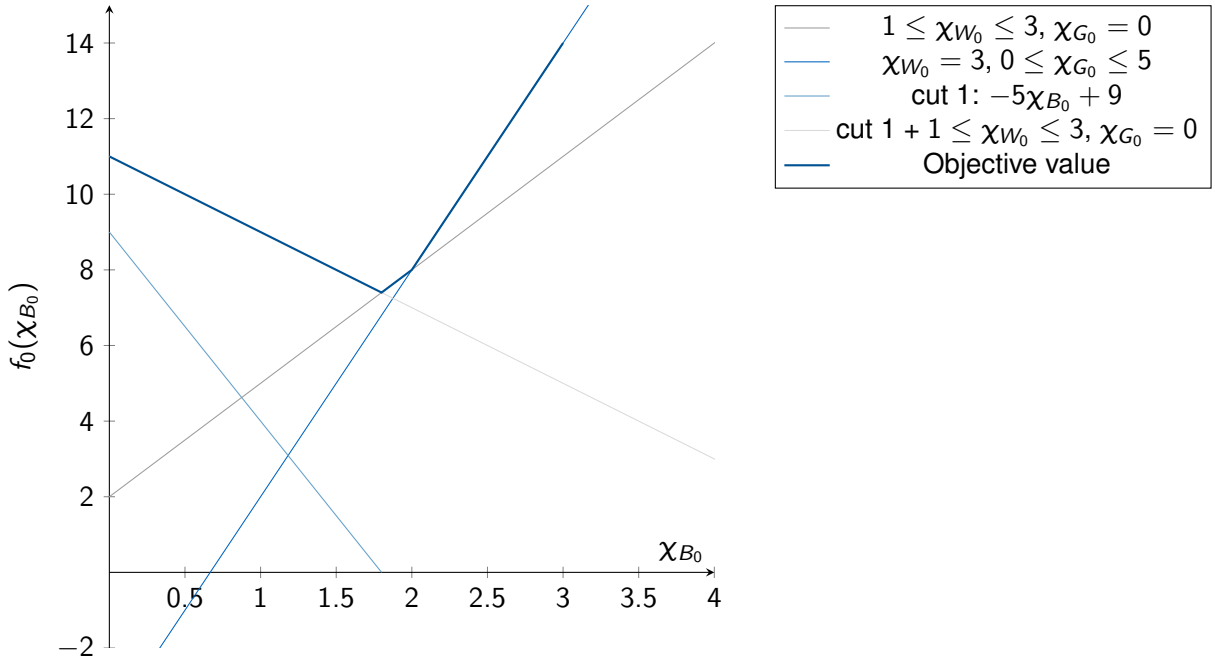


Figure 3.8: Objective function value dependent on outcome of χ_{B_0} of example 3.2.1 after adding the first cut: Next to the two original price driving curves, explained in figure 3.5, the first cut (light blue) has to be added. It ensures that a better estimate of the future costs is added to the optimization of the first stage. The cut leads to three pieces for the objective function of the first stage, as the information that saving less in the first stage leads to higher costs in the second stage (due to the necessary use of the gas turbine).

It is important to note that in case of a multi-stage optimization problem, the cuts added to the problem have to be dualized as well. So, their dual variables are a part of the cut for the preceding problem. I.e. in case of a three stage problem, the dualized problem of the second stage (with the cut coming from the third stage), includes information in their dual variables for the cut for the first stage. As in the given example, the second stage corresponds to the last stage, there are no cuts in the second stage to be dualized.

For understanding the outcome of the dual problem in stage two, it is stated again:

$$\begin{aligned} \alpha(\chi_{B_0} = 1.8) = \max_{\gamma,\gamma_W,\gamma_G,\gamma_B} \; & -1.2\gamma - 2\gamma_W - 5\gamma_G - 10\gamma_B \\ \text{s.t. } & \gamma \geq -\gamma_W - 2 \\ & \gamma \geq -\gamma_G - 5 \\ & \gamma \leq \gamma_B + 1 \end{aligned}$$

In the right subplot of figure 3.6, the cost function with respect to $\chi_{B_0} = 1.8$ can be seen. The optimum now lies on the vertex $P_2 = (-2, 0, 0, 0)$. So, the second cut results in:

$$\begin{aligned} \alpha_0 &\geq -2(\chi_{B_0} - 3) \\ &\geq -2\chi_{B_0} + 6 \end{aligned}$$

The visualization of the approximated future cost function $\alpha_0(\chi_{B_0})$ is shown in figure 3.9. The

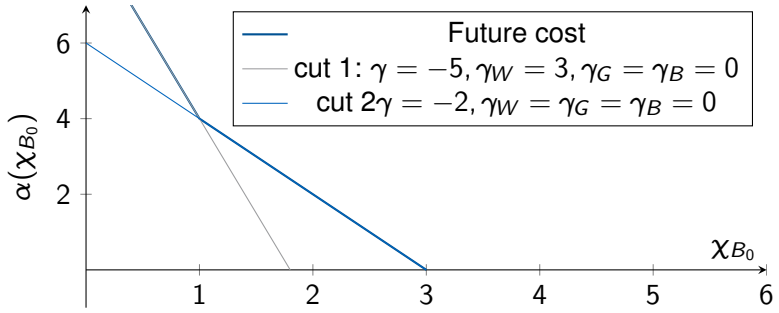


Figure 3.9: Future cost function in the second iteration after adding the second cut (example 3.2.1)

first stage problem, therefore, reads as follows:

$$\begin{aligned}
\min_{\chi_{W_0},\chi_{G_0},\chi_{B_0},\alpha_0} \quad & 2\chi_{W_0} + 5\chi_{G_0} + \chi_{B_0} + \alpha_0 \\
\textit{s.t.} \quad & \chi_{W_0} + \chi_{G_0} - \chi_{B_0} = 1 \\
& \chi_{W_0} \leq 3 \\
& \chi_{G_0} \leq 5 \\
& \chi_{B_0} \leq 10 \\
& \alpha_0 \geq -5\chi_{B_0} + 9 \\
& \alpha_0 \geq -2\chi_{B_0} + 6 \\
& \chi_{P_0} \geq 0 \; \forall P \in \{W, G, B\} \\
& \alpha_0 \geq 0
\end{aligned}$$

Due to the new cut, the minimum of the objective value shifts from $\chi_{B_0} = 1.8$ to $\chi_{B_0} = 1$. The complete solution is given by $\boldsymbol{\chi}_{\boldsymbol{P_0}}^T = \begin{bmatrix} 2 & 0 & 1 \end{bmatrix}$, $\alpha_0 = 4$ and objective value 9. The lower bound gets updated to $\underline{z} = 9$. Unfortunately, as the upper bound is still $\bar{z} = 9.8$, the forward simulation has to be started again. The second stage problem reads as follows:

$$\begin{aligned}
\alpha(\chi_{B_0} = 1) = \min_{\chi_{W_1},\chi_{G_1},\chi_{B_1}} \quad & 2\chi_{W_1} + 5\chi_{G_1} + \chi_{B_1} \\
\textit{s.t.} \quad & 1 + \chi_{W_1} + \chi_{G_1} - \chi_{B_1} = 3 \\
& \chi_{W_1} \leq 2 \\
& \chi_{G_1} \leq 5 \\
& \chi_{B_1} \leq 10 \\
& \chi_{P_1} \geq 0 \; \forall P \in \{W, G, B\}
\end{aligned}$$

As it can be seen, the wind power available in the second time step suffices with the remaining storage content to fulfill the demand. Therefore, the solution can easily be derived to be $\boldsymbol{\chi}_{\boldsymbol{P_1}}^T = \begin{bmatrix} 2 & 0 & 0 \end{bmatrix}$ and objective value 4. The upper bound updates to $\bar{z} = 5 + 4 = 9$. The

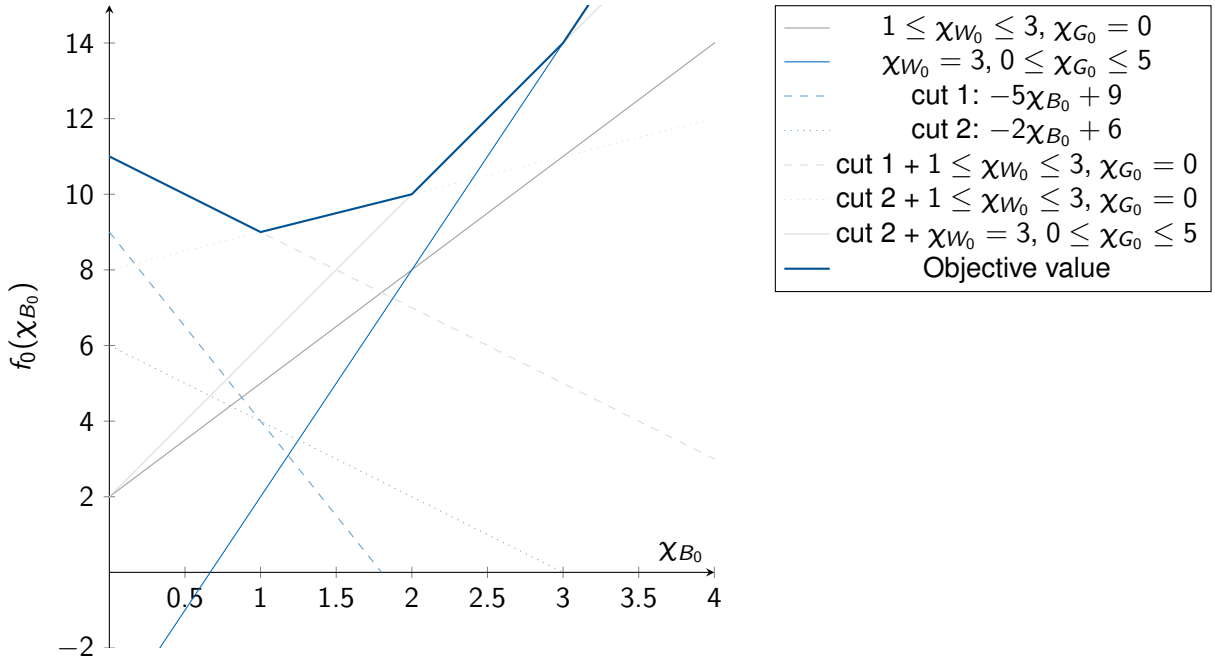


Figure 3.10: Objective function value dependent on outcome of χ_{B_0} of example 3.2.1 after adding the second cut: The new cut (dotted light blue) again provides more information about the additional costs by saving/not saving power in the storage for the second stage. The four parts of the piecewise linear objective function are derived as follows: Part I is gained by adding the first cut to the costs from the first stage problem (c.f. figure 3.5), for part II the second cut is now relevant (c.f. figure 3.9). Part III results from using all possible wind power ($\chi_{W_0} = 3$) added to the second cut.

algorithm continues with the backward simulation using the dual variables of the second stage:

$$\begin{aligned} \alpha(\chi_{B_0} = 1) = \max_{\gamma,\gamma_W,\gamma_G,\gamma_B} \quad & -2\gamma - 2\gamma_W - 5\gamma_G - 10\gamma_B \\ \text{s.t. } & \gamma \geq -\gamma_W - 2 \\ & \gamma \geq -\gamma_G - 5 \\ & \gamma \leq \gamma_B + 1 \end{aligned}$$

As the cost vector is now perpendicular to the constraint connecting the previous calculated vertices, the solution of the second stage optimization problem in dual space does not have a unique solution. All points on the line correspond to solutions. As both extreme points of this constraint have already been used to construct cuts for the future cost function, there is no new information to be added to the approximation.

Hence, the objective value of the first stage problem stays the same as well as the lower bound. Fortunately, the lower and upper bound are equal ($\bar{z} - \underline{z} = 9 - 9 = 0$) and the algorithm terminates with the assumed solution given in table 3.2.

For the extension to the multistage case, the cut has to be adapted as note by Velásquez, Restrepo, and Campo [108]. It has already been mentioned that the dual variable of the cut

has to be taken into account for the cut generation. This can be easily shown by writing down the problems of a multistage case:

$$\min_{\boldsymbol{\chi}_0,\ldots,\boldsymbol{\chi}_t,\ldots\boldsymbol{\chi}_T} \sum_{t=1}^{T} \boldsymbol{c}_t^{\mathsf{T}}\boldsymbol{\chi}_t \tag{3.31}$$

$$\text{s.t. } \boldsymbol{A}_0\boldsymbol{\chi}_0 \geq \boldsymbol{b}_0 \tag{3.32}$$

$$\boldsymbol{E}_{t-1}\boldsymbol{\chi}_{t-1} + \boldsymbol{A}_t\boldsymbol{\chi}_t \geq \boldsymbol{b}_t \; \forall t \in \{1,\ldots,T\} \tag{3.33}$$

$$\boldsymbol{\chi}_t \geq \boldsymbol{0} \; \forall t \in \{0,\ldots,T\} \tag{3.34}$$

Hence, stage $T-1$ of problem (3.31)-(3.34) after one iteration looks as follows:

$$\min_{\boldsymbol{\chi}_{T-1}} \boldsymbol{c}_{T-1}^{\mathsf{T}}\boldsymbol{\chi}_{T-1} + \alpha_{T-1}(\boldsymbol{\chi}_{T-1}) \tag{3.35}$$

$$\text{s.t. } \boldsymbol{A}_{T-1}\boldsymbol{\chi}_t \geq \boldsymbol{b}_{T-1} - \boldsymbol{E}_{T-2}\boldsymbol{\chi}_{T-2} \tag{3.36}$$

$$\alpha_{T-1} \geq \boldsymbol{\gamma}_T^{\mathsf{T}}\left(\boldsymbol{b}_{T-1} - \boldsymbol{E}_{T-1}\boldsymbol{\chi}_{T-1}\right) \tag{3.37}$$

$$\boldsymbol{\chi}_{T-1} \geq \boldsymbol{0} \tag{3.38}$$

To get the cut for stage $T-2$, the dual problem is taken. Hence, also cut (3.37) has to be dualized. In general, the cut for stage t will look like the following:

$$\alpha_t \geq -\boldsymbol{\gamma}_{t+1}^{\mathsf{T}}\boldsymbol{E}_t\boldsymbol{\chi}_t + \varsigma_t. \tag{3.39}$$

With ς_t taken from [108] given by:

$$\varsigma_t = \begin{cases} \boldsymbol{\gamma}_{t+1}^{\mathsf{T}}\boldsymbol{b}_{t+1} & t = T-1 \\ \boldsymbol{\gamma}_{t+1}^{\mathsf{T}}\boldsymbol{b}_{t+1} + \gamma_{t+1,\alpha}\varsigma_{t+1} & t \in \{0,\ldots,T-2\} \end{cases}. \tag{3.40}$$

As the general idea behind DDP is now clear, one can proceed to introduce the stochastic variant of the algorithm.

Extension to stochastic case For the stochastic case, there are more optimization problems in the second and following stages. The deterministic two-stage problem (3.18)-(3.21) can be extended as follows, if the second-stage becomes uncertain:

$$\min_{\boldsymbol{\chi}_0,\boldsymbol{\chi}_{1r}} \boldsymbol{c}_0^{\mathsf{T}}\boldsymbol{\chi}_0 + \sum_{r=1}^{R} p_{1r}\boldsymbol{c}_1^{\mathsf{T}}\boldsymbol{\chi}_{1r} \tag{3.41}$$

$$\text{s.t. } \boldsymbol{A}_0\boldsymbol{\chi}_0 \geq \boldsymbol{b}_0 \tag{3.42}$$

$$\boldsymbol{E}_0\boldsymbol{\chi}_0 + \boldsymbol{A}_{1r}\boldsymbol{\chi}_{1r} \geq \boldsymbol{b}_{1r} \; \forall r \in \{1,\ldots,R\} \tag{3.43}$$

$$\boldsymbol{\chi}_0, \boldsymbol{\chi}_{1r} \geq \boldsymbol{0} \; \forall r \in \{1,\ldots,R\} \tag{3.44}$$

The parameter p_{1r} denotes the probability of the r-th realization of the second stage. There are in total R realizations, which the second stage can take. The subscript r for the matrix $\boldsymbol{A}_{1r}$ and the vectors $\boldsymbol{b}_{1r}$, $\boldsymbol{\chi}_{1r}$ represents the dependence of the parameters and therefore the variables on the realization. This notation corresponds to problem (3.13)-(3.17) for a discrete

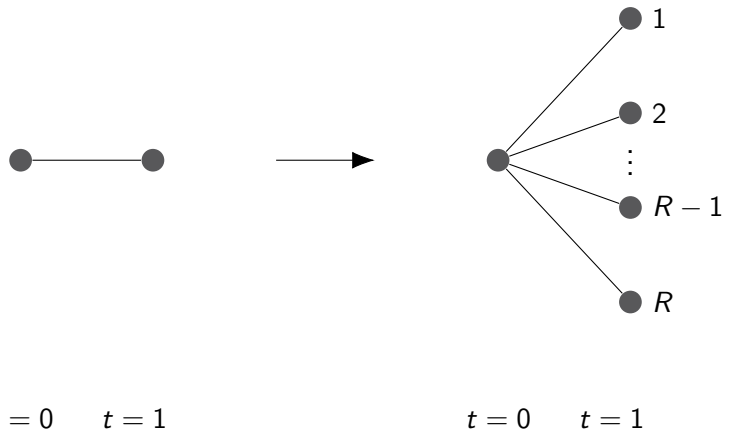


Figure 3.11: Realization representation for two-stage DDP and SDDP

random variable, where the realization was denoted by ω_t. The superscript notation is used further on for clarity and brevity.

The uncertain parameters can be in the matrix $\boldsymbol{A}_{1r}$ as well as in the right-hand side $\boldsymbol{b}_{1r}$ of the constraint (3.43). In conclusion, the number of variables for one uncertain stage is growing by the numbers of realizations R. In the multi-stage case, the growth is hence exponential, as for every realization of the previous step, all realizations of the current step get their own variable.

Figure 3.11 represents the extension from the deterministic DDP case to the uncertain SDDP case for two-stage problems. Each node in time step $t = 1$ represents one of the R realizations the uncertain parameters can become.

One simplification has been included without being mentioned yet: By writing the stochastic two-stage problem as it has been done in (3.41)-(3.44), a discrete probability distribution is used. While it is not necessarily the case that the uncertain parameters are discretely distributed, a discretization will however be required if the uncertain optimization problem is to be formulated as a tractable one.

In the case of a continuous distribution function instead of the summation an integral would be used. To get from the continuous to the discrete case, samples have to be taken to represent the original structure of the distribution function. The resulting problem is the so-called Sample Average Approximation (SAA) based on Monte Carlo. For the construction of this problem, the interested reader is referred to [101, Chapter 9.1], [109] and [102, Chapter 5].

The algorithm for the stochastic case depends on what assumptions can be made for the underlying stochastic process. If, for example, independence from one time step to the next can be assumed, the algorithm is easier due to the fact that cuts for the approximation of the future cost function can be used multiple times.

In the following, the derivation of the cuts is done for the general case. Afterwards the implications for the independent case and the assumption of a Markov process is taken into account for a multi stage problem.

SDDP procedure The main idea of SDDP is to approximate the expected value of the future costs, instead of calculating the costs for every realization. As a reminder, the expected future

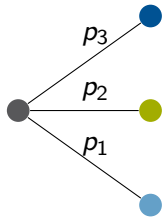


Figure 3.12: Representation of two-stage stochastic problem with $R = 3$ realizations

costs are given by:

$$\min_{\boldsymbol{\chi}_{1r}} \sum_{r=1}^{R} p_{1r} \boldsymbol{c}_1^{\mathsf{T}} \boldsymbol{\chi}_{1r} \tag{3.45}$$

$$\text{s.t. } \boldsymbol{A}_{1r} \boldsymbol{\chi}_{1r} \geq \boldsymbol{b}_{1r} - \boldsymbol{E}_0 \boldsymbol{\chi}_0 \tag{3.46}$$

$$\boldsymbol{\chi}_{1r} \geq \boldsymbol{0} \tag{3.47}$$

This corresponds to the second stage optimization problem of the two-stage case. In general, the procedure of the SDDP algorithm is quite similar to the DDP case. The main difference lies within the amount of realizations to build one cut out. Instead of creating one cut per scenario, all possible outcomes in one time step are used to create a weighted cut. The weighting is done via the probabilities of each realization.

Figure 3.12 shows an example of a two-stage stochastic optimization problem with $R = 3$ different realizations for the second time step. The probabilities to reach each of the realizations are given by p_1, p_2 and p_3. For each of the realizations, the second stage problem for the r-th realization looks like follows:

$$\min_{\boldsymbol{\chi}_{1r}} \boldsymbol{c}_1^{\mathsf{T}} \boldsymbol{\chi}_{1r} \tag{3.48}$$

$$\text{s.t. } \boldsymbol{A}_{1r} \boldsymbol{\chi}_{1r} \geq \boldsymbol{b}_{1r} - \boldsymbol{E}_0 \boldsymbol{\chi}_0 \tag{3.49}$$

$$\boldsymbol{\chi}_{1r} \geq \boldsymbol{0} \tag{3.50}$$

Problem (3.48)-(3.50) is basically just one of the sum terms of the future cost function (3.45)-(3.47). All the R problems of the second stage can be solved like the problems in the DDP case. The resulting cuts would then be summed up and weighted by the probabilities. For R realizations, the cut for the first stage problem would then be given by:

$$\alpha_0 \geq \sum_{r=1}^{R} p_{1r} \boldsymbol{\gamma}_{1r}^{*,T} \left(\boldsymbol{b}_{1r} - \boldsymbol{E}_0 \boldsymbol{\chi}_0 \right) \tag{3.51}$$

The generation of the cuts is one main difference to the classic stochastic approach. Instead of calculating the scenario tree as a whole, the cuts, created by the later time steps are added to the former ones, combine the information one gains by solving the later problems. This is represented in figure 3.13. The left graph gives the classic stochastic structure with three possible realizations per time step. Each of the realizations in time step $t = 2$ is treated

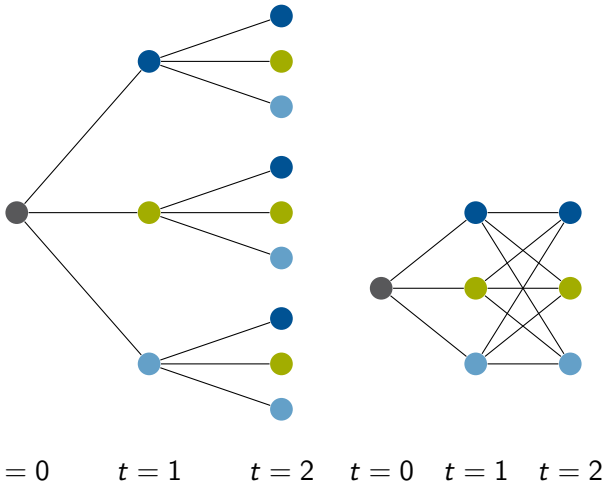


Figure 3.13: Stochastic case with classic scenario tree and SDDP representation of combining information of the tree

separately. In the right graph, the information of each realization of time step $t = 2$ is used for cut generation of the realizations in time step $t = 1$.

One has to note that the benefits coming from SDDP compared to the classic stochastic approach mostly arise from the fact that not the whole history of a scenario is necessary to represent it. This comes to the price of having a more complicated solution algorithm than solving the deterministic equivalent used for stochastic optimization. The benefits are especially visible in a four-stage scenario tree, illustrated in figure 3.14. The more history is needed to represent the dark blue scenarios or their probability in the last time step, the less can their similarities be exploited. Some benefits in creating information about the future costs from the last time step can always be used for different scenarios of the previous time step, but still computational effort has to be included due to different probabilities for equal realizations. This will be explained more detailed for the independent and Markov stochastic process.

Similar to the DDP algorithm 3.2, the SDDP procedure is structured in the forward and the backward simulation. In the forward simulation, one path in the SDDP tree is selected, choosing one realization in each time step randomly, represented in with blue edges in the SDDP lattice in figure 3.15. For the chosen path (root-mid-high), each time step will be calculated separately with fixed solutions from the time step before (same as in the forward simulation of the DDP algorithm 3.2). In the backward step, all realizations of time step $t = 2$ are taken and their dual variables evaluated to build the weighted cut (3.51).

In general, the cut can only be used for the previous scenario on the path. Depending on the assumptions of the underlying stochastic process, the cut needs to be weighted differently for using it for other scenarios. This is where the advantages of SDDP arise.

To illustrate the idea of the forward and backward step in the SDDP case, the following example is taken:

Example 3.2.2 *The wind output of a turbine is uncertain and has three different realizations per time step: a low, mid and high wind scenario. For the explanation of the forward and*

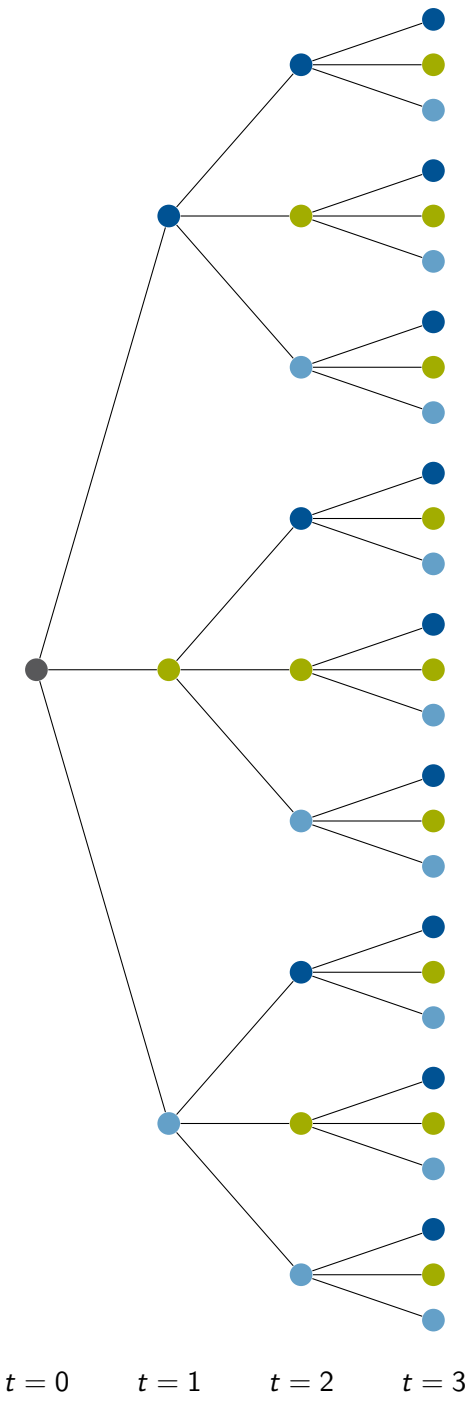


Figure 3.14: Illustration of the benefits and drawbacks of SDDP via a four-stage scenario tree: The less the history of the scenarios in the last stage $t = 3$ is important, the more the tree can be nested. If the stages are completely independent of their predecessor, all scenarios of the last stage boil down to three.

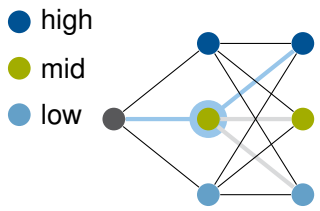


Figure 3.15: Selection of one path in the forward iteration and calculating one cut for the *mid* scenario in the backward iteration

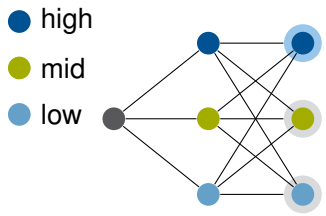


Figure 3.16: Probabilities for the scenarios of example 3.2.2

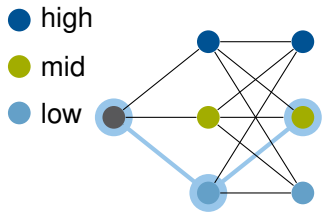


Figure 3.17: Randomly picked path (root-low-mid) in forward simulation of example 3.2.2

backward step, three time steps are taken into account. The respective nested scenario tree is represented in figure 3.16.

The algorithm starts similar to the DDP algorithm 3.2 with an initialization, mostly by setting the expected future costs to zero and solving the certain and first time step. In the forward simulation of the SDDP algorithm, one random path of the tree is chosen and optimized step by step like in the DDP case.

In case of the three time step example, in the first iteration, (root-low-mid) is the chosen path, represented in figure 3.17. The result of the root-node, which is equal to the certain problem in the first time step ($t = 0$), is used as an input for the scenario low in time step $t = 1$. The output of this optimization problem is then taken as input for the optimization problem in time step $t = 2$ for the mid scenario. This is similar to the process of the DDP algorithm.

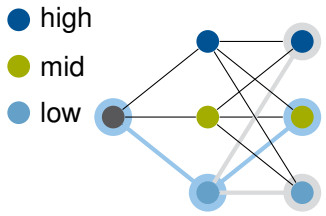


Figure 3.18: Calculating cut for time step $t = 1$ out of all realizations in time step $t = 2$ (example 3.2.2)

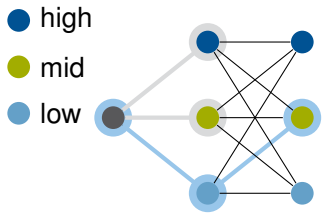


Figure 3.19: Calculating cut for time step $t = 0$ out of all realizations in time step $t = 1$ (example 3.2.2)

The probabilities of the respective realizations are not used in calculating the outcomes of the respective path in the forward algorithm.

The backward recursion is again responsible for creating the cut. For the cut generation, the probabilities are needed as seen in equation (3.51)*. Using the result of the low scenario in time step* $t = 1$*, the remaining two scenarios high and low of time step* $t = 3$ *are optimized. The cut is then generated via the separate cut parts of each of the three realizations in time step* $t = 2$ *weighted by the probabilities like in equation* (3.51)*. The weighted cut is then added to the mid scenario in time step* $t = 1$*. The backward recursion is still not completed, as the same procedure is done for the cut generated by all possible realizations in time step* $t = 1$ *and added to the root-node in time step* $t = 0$*, visualized by figure 3.19.*

In the last step of the backward recursion, the root problem is again optimized including the new weighted cut.

Dependent on the assumptions, which hold for the stochastic process (e.g. independence or Markov), the cuts can not only be added to the respective realizations of the path, but to all realizations of one time step. This is where the benefit of SDDP compared to the classic stochastic approach comes into play. The two main assumptions, which are used in context of SDDP, namely independent process and Markov process, and their impact on the structure of the algorithm are discussed in a separate paragraph.

The next step in deriving the SDDP algorithm lies within the concrete cut generation. Therefore, a numerical example is elaborated in the following.

Plant/Storage	Costs per unit	Capacity
Wind	2	variable
Gas	5	5
Battery	1	10

Time step	Demand	Capacity Wind
1	1	3 (certain)
2	3	low: 1, mid: 2, high: 3

Table 3.3: Data for SDDP example 3.2.3

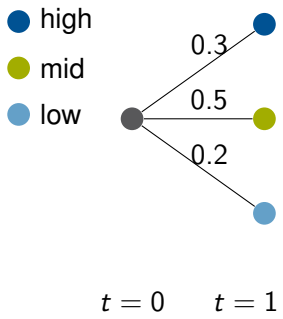


Figure 3.20: Probabilities for the scenarios of example 3.2.3

Example 3.2.3 *The power system of the company already known from example 3.2.1 of the DDP algorithm is now faced with an uncertain maximum available capacity for the wind turbine in time step $t = 1$. The input data of the example is given in table 3.3 and the probabilities of the respective scenarios are illustrated in figure 3.20. The stochastic optimization problem reads as follows:*

$$
\begin{aligned}
\min_{\boldsymbol{\chi}_W, \boldsymbol{\chi}_G, \boldsymbol{\chi}_B} \; & 2\chi_{W_0} + 5\chi_{G_0} + \chi_{B_0} + 0.2 \cdot \left(2\chi_{W_{1,low}} + 5\chi_{G_{1,low}} + \chi_{B_{1,low}}\right) \\
& + 0.5 \cdot \left(2\chi_{W_{1,mid}} + 5\chi_{G_{1,mid}} + \chi_{B_{1,mid}}\right) + 0.3 \cdot \left(2\chi_{W_{1,high}} + 5\chi_{G_{1,high}} + \chi_{B_{1,high}}\right) \\
s.t. \; & \chi_{W_0} + \chi_{G_0} - \chi_{B_0} = 1 \\
& \chi_{B_0} + \chi_{W_{1,low}} + \chi_{G_{1,low}} - \chi_{B_{1,low}} = 3 \\
& \chi_{B_0} + \chi_{W_{1,mid}} + \chi_{G_{1,mid}} - \chi_{B_{1,mid}} = 3 \\
& \chi_{B_0} + \chi_{W_{1,high}} + \chi_{G_{1,high}} - \chi_{B_{1,high}} = 3 \\
& \chi_{W_0} \leq 3 \\
& \chi_{W_{1,low}} \leq 1 \\
& \chi_{W_{1,mid}} \leq 2 \\
& \chi_{W_{1,high}} \leq 3 \\
& \boldsymbol{\chi}_G \leq \mathbf{5} \\
& \boldsymbol{\chi}_B \leq \mathbf{10} \\
& \boldsymbol{\chi}_P \geq \mathbf{0} \; \forall P \in \{W, G, B\}
\end{aligned}
$$

In contrast to the DDP example 3.2.1, the solution is not immediately visible. Instead of three variables for the second time step, nine variables are necessary to represent the stochastic problem in a deterministic form.

As the right-hand-side of the optimization problem is uncertain, the properties fixed recourse, fixed technology and fixed objective coefficients hold (c.f. section 3.2.2.1).

The optimization problem can be decomposed into four separate optimization problems, one for the first time step $t = 0$ and three for the second. Before starting with the forward iteration, initialization and optimization of the first stage take place:

$$
\begin{aligned}
\min_{\chi_{W_0}, \chi_{G_0}, \chi_{B_0}} \; & 2\chi_{W_0} + 5\chi_{G_0} + \chi_{B_0} \\
s.t. \; & \chi_{W_0} + \chi_{G_0} - \chi_{B_0} = 1 \\
& \chi_{W_0} \leq 3 \\
& \chi_{G_0} \leq \mathbf{5} \\
& \chi_{B_0} \leq \mathbf{10} \\
& \boldsymbol{\chi}_{P_0} \geq \mathbf{0} \; \forall P \in \{W, G, B\}
\end{aligned}
$$

It should be noted that the expected future costs are initialized to zero, similar to the DDP case. The result of the first stage optimization problem can now be taken from the deterministic version of the example, so the optimal solution is given by $\boldsymbol{\chi}_{P_0}^T = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$ and objective value 2. The lower bound $\underline{z}$ is set to the value of the first stage optimum $\underline{z} = 2$.

For the stochastic case, the forward simulation starts by picking a random path, in this case picking a random realization for the second stage $t = 1$. It may be assumed that the realization mid gets chosen and therefore, the second stage optimization problem for scenario

mid is calculated:

$$\begin{aligned}\alpha_{mid}(\chi_{B_0}) = \min_{\chi_{W_{1,mid}},\chi_{G_{1,mid}},\chi_{B_{1,mid}}} \; & 2\chi_{W_{1,mid}} + 5\chi_{G_{1,mid}} + \chi_{B_{1,mid}} \\ s.t. \; & \chi_{B_0} + \chi_{W_{1,mid}} + \chi_{G_{1,mid}} - \chi_{B_{1,mid}} = 3 \\ & \chi_{W_{1,mid}} \leq 2 \\ & \chi_{G_{1,mid}} \leq 5 \\ & \chi_{B_{1,mid}} \leq 10 \\ & \chi_{P_{1,mid}} \geq 0 \; \forall P \in \{W, G, B\}\end{aligned}$$

The corresponding dual problem reads as follows:

$$\begin{aligned}\alpha_{mid}(\chi_{B_0}) = \max_{\gamma_{mid},\gamma_{W_{mid}},\gamma_{G_{mid}},\gamma_{B_{mid}}} \; & \gamma(\chi_{B_0} - 3) - 2\gamma_{W_{mid}} - 5\gamma_{G_{mid}} - 10\gamma_{B_{mid}} \\ s.t. \; & \gamma_{mid} \geq -\gamma_{W_{mid}} - 2 \\ & \gamma \geq -\gamma_{G_{mid}} - 5 \\ & \gamma \leq \gamma_{B_{mid}} + 1\end{aligned}$$

By setting $\chi_{B_0} = 0$ from optimizing the first stage, the result is given by the same results from before: $\boldsymbol{\chi}_{\boldsymbol{P_1}}^T = \begin{bmatrix} 2 & 1 & 0 \end{bmatrix}$ and objective value 9. In contrast to the DDP case, the upper bound is not set, yet. In fact, explaining the creation of the upper bound is done in the respective paragraph.

To get the cut, all scenarios in the second time step $t = 1$ have to be calculated. The cut-part from the mid realization can already be stated by:

$$\begin{aligned}\alpha_{0,mid} &\geq -5(\chi_{B_0} - 3) - 2 \cdot 3 \\ &\geq -5\chi_{B_0} + 9\end{aligned}$$

The two remaining second stage problems read as follows:

$$\begin{aligned}\alpha_{low}(\chi_{B_0}) = \min_{\chi_{W_{1,low}},\chi_{G_{1,low}},\chi_{B_{1,low}}} \; & 2\chi_{W_{1,low}} + 5\chi_{G_{1,low}} + \chi_{B_{1,low}} \\ s.t. \; & \chi_{B_0} + \chi_{W_{1,low}} + \chi_{G_{1,low}} - \chi_{B_{1,low}} = 3 \\ & \chi_{W_{1,low}} \leq 1 \\ & \chi_{G_{1,low}} \leq 5 \\ & \chi_{B_{1,low}} \leq 10 \\ & \chi_{P_{1,low}} \geq 0 \; \forall P \in \{W, G, B\}\end{aligned}$$

$$\begin{aligned}\alpha_{high}(\chi_{B_0}) = \min_{\chi_{W_{1,high}},\chi_{G_{1,high}},\chi_{B_{1,high}}} \; & 2\chi_{W_{1,high}} + 5\chi_{G_{1,high}} + \chi_{B_{1,high}} \\ s.t. \; & \chi_{B_0} + \chi_{W_{1,high}} + \chi_{G_{1,high}} - \chi_{B_{1,high}} = 3 \\ & \chi_{W_{1,high}} \leq 3 \\ & \chi_{G_{1,high}} \leq 5 \\ & \chi_{B_{1,high}} \leq 10 \\ & \chi_{P_{1,high}} \geq 0 \; \forall P \in \{W, G, B\}\end{aligned}$$

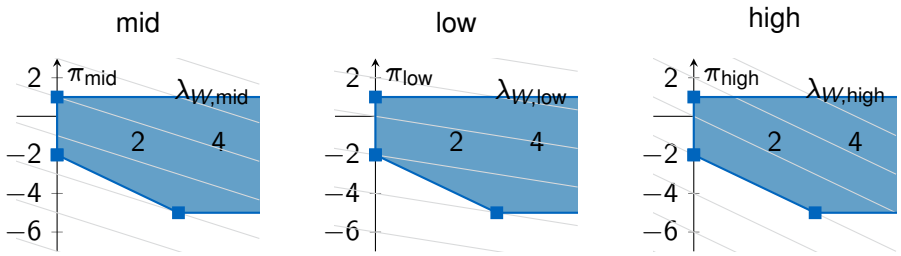


Figure 3.21: Projections of second stage dual problem of example 3.2.3 for all realizations (mid, low, high) in respective two-dimensional space with $\gamma_{G,r} = 0, \gamma_{B,r} = 0, \ \forall r \in \{\text{low, mid, high}\}$ for first stage decision $\chi_{B_0} = 0$.

In the primal space, both objective functions stay the same. Transforming to the dual space means again that the uncertain parameters switch to the objective function, while the set of feasible solutions stays the same:

$$\begin{aligned}\alpha_{low}(\chi_{B_0} = 0) = \max_{\gamma_{low},\gamma_{W_{low}},\gamma_{G_{low}},\gamma_{B_{low}}} \quad & -3\gamma_{low} - \gamma_{W_{low}} - 5\gamma_{G_{low}} - 10\gamma_{B_{low}} \\ \text{s.t. } \gamma &\geq -\gamma_{W_{low}} - 2 \\ \gamma &\geq -\gamma_{G_{low}} - 5 \\ \gamma &\leq \gamma_{B_{low}} + 1\end{aligned}$$

$$\begin{aligned}\alpha_{high}(\chi_{B_0}) = \max_{\gamma_{high},\gamma_{W_{high}},\gamma_{G_{high}},\gamma_{B_{high}}} \quad & -3\gamma_{high} - 3\gamma_{W_{high}} - 5\gamma_{G_{high}} - 10\gamma_{B_{high}} \\ \text{s.t. } \gamma &\geq -\gamma_{W_{high}} - 2 \\ \gamma &\geq -\gamma_{G_{high}} - 5 \\ \gamma &\leq \gamma_{B_{high}} + 1\end{aligned}$$

Figure 3.21 visualizes all variants of the objective function for the second stage dependent on the realization. As the dual problem of the high realization yields not a unique solution, which means the solution to the primal problem is degenerated, i.e. more constraints than "necessary" are active in the optimum.

Dependent on the implementation of the linear solver algorithm, either extreme point of the constraint can be picked, i.e. $P_2 = (-2, 0, 0, 0)$ or $P_3 = (-5, 3, 0, 0)$ for the high scenario. Therefore, two possible cut-parts arise from the derivation:

$$\begin{aligned}\alpha_{0,high}(P_2) &\geq -2(\chi_{B_0} - 3) \\ &\geq -2\chi_{B_0} + 6 \\ \alpha_{0,high}(P_3) &\geq -5(\chi_{B_0} - 3) - 3 \cdot 3 \\ &\geq -5\chi_{B_0} + 6\end{aligned}$$

In case of a low wind scenario, the resulting cut is given by:

$$\begin{aligned}\alpha_{0,low} &\geq -5(\chi_{B_0} - 3) - 3 \\ &\geq -5\chi_{B_0} + 12\end{aligned}$$

Plant/Storage	$t = 0$	$t = 1$ low	$t = 1$ mid	$t = 1$ high
Wind	2	1	2	2
Gas	0	1	0	0
Battery	1	0	0	0

Table 3.4: Solution for SDDP example 3.2.3

Hence, the two possibilities for the complete weighted cut added to the first stage are given by:

$$\begin{aligned}\alpha_0(P_2) &\geq 0.2(-5\chi_{B_0} + 12) + 0.5(-5\chi_{B_0} + 9) - 0.3(-2\chi_{B_0} + 6)\\ &\geq -4.1\chi_{B_0} + 8.7\\ \alpha_0(P_3) &\geq 0.2(-5\chi_{B_0} + 12) + 0.5(-5\chi_{B_0} + 9) - 0.3(-5\chi_{B_0} + 6)\\ &\geq -5\chi_{B_0} + 8.7\end{aligned}$$

Compared to the DDP example 3.2.1, the resulting cut for P_3 leads to a slightly cheaper future cost function (8.7 instead of 9). This is due to the fact that the high wind scenario is slightly more probable than the low wind case. Dependent on the choice of the linear solver, meaning dependent on which extreme point was chosen, the respective cut is added to the first-stage problem and the algorithm starts again. In the end, the optimal dispatch is stated in table 3.4 with resulting costs of 9.6.

In 70 % of the cases, the capacity of the wind turbine in the second time step is not sufficient to fulfill the demand, meaning that the optimal strategy of the company is to store one unit in the battery at the first time step. In case of low wind, the gas plant is still needed, but in both more likely scenarios mid and high, the wind turbine combined with the stored unit is sufficient for the increased demand of 3.

The calculated costs of 9.6 must be seen as expected costs and not the costs which have to be paid. The costs for the company in the end are either 9 (mid, high) or 12 (low).

The steps of the SDDP procedure so far are collected in algorithm 3.3. As the specific calculation of the upper bound was not mentioned so far, it is left out of the algorithm.

Independent and Markov processes As noted before, the advantages of the SDDP algorithm lies within combining information of future scenarios. The underlying stochastic process of the uncertainty defines the degree of how much information can be used. The two most important assumptions therefore, are either independence of the stochastic process or a Markov process.

In literature for SDDP, the independence assumption is most often found. It is the basis for most arguments on why the algorithm gives a computational advantage over the classic stochastic approach. In case of an independent stochastic process, the history of how a scenario is reached, is not important at all. Figure 3.22 illustrates that concept: All probabilities to reach a scenario in time step $t = 2$ are the same for each scenario, independent of which scenario has been attained in time step $t = 1$.

As all probabilities for each scenario in one time step regardless of their history are the same, the cuts which are generated out of the scenarios are not only applicable to the

Algorithm 3.3: SDDP algorithm without upper bound calculation

Data: $\boldsymbol{A}_{tr}$, $\boldsymbol{b}_{tr}$, $\boldsymbol{c}_{tr}$, $\boldsymbol{E}_{tr}$, p_{tr}, T

Initialization:

$\bar{z} \leftarrow \infty$

$\alpha^*_{tr} \leftarrow 0 \; \forall t = 0, \dots T$

`// Solve approximate first stage problem`

$\boldsymbol{x}^*_0 \leftarrow \arg\min_{\mathcal{X}_0} \boldsymbol{c}_0 \boldsymbol{\chi}_0 + \alpha^*_0$ with $\mathcal{X}_0 = \{\boldsymbol{A}_0 \boldsymbol{\chi}_0 \geq \boldsymbol{b}_0\}$

`// Calculate lower bound`

$\underline{z} \leftarrow \boldsymbol{c}_0 \boldsymbol{\chi}^*_0$

while $\bar{z} - \underline{z} > \epsilon$ **do**

`// Forward simulation`

for $t \leftarrow 2, \dots T$ **do**

`// Pick random realization` r

$\boldsymbol{\chi}^*_{tr}, \alpha^*_{tr} \leftarrow \arg\min_{\mathcal{X}_{tr}} \boldsymbol{c}_t \boldsymbol{\chi}_{tr} + \alpha_{tr}$ with

$$\mathcal{X}_{tr} = \left\{ \boldsymbol{A}_{tr} \boldsymbol{\chi}_{tr} \geq \boldsymbol{b}_{tr} - \boldsymbol{E}_{t-1,r_{t-1}} \boldsymbol{\chi}^*_{t-1,r_{t-1}} \right\}$$

`// Save picked realization` r `of respective time step` t `in` r_t

end

`// Calculation of upper bound`

`// Will be discussed in a separate paragraph.`

`// Backward simulation`

for $t \leftarrow T, \dots 2$ **do**

`// Iterate through all realizations of that time step dependent on the chosen path`

for $r \in R_t$ **do**

$\boldsymbol{\gamma}^*_{tr} \leftarrow \arg\min_{\mathcal{X}_{tr}} \boldsymbol{c}_t \boldsymbol{\chi}_{tr} + \alpha^*_{tr}$ with

$$\mathcal{X}_{tr} = \left\{ \boldsymbol{A}_{tr} \boldsymbol{\chi}_{tr} \geq \boldsymbol{b}_{tr} - \boldsymbol{E}_{t-1,r_{t-1}} \boldsymbol{\chi}^*_{t-1,r_{t-1}} \; : \boldsymbol{\gamma}_{tr} \right\}$$

`// with` r_{t-1} `as saved realization from forward simulation`

end

`// Calculate weighted cut and deploy only to realization of previous time step according to path`

$$\alpha^*_{t-1,r_{t-1}} \geq \sum_{r=1}^{R_t} p_{tr} \left(\boldsymbol{\gamma}^{*,T}_{tr} \boldsymbol{E}_{t-1,r_{t-1}} \boldsymbol{\chi}_{t-1,r_{t-1}} + \varsigma_{t-1,r_{t-1}} \right)$$

end

`// Solve approximate first stage problem`

$\boldsymbol{\chi}^*_0, \alpha^*_0 \leftarrow \arg\min_{\mathcal{X}_0} \boldsymbol{c}_0 \boldsymbol{\chi}_0 + \alpha_0$ with

$$\mathcal{X}_0 = \left\{ \boldsymbol{A}_0 \boldsymbol{\chi}_0 \geq \boldsymbol{b}_0, \alpha_0 \geq \sum_{r=1}^{R_1} p_{1r} \left(\boldsymbol{\gamma}^{*,T}_{tr} \boldsymbol{E}_{t-1,r_{t-1}} \boldsymbol{\chi}_{t-1,r_{t-1}} + \varsigma_{t-1,r_{t-1}} \right) \right\}$$

`// Calculate lower bound`

$\underline{z} \leftarrow \boldsymbol{c}_0 \boldsymbol{\chi}^*_0 + \alpha^*_0$

end

return $\boldsymbol{\chi}_t$, z

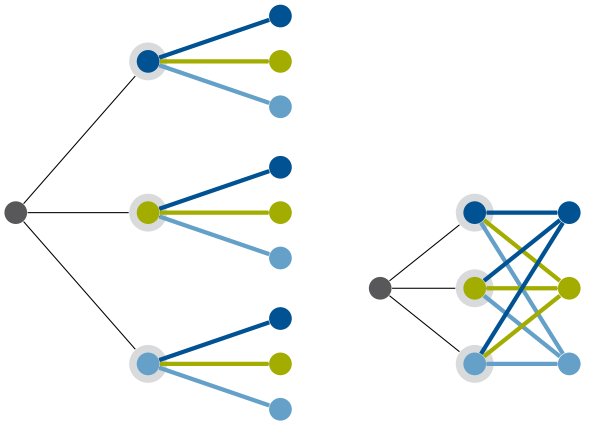


Figure 3.22: Visualization of independence assumption of the underlying stochastic process: thick edges of the same color represent the same probability to reach the following scenario; the highlighted realizations in time step $t = 1$ get the same cut.

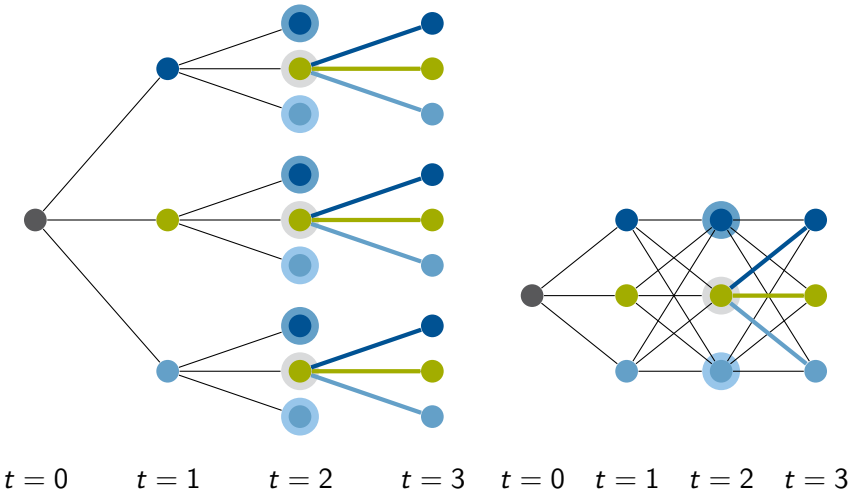


Figure 3.23: Visualization of Markov assumption of the underlying stochastic process: thick edges of the same color represent the same probability to reach the following scenario; the highlighted realizations of the same color in time step $t = 1$ get the same cut.

realization of the chosen path of the forward simulation, but to all parallel realizations of that time step. For the case, which is represented by figure 3.22, this means that the cut generated from the realizations in time step $t = 2$ can be applied to all three scenarios of time step $t = 1$ with the same weights. The expected future costs for time step $t = 2$ are the same, as all realizations in time step $t = 1$ have the same future.

The scenario tree can for the independent case be represented in a lattice with the edges representing the same probabilities, regardless of the number of stages. The resulting procedure for the independent case is summarized in algorithm 3.4.

In case of an underlying Markov process, the situation is different. The probability of a scenario is now dependent on the scenario of the previous time step. This is illustrated in

Algorithm 3.4: SDDP algorithm with independent process without upper bound calculation

Data: $\boldsymbol{A}_{tr}$, $\boldsymbol{b}_{tr}$, $\boldsymbol{c}_{tr}$, $\boldsymbol{E}_{tr}$, p_{tr}, T

Initialization:

$\bar{z} \leftarrow \infty$

$\alpha^*_{tr} \leftarrow 0 \; \forall t = 0, \dots T$

// Solve approximate first stage problem

$\boldsymbol{x}^*_0 \leftarrow \arg\min_{\mathcal{X}_0} \boldsymbol{c}_0\boldsymbol{\chi}_0 + \alpha^*_0$ with $\mathcal{X}_0 = \{\boldsymbol{A}_0\boldsymbol{\chi}_0 \geq \boldsymbol{b}_0\}$

// Calculate lower bound

$\underline{z} \leftarrow \boldsymbol{c}_0\boldsymbol{\chi}^*_0$

while $\bar{z} - \underline{z} > \epsilon$ **do**

  // Forward simulation

  for $t \leftarrow 2, \dots T$ **do**

    // Pick random realization r

    $\boldsymbol{\chi}^*_{tr}, \alpha^*_{tr} \leftarrow \arg\min_{\mathcal{X}_{tr}} \boldsymbol{c}_t\boldsymbol{\chi}_{tr} + \alpha_{tr}$ with

    $\mathcal{X}_{tr} = \left\{\boldsymbol{A}_{tr}\boldsymbol{\chi}_{tr} \geq \boldsymbol{b}_{tr} - \boldsymbol{E}_{t-1,r_{t-1}}\boldsymbol{\chi}^*_{t-1,r_{t-1}}\right\}$

    // Save picked realization r of respective time step t in r_t

  end

  // Calculation of upper bound

  // Will be discussed in a separate paragraph.

  // Backward simulation

  for $t \leftarrow T, \dots 2$ **do**

    // Iterate through all realizations of that time step

    for $r \in R_t$ **do**

      $\boldsymbol{\gamma}^*_{tr} \leftarrow \arg\min_{\mathcal{X}_{tr}} \boldsymbol{c}_t\boldsymbol{\chi}_{tr} + \alpha^*_{tr}$ with

      $\mathcal{X}_{tr} = \left\{\boldsymbol{A}_{tr}\boldsymbol{\chi}_{tr} \geq \boldsymbol{b}_{tr} - \boldsymbol{E}_{t-1,r_{t-1}}\boldsymbol{\chi}^*_{t-1,r_{t-1}} \; : \boldsymbol{\gamma}_{tr}\right\}$

        // with r_{t-1} as saved realization from forward simulation

    end

    // Iterate through all realizations of previous time step

    for $r_{t-1} \in R_{t-1}$ **do**

      // Calculate weighted cut and deploy to all realizations of previous time step

      $\alpha^*_{t-1,r_{t-1}} \geq \sum_{r=1}^{R_t} p_{tr}\left(\boldsymbol{\gamma}^{*,T}_{tr}\boldsymbol{E}_{t-1,r_{t-1}}\boldsymbol{\chi}_{t-1,r_{t-1}} + \varsigma_{t-1,r_{t-1}}\right)$

    end

  end

  // Solve approximate first stage problem

  $\boldsymbol{\chi}^*_0, \alpha^*_0 \leftarrow \arg\min_{\mathcal{X}_0} \boldsymbol{c}_0\boldsymbol{\chi}_0 + \alpha_0$ with

  $\mathcal{X}_0 = \left\{\boldsymbol{A}_0\boldsymbol{\chi}_0 \geq \boldsymbol{b}_0, \alpha_0 \geq \sum_{r=1}^{R_1} p_{1r}\left(\boldsymbol{\gamma}^{*,T}_{1r}\boldsymbol{E}_0\boldsymbol{\chi}_0 + \varsigma_0\right)\right\}$

  // Calculate lower bound

  $\underline{z} \leftarrow \boldsymbol{c}_0\boldsymbol{\chi}^*_0 + \alpha^*_0$

end

return $\boldsymbol{\chi}_t$, z

figure 3.23. All realizations in time step $t = 3$ with the same predecessor are equally probable. Therefore, a cut which is generated for a green realization in time step $t = 2$ can be used for both other green realizations in time step $t = 2$. Hence, the tree can be represented in a nested structure, but the probabilities in arriving in one scenario are not the same anymore. An extensive introduction to Markov processes can be found by Puterman [110] and an application in SDDP in [111].

Algorithm 3.5 summarizes the procedure for an underlying Markov process. In contrast to the independent case, it is important to implement the subproblems in such a way that they know their predecessors. The recent history is still important in case of Markov.

Bounds The difficulty in finding an upper bound especially lies within the structure of the multistage stochastic optimization problem. The question arises how to bound an expectation value. A common approach is to pick random samples and calculate the expectation value for them however, based on the picking of the samples, the bound can also be below the lower bound and therefore not realistic. To ensure sensible results, different approaches have been tested: see for instance Homem-de-Mello, Matos, and Finardi [112] for the development of a hypothesis test to create a confidence interval for the goodness of the bound. The used implementation will be presented later in the chapter after introducing the more elaborate cut generation from Benders decomposition.

Parallelization It can be noted that in backward recursion the iteration through all realizations per time step can be parallelized. Starting with the last time step, all realizations in that time step are optimized in parallel. The previous time step then gets per realization the new cut based on the chosen uncertainty assumption (independence, Markov) and can be solved again in parallel.

3.3 Combining decomposition and uncertainty methods

It has already been mentioned in the derivation of SDDP that the procedure is basically a so-called nested Benders decomposition algorithm: Hence, the idea on how to speed up Benders decomposition by using an improved cut generation can be applied as well to the SDDP algorithm. In the following subsections the basic ideas on what improved cuts can be, how to derive them and their properties are explained. Afterwards, the implications on using this method on SDDP are highlighted and implementation ideas are given.

3.3.1 Facet and Pareto-optimal cuts

The following section is based on the dissertation of Stursberg [113, Chapter 3]. The section summarizes the key ideas developed in the thesis and gives insight into the used cut generation.

In section 3.2.2.2, the cut generation is seen as creating a new section of the piecewise linear future cost function. The following derivations mostly use knowledge from applied geometry like separating hyperplanes and set theory. The important set in case of SDDP or Benders decomposition is the epigraph of the future cost function. In the following, the optimal

Algorithm 3.5: SDDP algorithm with Markov process without upper bound calculation

Data: $\boldsymbol{A}_{tr}$, $\boldsymbol{b}_{tr}$, $\boldsymbol{c}_{tr}$, $\boldsymbol{E}_{tr}$, p_{tr}, T

Initialization:

$\bar{z} \leftarrow \infty$

$\alpha^*_{tr} \leftarrow 0 \; \forall t = 0, \dots T$

// Solve approximate first stage problem

$\boldsymbol{x}^*_0 \leftarrow \arg\min_{\mathcal{X}_0} \boldsymbol{c}_0\boldsymbol{\chi}_0 + \alpha^*_0$ with $\mathcal{X}_0 = \{\boldsymbol{A}_0\boldsymbol{\chi}_0 \geq \boldsymbol{b}_0\}$

// Calculate lower bound

$\underline{z} \leftarrow \boldsymbol{c}_0\boldsymbol{\chi}^*_0$

while $\bar{z} - \underline{z} > \epsilon$ **do**

  // Forward simulation

  for $t \leftarrow 2, \dots T$ **do**

    // Pick random realization r

    $\boldsymbol{\chi}^*_{tr}, \alpha^*_{tr} \leftarrow \arg\min_{\mathcal{X}_{tr}} \boldsymbol{c}_t\boldsymbol{\chi}_{tr} + \alpha_{tr}$ with

    $\mathcal{X}_{tr} = \left\{\boldsymbol{A}_{tr}\boldsymbol{\chi}_{tr} \geq \boldsymbol{b}_{tr} - \boldsymbol{E}_{t-1,r_{t-1}}\boldsymbol{\chi}^*_{t-1,r_{t-1}}\right\}$

    // Save picked realization r of respective time step t in r_t

  end

  // Calculation of upper bound

  // Will be discussed in a separate paragraph.

  // Backward simulation

  for $t \leftarrow T, \dots 2$ **do**

    // Iterate through all realizations of that time step

    for $r \in R_t$ **do**

      $\boldsymbol{\gamma}^*_{tr} \leftarrow \arg\min_{\mathcal{X}_{tr}} \boldsymbol{c}_t\boldsymbol{\chi}_{tr} + \alpha^*_{tr}$ with

      $\mathcal{X}_{tr} = \left\{\boldsymbol{A}_{tr}\boldsymbol{\chi}_{tr} \geq \boldsymbol{b}_{tr} - \boldsymbol{E}_{t-1,r_{t-1}}\boldsymbol{\chi}^*_{t-1,r_{t-1}} \; : \boldsymbol{\gamma}_{tr}\right\}$

        // with r_{t-1} as saved realization from forward simulation

    end

    // Iterate only through same realizations of previous time step

    for $r_{t-1} \in R_{t-1,p}$ **do**

      // Calculate weighted cut and deploy to all realizations of previous time step

      $\alpha^*_{t-1,r_{t-1}} \geq \sum_{r=1}^{R_t} p_{tr}\left(\boldsymbol{\gamma}^{*,T}_{tr}\boldsymbol{E}_{t-1,r_{t-1}}\boldsymbol{\chi}_{t-1,r_{t-1}} + \varsigma_{t-1,r_{t-1}}\right)$

    end

  end

  // Solve approximate first stage problem

  $\boldsymbol{\chi}^*_0, \alpha^*_0 \leftarrow \arg\min_{\mathcal{X}_0} \boldsymbol{c}_0\boldsymbol{\chi}_0 + \alpha_0$ with

  $\mathcal{X}_0 = \left\{\boldsymbol{A}_0\boldsymbol{\chi}_0 \geq \boldsymbol{b}_0, \alpha_0 \geq \sum_{r=1}^{R_1} p_{1r}\left(\boldsymbol{\gamma}^{*,T}_{1r}\boldsymbol{E}_0\boldsymbol{\chi}_0 + \varsigma_0\right)\right\}$

  // Calculate lower bound

  $\underline{z} \leftarrow \boldsymbol{c}_0\boldsymbol{\chi}^*_0 + \alpha^*_0$

end

return $\boldsymbol{\chi}_t$, z

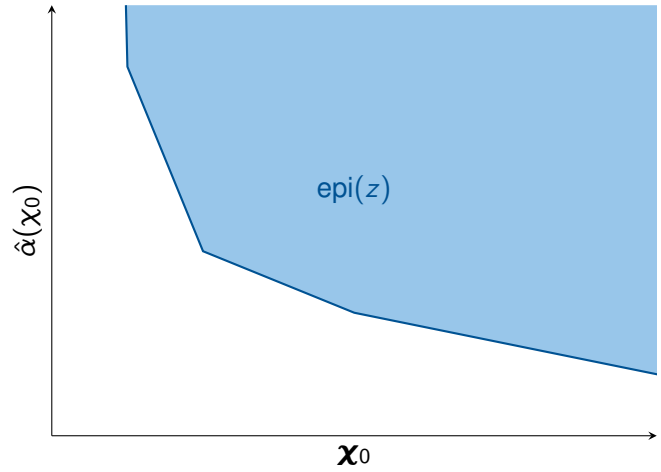


Figure 3.24: Epigraph of the future cost function

value of the second stage will be denoted by z:

$$z = \min_{\boldsymbol{\chi}_1} \boldsymbol{c}_1^\mathsf{T} \boldsymbol{\chi}_1 \tag{3.52}$$

$$\text{s.t. } \boldsymbol{A}_1 \boldsymbol{\chi}_1 \geq \boldsymbol{b}_1 - \boldsymbol{E}_0 \boldsymbol{\chi}_0. \tag{3.53}$$

The epigraph of z is equal to the epigraph of the future cost function α. It is expressed by the following (c.f. definition A.12 [113]):

$$\text{epi}(z) = \left\{ (\boldsymbol{\chi}_0, \alpha) \mid \exists \boldsymbol{\chi}_1 : \boldsymbol{A}_1 \boldsymbol{\chi}_1 \geq \boldsymbol{b}_1 - \boldsymbol{E}_0 \boldsymbol{\chi}_0,\ \boldsymbol{c}_1^\mathsf{T} \boldsymbol{\chi}_1 \leq \alpha \right\}. \tag{3.54}$$

A visualization of the epigraph can be seen in figure 3.24. In 2D, the area above the future cost function is the epigraph on which most theory applies.

As mentioned earlier, in classic Benders decomposition, cuts are distinguished between *optimality* and *feasibility cuts*. Fischetti, Salvagnin, and Zanette [114] published a different method on interpreting the cut-generating problem and instead of separating the optimality and feasibility choice, they reformulate the cut-generation problem into a pure problem of feasibility. It is equivalent to finding a hyperplane separating a given point from a set. In case of Benders decomposition, the point represents the infeasible solution, which lies outside of the given solution space, from which it should be separated by the cut. In the context of the aforementioned epigraph (c.f. figure 3.24), this corresponds to the calculated point $(\boldsymbol{\chi}_0, \alpha)$ not lying in epi(z). Therefore, a cut is a hyperplane, which separates the calculated optimum from epi(z).

This classical separation problem is also the topic of the paper by Cornuéjols and Lemaréchal [115]. They use the so-called *reverse polar set* to generate *facet cuts*, which separate the point from the given set. This is also the underlying idea of Fischetti, Salvagnin, and Zanette [114], when they developed their unified cut. The idea of the reverse polar set was originally introduced by Balas [116]. For a more detailed and descriptive introduction into his concepts and visualization, the interested reader is referred to his recent book [117].

Broadly speaking, a facet cut can be described as a cut separating the point from the respective set while touching the surface of the set. The feasible set in Benders decomposition

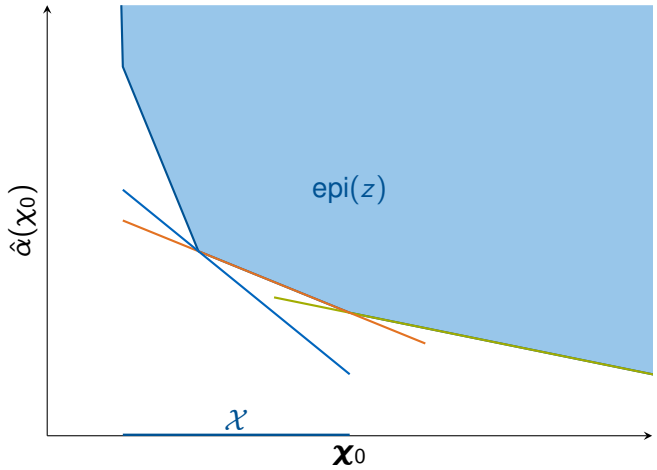


Figure 3.25: Visualization of facet and Pareto-optimal cut: Both, the orange and the green cut are facet cuts of epi(z). Only the orange cut is Pareto-optimal, as it cannot be dominated by any other cut at domain $\mathcal{X}$. The blue cut is not a facet cut but still Pareto-optimal as it leads to higher values of α in the left part of domain $\mathcal{X}$ and hence is not dominated by the orange cut.

is replaced by a linear approximation, generated from the cuts. Basically, this is the same idea as for the SDDP algorithm. A cut is often more effective if it touches a face of maximal dimension of a set, e.g. for a two-dimensional set, a cut which touches a side of the set is in most cases more effective than a cut which touches the surface just in one point.

A Pareto-optimal cut is a cut, which is not dominated by any other cut [118]. In contrast to facet cuts, Pareto-optimal cuts take the domain of the complicating variable into account. A facet cut, which does not touch the future cost function in the domain, on which $\boldsymbol{x}$ is feasible, is not Pareto-optimal. This concept is visualized in figure 3.25. A cut, which supports the epigraph in that specific domain, is most likely Pareto-optimal even if it does not support a whole surface of the epigraph. This comes due to the fact, that domination of cuts takes into account that they have to enforce a higher or equal value for α on the complete domain and strictly higher at least at one point.

Hence, a cut which would fulfill both properties (facet and Pareto-optimal) would be the most beneficial. In chapter 3 of the dissertation, Stursberg [113] focuses on the derivation of these cuts by using the definitions of the reverse polar set and the relaxed alternative polyhedron based on the approaches of Fischetti, Salvagnin, and Zanette [114] and Cornuéjols and Lemaréchal [115]. The main steps in understanding the relation between these sets and the implications for the formulation of the second stage problem are highlighted in the following. For a more detailed mathematical derivation, the interested reader is referred to the thesis of Stursberg [113].

Stursberg [113] and Cornuéjols and Lemaréchal [115] defined the reverse polar set $\mathcal{S}^-$ for a given set $\mathcal{S}$ as (which is actually a point reflection of the definition by Balas [117]):

$$\mathcal{S}^- := \left\{ \boldsymbol{\pi} \in \mathbb{R}^n \;\middle|\; \boldsymbol{\pi}^\mathsf{T} \boldsymbol{s} \leq -1 \; \forall \boldsymbol{s} \in \mathcal{S} \right\} \tag{3.55}$$

To illustrate the idea behind it, a cone set $\mathcal{S} = \{(\chi_1, \chi_2) \in \mathbb{R}^2 \mid \chi_1 \geq 1, \chi_2 \geq 1\}$ is

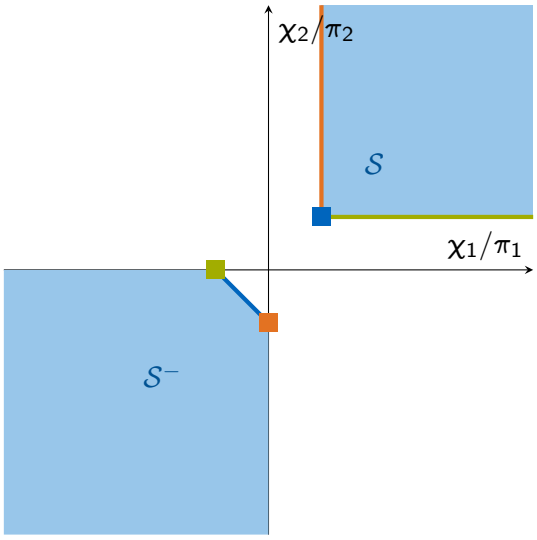


Figure 3.26: Set $\mathcal{S}$ and the corresponding reverse polar set $\mathcal{S}^-$: Vertices of $\mathcal{S}^-$ correspond to facets of $\mathcal{S}$ and vice versa, indicated by the colors. While set $\mathcal{S}$ is defined over χ_i, set $\mathcal{S}^-$ is defined over π_i, hence the labeling of the axes.

visualized in the first quadrant of figure 3.26. The corresponding reverse polar set is then given by:

$$\mathcal{S}^- := \left\{ \begin{bmatrix} \pi_1 \\ \pi_2 \end{bmatrix} \in \mathbb{R}^2 \;\middle|\; \pi_1 + \pi_2 \leq -1,\ \pi_1 \leq 0,\ \pi_2 \leq 0 \right\}. \tag{3.56}$$

The vertices of the reverse polar set $\mathcal{S}^-$ correspond to the facets of set $\mathcal{S}$. For an example, the scalar product of any vector ending at the blue facet of the reverse polar and the blue vertex of set $\mathcal{S}$ fulfill exactly the inequality (3.55). Vice versa, the facets of set $\mathcal{S}$ can be easily represented by the vertices of the reverse polar set.

The reverse polar set of interest of a given optimal point $(\boldsymbol{\chi}^*, \alpha^*)$ is derived by Stursberg [113, p. 105] as follows:

$$\mathcal{R} = (\text{epi}(z) - (\boldsymbol{\chi}^*, \alpha^*))^- \tag{3.57}$$

$$= \left\{ \begin{bmatrix} \boldsymbol{\pi} \\ \pi_0 \end{bmatrix} \;\middle|\; \exists \boldsymbol{\gamma} :\ \begin{bmatrix} \boldsymbol{\pi}^\mathsf{T} & \pi_0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\chi}_0^* \\ \alpha \end{bmatrix} - \boldsymbol{\gamma}^\mathsf{T} \boldsymbol{b}_1 \geq 1, \boldsymbol{\gamma}^\mathsf{T} \boldsymbol{A}_1 - \pi_0 \boldsymbol{c}_1^\mathsf{T} = \boldsymbol{0},\ \boldsymbol{\gamma}^\mathsf{T} \boldsymbol{E}_0 = \boldsymbol{\pi}^\mathsf{T} \right\}. \tag{3.58}$$

For the derivation of the reverse polar set, Stursberg uses theorem 3.6 [113, p. 101] in which all normal vectors are separating a point $(\boldsymbol{\chi}^*, \alpha^*)$ from the epigraph are characterized. With help of the theorem and the fact that it is assumed that the epigraph is homogeneous the mathematical representation (3.58) can be stated. As the derivation of the formulation is not as important for the derivation of the cuts as the representation itself, it is not explained in more detail. The mathematical representation sets the origin of the reverse polar set dependent on the current solution $(\boldsymbol{\chi}^*, \alpha^*)$.

Comparing figure 3.26 with a representation 3.2 of the future cost function $\hat{\alpha}$, one might notice that it would be helpful for the linear approximation to get the facets of the epigraph of the future cost function. As presented, vertices of the reverse polar lead to facets of the original

set. Hence, if there would be a method on how to get the reverse polar of the epigraph of the future cost function, the linear approximation could be attained more easily.

Unfortunately, the mathematical representation of the reverse polar set of a given feasible set is mostly difficult to get. Fischetti, Salvagnin, and Zanette [114] on the other hand, do not use the reverse polar set, but the alternative polyhedron to calculate cuts:

$$\mathcal{P} = \left\{ \begin{bmatrix} \boldsymbol{\gamma} \\ \gamma_0 \end{bmatrix} \geq \mathbf{0} \;\middle|\; \boldsymbol{\gamma}^\mathsf{T} \boldsymbol{A}_1 + \gamma_0 \boldsymbol{c}_1^\mathsf{T} = \mathbf{0},\ \boldsymbol{\gamma}^\mathsf{T}(\boldsymbol{b}_1 - \boldsymbol{E}_0 \boldsymbol{\chi}_0^*) + \gamma_0 \alpha^* = -1 \right\} \tag{3.59}$$

The point $(\boldsymbol{\chi}^*, \alpha^*)$ represents the current optimal solution of the master problem of Benders decomposition. The set contains all valid Farkas certificates (c.f. [113, p. 101]). These certificates can be derived from Farkas' Lemma [119, 120]. It basically states that for a given matrix $\boldsymbol{A}$ and a given vector $\boldsymbol{b}$ exactly one of the two statements holds:

1. There exists a vector $\boldsymbol{\chi} \geq \mathbf{0}$ for which the equation system $\boldsymbol{A}\boldsymbol{\chi} = \boldsymbol{b}$ is valid.
2. There exists a vector $\boldsymbol{\upsilon}$ for which the equation systems $\boldsymbol{A}^\mathsf{T}\boldsymbol{\upsilon} \geq \mathbf{0}$ and $\boldsymbol{b}^\mathsf{T}\boldsymbol{\upsilon} < \mathbf{0}$ are valid.

The lemma can be interpreted as follows: If vector $\boldsymbol{b}$ does not lie in the cone spanned by the columns of matrix $\boldsymbol{A}$, there exists a hyperplane with normal vector $\boldsymbol{\upsilon}$ which separates vector $\boldsymbol{b}$ from the cone. For the set of feasible solutions $\mathcal{F}$ for subproblem (3.52)-(3.53) holds:

$$\mathcal{F} = \left\{ \boldsymbol{\chi}_1 \;\middle|\; \boldsymbol{A}_1\boldsymbol{\chi}_1 \geq \boldsymbol{b}_1 - \boldsymbol{E}_0\boldsymbol{\chi}_0^*,\ \boldsymbol{c}_1^\mathsf{T}\boldsymbol{\chi}_1 \leq \alpha^* \right\}. \tag{3.60}$$

This equation system is only valid, if there exists no hyperplane with normal vector $\begin{bmatrix} \boldsymbol{\gamma} & \gamma_0 \end{bmatrix}$, separating vector $\begin{bmatrix} \boldsymbol{b} - \boldsymbol{E}_0\boldsymbol{\chi}_0^* & \alpha^* \end{bmatrix}^\mathsf{T}$ from the cone spanned by the columns of $\boldsymbol{A}_1$. All of these normal vectors are collected in the alternative polyhedron $\mathcal{P}$ (3.59). Instead of forcing the strictly smaller than zero property of the second statement of Farkas' Lemma, the normal vectors are scaled in such a way that they equal –1.

Even if definition (3.55) looks more neat than definition (3.59), the alternative polyhedron for a given solution is easier to calculate than the reverse polar $\mathcal{R}$ (3.58).

To connect both the reverse polar set and the alternative polyhedron, Stursberg [113, p. 107] first refers to the relaxed alternative polyhedron:

$$\mathcal{P}^{\leq} = \left\{ \begin{bmatrix} \boldsymbol{\gamma} \\ \gamma_0 \end{bmatrix} \geq \mathbf{0} \;\middle|\; \boldsymbol{\gamma}^\mathsf{T} \boldsymbol{A}_1 + \gamma_0 \boldsymbol{c}_1^\mathsf{T} = \mathbf{0},\ \boldsymbol{\gamma}^\mathsf{T}(\boldsymbol{b}_1 - \boldsymbol{E}_0 \boldsymbol{\chi}_0^*) + \gamma_0 \alpha^* \leq -1 \right\}. \tag{3.61}$$

For the case, in which the inequality is fulfilled with equality, the relaxed version matches the original polyhedron $\mathcal{P}$. The relation between the reverse polar set and the relaxed alternative polyhedron can then be expressed by:

$$\mathcal{R} = \begin{bmatrix} \boldsymbol{E}_0^\mathsf{T} & \mathbf{0} \\ \mathbf{0}^\mathsf{T} & -1 \end{bmatrix} \cdot \mathcal{P}^{\leq} \tag{3.62}$$

$$= \left\{ \begin{bmatrix} \boldsymbol{E}_0^\mathsf{T}\boldsymbol{\gamma} \\ -\gamma_0 \end{bmatrix} \;\middle|\; \begin{bmatrix} \boldsymbol{\gamma} \\ \gamma_0 \end{bmatrix} \geq \mathbf{0},\ \boldsymbol{\gamma}^\mathsf{T} \boldsymbol{A}_1 + \gamma_0 \boldsymbol{c}_1^\mathsf{T} = \mathbf{0},\ \boldsymbol{\gamma}^\mathsf{T}(\boldsymbol{b}_1 - \boldsymbol{E}_0 \boldsymbol{\chi}_0^*) + \gamma_0 \alpha^* \leq -1 \right\} \tag{3.63}$$

$$= \left\{ \begin{bmatrix} \boldsymbol{\pi} \\ \pi_0 \end{bmatrix} \;\middle|\; \boldsymbol{\pi} = \boldsymbol{E}_0^\mathsf{T}\boldsymbol{\gamma},\ \boldsymbol{\gamma}^\mathsf{T} \boldsymbol{A}_1 - \pi_0 \boldsymbol{c}_1^\mathsf{T} = \mathbf{0},\ \boldsymbol{\gamma}^\mathsf{T}\boldsymbol{b}_1 - \boldsymbol{\pi}^\mathsf{T}\boldsymbol{\chi}_0^* - \pi_0 \alpha^* \leq -1 \right\} \tag{3.64}$$

This relation shows that the reverse polar set is nothing else than a linear transformation of the relaxed alternative polyhedron. If matrix $\boldsymbol{E}_0$ has more rows than columns, the linear transformation can be understood to be a projection of the relaxed alternative polyhedron into the lesser reverse polar set space. Hence, finding a vertex of the relaxed alternative polyhedron does not necessarily lead to a vertex of the reverse polar set.

To get an optimal supporting cut, one has to refer to getting the optimal vertex of the reverse polar set $\mathcal{R}$ with respect to the current solution (c.f. Theorem 3.15 [113], rephrased Theorem 2.3 [115]):

$$\max_{\boldsymbol{\pi},\pi_0} \boldsymbol{\omega}^\mathsf{T}\boldsymbol{\pi} + \omega_0\pi_0 \tag{3.65}$$

$$\text{s.t. } \begin{bmatrix}\boldsymbol{\pi}\\ \pi_0\end{bmatrix} \in \mathcal{R} \tag{3.66}$$

The direction of the cost vector $\begin{bmatrix}\boldsymbol{\omega} & \omega_0\end{bmatrix}^\mathsf{T}$ is important for the unbounded reverse polar set. If the cost vector is chosen from the closed conical hull of the shifted epigraph $(\text{epi}(z) - (\boldsymbol{\chi}^*, \alpha^*))$, then optimization problem (3.65)-(3.66) is feasible. While Fischetti, Salvagnin, and Zanette [114] choose $\begin{bmatrix}\boldsymbol{\omega}\\ \omega_0\end{bmatrix} = \mathbf{1}$ and receive promising results, Stursberg [113] derives that varying the cost vector can lead not only to facet cuts but also to Pareto-optimal cuts. As the choice of $\boldsymbol{\omega}$ also affects the calculation of the upper bound, variations of $\boldsymbol{\omega}$ are not taken into account for this thesis.

Due to the relation of relaxed alternative polyhedron and reverse polar set (3.62), if the dual variables $\boldsymbol{\gamma}$ can be mapped to a vertex $\boldsymbol{\pi}$ of the reverse polar $\mathcal{R}$, the solution leads to a facet cut. Stursberg [113] derives the following optimization problem, which is a relaxed version of the original subproblem (3.52)-(3.53):

$$\min_{\lambda,\boldsymbol{\chi}_1} \lambda \tag{3.67}$$

$$\text{s.t. } \boldsymbol{A}_1\boldsymbol{\chi}_1 \geq \boldsymbol{b}_1 - \boldsymbol{E}_0\left(\boldsymbol{\chi}_0^* + \lambda\boldsymbol{\omega}\right) \tag{3.68}$$

$$\boldsymbol{c}_1^\mathsf{T}\boldsymbol{\chi}_1 \leq \alpha^* + \omega_0\lambda \tag{3.69}$$

Both the objective (3.52) and the constraints (3.53) are relaxed by the scalar λ. If the first stage variable $\boldsymbol{\chi}_0$ is also optimized in the second stage and set in a separate constraint $\boldsymbol{\chi}_0 = \boldsymbol{\chi}_0^*$, optimization problem (3.67)-(3.69) becomes:

$$\min_{\lambda,\boldsymbol{\chi}_0,\boldsymbol{\chi}_1} \lambda \tag{3.70}$$

$$\text{s.t. } \boldsymbol{A}_1\boldsymbol{\chi}_1 + \boldsymbol{E}_0\boldsymbol{\chi}_0 \geq \boldsymbol{b}_1 \tag{3.71}$$

$$\boldsymbol{\chi}_0 = \boldsymbol{\chi}_0^* + \lambda\boldsymbol{\omega} \tag{3.72}$$

$$\boldsymbol{c}_1^\mathsf{T}\boldsymbol{\chi}_1 \leq \alpha^* + \omega_0\lambda \tag{3.73}$$

By using the first stage variable, the structure of the optimization problem lets the choice of $\boldsymbol{\omega}$ get easier, as every value is possible and still leads almost always to a facet cut.

The corresponding cut for this optimization problem is given by:

$$\boldsymbol{\chi}_0^\mathsf{T}\boldsymbol{\gamma}_\omega + \alpha\gamma_\alpha \geq \boldsymbol{b}_1^\mathsf{T}\boldsymbol{\gamma}. \tag{3.74}$$

It can be easily rewritten in the typical feasibility cut representation by using:

$$\boldsymbol{b}_1^{\mathsf{T}}\boldsymbol{\gamma} - \boldsymbol{\chi}_0^{*,\mathsf{T}}\boldsymbol{\gamma}_\omega - \alpha^*\gamma_\alpha = \lambda^*. \tag{3.75}$$

Therefore, $\boldsymbol{b}_1^{\mathsf{T}}\boldsymbol{\gamma}$ can be replaced by the optimal values:

$$\boldsymbol{b}_1^{\mathsf{T}}\boldsymbol{\gamma} = \lambda^* + \boldsymbol{\chi}_0^{*,\mathsf{T}}\boldsymbol{\gamma}_\omega + \alpha^*\gamma_\alpha. \tag{3.76}$$

Hence, the cut is given by:

$$\lambda^* \leq \boldsymbol{\gamma}_\omega^{\mathsf{T}}(\boldsymbol{\chi}_0 - \boldsymbol{\chi}_0^*) + \boldsymbol{\gamma}_\alpha^{\mathsf{T}}(\alpha - \alpha^*). \tag{3.77}$$

The resulting procedure is summarized in algorithm 3.6. How this affects the choosing of cuts is illustrated by the following example:

Example 3.2.1 (continuing from p. 47) *Based on the DDP example with two stages and the company, which wants to optimally fulfill its demand, the idea of the relaxed problem and the resulting cut generation is shown. After optimizing the first stage problem as already shown, the optimal solutions for the storage $\chi_{B_1}^* = 0$ and for the future costs is given by $\alpha^* = 0$. Hence, the relaxed second stage problem is given by:*

$$\begin{aligned}
\min_{\lambda,\chi_{B_0},\chi_{W_1},\chi_{G_1},\chi_{B_1}} \quad & \lambda \\
s.t. \quad & \chi_{B_0} + \chi_{W_1} + \chi_{G_1} - \chi_{B_1} = 3 \\
& \chi_{W_1} \leq 2 \\
& \chi_{G_1} \leq 5 \\
& \chi_{B_t} \leq 10,\ \forall t \in \{0,1\} \\
& \chi_{B_1} = \chi_{B_1}^* + \lambda\omega \\
& 2\chi_{W_1} + 5\chi_{G_1} + \chi_{B_1} \leq \alpha^* + \lambda\omega_0 \\
& \chi_{P_1} \geq 0\ \forall P \in \{W, G, B\} \\
& \chi_{B_0} \geq 0
\end{aligned}$$

The optimization of the relaxed problem leads to $\lambda^* = 2$, $\chi_{W_1}^* = 1$, $\chi_{G_1}^* = \chi_{B_1}^* = \chi_{B_0}^* = 0$, $\gamma_\omega^* = \frac{2}{3}$, $\gamma_\alpha^* = \frac{1}{3}$. Hence, the cut after one iterations reads as follows:

$$\begin{aligned}
2 &\leq \frac{2}{3}\chi_{B_0} + \frac{1}{3}\alpha \\
\alpha &\geq -2\chi_{B_0} + 6
\end{aligned}$$

This equals the cut, the second iteration has yielded in the DDP section. If another iteration with the relaxed problem is executed, the cut calculated at the first iteration would be the result.

3.3.2 Implementation of SDDP in power system analysis

Implementing the cut idea and the structural base for it into a SDDP algorithm is straight forward. However, there are some issues when it is applied to power system analysis. This subsection highlights specifics which have been faced by implementing the SDDP algorithm with advanced cut generation introduced before into the *urbs* framework [60, 121, 82].

Algorithm 3.6: SDDP algorithm with combined cuts

Data: $\boldsymbol{A}_{tr}$, $\boldsymbol{b}_{tr}$, $\boldsymbol{c}_{tr}$, $\boldsymbol{E}_{tr}$, p_{tr}, T

Initialization:

$\bar{z} \leftarrow \infty$

$\alpha^*_{tr} \leftarrow 0 \; \forall t = 0, \dots T$

// Solve approximate first stage problem

$\boldsymbol{x}^*_0 \leftarrow \arg\min_{\mathcal{X}_0} \boldsymbol{c}_0\boldsymbol{\chi}_0 + \alpha^*_0$ with $\mathcal{X}_0 = \{\boldsymbol{A}_0\boldsymbol{\chi}_0 \geq \boldsymbol{b}_0\}$

// Calculate lower bound

$\underline{z} \leftarrow \boldsymbol{c}_0\boldsymbol{\chi}^*_0$

while $\bar{z} - \underline{z} > \epsilon$ **do**

  // Forward simulation

  for $t \leftarrow 2, \dots T$ **do**

    // Pick random realization r

    $\boldsymbol{\chi}^*_{tr}, \alpha^*_{tr} \leftarrow \arg\min_{\mathcal{X}_{tr}} \boldsymbol{c}_t\boldsymbol{\chi}_{tr} + \alpha_{tr}$ with

    $\mathcal{X}_{tr} = \left\{\boldsymbol{A}_{tr}\boldsymbol{\chi}_{tr} \geq \boldsymbol{b}_{tr} - \boldsymbol{E}_{t-1,r_{t-1}}\boldsymbol{\chi}^*_{t-1,r_{t-1}}\right\}$

    // Save picked realization r of respective time step t in r_t

  end

  // Calculation of upper bound

  // Discussed in implementation

  // Backward simulation

  for $t \leftarrow T, \dots 2$ **do**

    // Iterate through all realizations of that time step

    for $r \in R_t$ **do**

      $\lambda^*_{tr}, \boldsymbol{\gamma}_{tr\alpha}, \boldsymbol{\gamma}^*_{tr\omega} \leftarrow \arg\min_{\lambda,\mathcal{X}_{tr}} \lambda_{tr}$ with $\mathcal{X}_{tr} =$

      $\Big\{\boldsymbol{A}_{tr}\boldsymbol{\chi}_{tr} \geq \boldsymbol{b}_{tr} - \boldsymbol{E}_{t-1,r_{t-1}}\boldsymbol{\chi}^*_{t-1,r_{t-1}} \; : \boldsymbol{\gamma}_{tr};\; \boldsymbol{c}_t\boldsymbol{\chi}_{tr} \leq \alpha^*_{tr} + \omega_0\lambda_{tr} \; : \gamma_{tr\alpha};$

      $\boldsymbol{\chi}_{t-1,r_{t-1}} = \boldsymbol{\chi}^*_{t-1,r_{t-1}} + \boldsymbol{\omega}\lambda_{tr} \; : \gamma_{tr\omega}\Big\}$

        // with r_{t-1} as saved realization from forward simulation

    end

    // Iterate through all realizations of previous time step

    for $r_{t-1} \in R_{t-1}$ **do**

      // Calculate weighted cut and deploy to all realizations of previous time step

      $\sum_{r=1}^{R_t} p_{tr}\lambda^*_{tr} \leq$

      $\sum_{r=1}^{R_t} p_{tr}\left(\boldsymbol{\gamma}^\top_{tr\omega}\boldsymbol{E}_0(\boldsymbol{\chi}_{t-1,r-1} - \boldsymbol{\chi}^*_{t-1,r_{t-1}}) + \boldsymbol{\gamma}^\top_{tr\alpha}(\alpha_{t-1,r_{t-1}} - \alpha^*_{t-1,r_{t-1}})\right)$

    end

  end

  // Solve approximate first stage problem

  $\boldsymbol{\chi}^*_0, \alpha^*_0 \leftarrow \arg\min_{\mathcal{X}_0} \boldsymbol{c}_0\boldsymbol{\chi}_0 + \alpha_0$ with $\mathcal{X}_0 =$

  $\left\{\boldsymbol{A}_0\boldsymbol{\chi}_0 \geq \boldsymbol{b}_0, \sum_{r=1}^{R_1} p_{1r}\lambda^*_{1r} \leq \sum_{r=1}^{R_1} p_{1r}\left(\boldsymbol{\gamma}^\top_{1r\omega}\boldsymbol{E}_0(\boldsymbol{\chi}_0 - \boldsymbol{\chi}^*_0) + \boldsymbol{\gamma}^\top_{1r\alpha}(\alpha_0 - \alpha^*_0)\right)\right\}$

  // Calculate lower bound

  $\underline{z} \leftarrow \boldsymbol{c}_0\boldsymbol{\chi}^*_0 + \alpha^*_0$

end

return $\boldsymbol{\chi}_t$, z

Variable	Code equivalent	Description
κ_{vp}	`cap_pro`	Process capacity
κ_{vt}	`cap_tra`	Transmission capacity
κ_{vs}^{c}	`cap_sto_c`	Storage content capacity
κ_{vs}^{p}	`cap_sto_p`	Storage power
$\epsilon_{vst}^{\mathrm{con}}$	`e_sto_con`	Storage energy content
$\rho_{vct}^{\mathrm{state}}$	`e_co_stock_state`	Summed consumed stock commodity amount

Table 3.5: Variables of subproblem set by the previous problem in SDDP implementation of *urbs*

3.3.2.1 Exchange variables

Using SDDP implies a decomposition in the time domain. Whether every time step builds its own subproblem or several time steps connect to one subproblem depends on the research question. Wozabal, Graf, and Hirschmann [122] highlight that not only short-term analysis but also mid- to long-term analysis might be of interest for stochastic methods. For price deviations of the day-ahead market, an hourly separation of the subproblems is more sensible than for a one year simulation.

In case of the linear optimization framework *urbs*, the critical variables and constraints are those which do connect several time steps. Typical examples for a classic temporal decomposition of these are CO_2 limits per region or state, installed capacities or storage contents. In case of SDDP, the master problem has no overview over all time steps just the start time steps. Therefore, the master problem optimizes the expansion of processes, e.g. power plants, and storages. All variables, which are set from one problem to the next, are summarized in table 3.5.

Each subproblem has its own copy of these variables. In case of the time dependent storage content $\epsilon_{vst}^{\mathrm{con}}$, the copy is only for the first time step of the respective subproblem. All of the listed variables are set via the result from the previous problem via the relaxed constraint (3.72).

3.3.2.2 Uncertainty

Deviations of the renewable power input because of uncertainty are controlled by two main parameters: the factor on how much the output varies from the original data and the probability for such an event. As the power output of renewable sources is given as a time series in the respective resolution, e.g. one hour, one factor is set for the whole time series per scenario. For three wind realizations *low*, *mid* and *high*, the deviation factor $\delta_r \in \{-1, 1\}$ states the convex combination of the original input and the maximum/minimum for each realization r:

$$s_r = \begin{cases} (1+\delta_r)s & \delta_r \leq 0 \\ \delta_r + (1-\delta_r)s & \delta_r > 0 \end{cases}. \tag{3.78}$$

For a given capacity factor s in time step t, $\delta_r > 0$ states how much percent of the remaining gap to a full usage of the installed renewable capacity can be used, $\delta_r < 0$ describes the

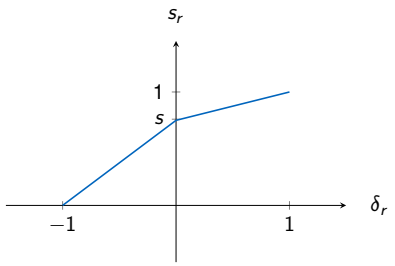


Figure 3.27: Deviation factor for uncertain renewable input generation. A $\delta_r > 0$ results in a higher output than the original data s, $\delta_t < 0$ linearly lowers the output.

percentage in case of a lower available capacity input. Figure 3.27 illustrates the concept on how the uncertain output s_r of realization r deviates dependent on δ_r.

3.3.2.3 Upper bound

Using the relaxed problem formulation (3.70)-(3.73) ensures facet cuts, but leads to problems with getting an upper bound. Stursberg [113, Section 3.4.1] derives how to develop a suitable upper bound, in case that the relaxed subproblems are not feasible without the relaxation (which is most likely the case). Typical examples of when his calculation comes into play is for relaxing the installed capacity of a process in the subproblem. This can lead to an infeasible subproblem if the relaxation is neglected as there might not be enough remaining process capacities to fulfill the given demand. The idea of Stursberg [113] is to calculate the costs for installing the relaxed process capacities and adding them to the current subproblem objectives.

In case of relaxed global constraints like CO_2 limits, the strategy is different. As it is not obvious which processes can replace the ones which generate too much CO_2 output, the costs cannot be easily determined. Hence, if the constraint violation is not as high anymore, the costs are calculated via choosing the most expensive power plant which has additional capacities left.

Storage content constraints are handled similarly to limits like CO_2: If the violation is relatively small, the additional costs are calculated by the most expensive remaining power plant.

Stursberg [113] mentions another technique to create an upper bound: By setting $\begin{bmatrix} \boldsymbol{\omega} \\ \omega_0 \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix}$ a classic feasibility cut can be created. As these feasibility cuts are not as helpful as the presented cuts, this is only a fallback option, if there is no other available upper bound.

For calculating a valid upper bound for SDDP, one has to keep in mind that by choosing a random path in the forward iteration, the upper bound created from it might be too low due to beneficial renewable input. For this specific implementation, the convergence criterion is firstly checked after a certain amount of iterations to ensure that several paths are taken into account. The upper bound is calculated via a one-sided confidence interval as stated in [102, p. 256].

3.4 Summary

In this chapter, an introduction to decomposition techniques for linear programs has been given. Additionally, an overview of uncertainty modeling techniques and a solution algorithm for probabilistic modeled uncertainty – SDDP – has been explained. The algorithm has been extended by a new cut strategy for Benders decomposition.

The chapter can be seen as a starting point for using SDDP in linear optimization programs in power system analysis. As the code is published under an open license, other researchers can easily build new ideas on top of it. There are various extensions which can be incorporated, like scenario reduction via preprocessing [123], advanced sampling strategies [124, 112] or an application-dependent refinement of the upper bound implementation.

Chapter 4

Case studies

In the previous chapters, energy system models based on open data for Bavaria and Germany have been presented. Additionally, methods on how to model uncertainty in linear optimization problems have been highlighted.

This chapter applies the presented methods to the German model and presents numerical results. Firstly, the presented models get validated before motivating the choice of the wind capacity factor as uncertain parameter. The chapter continues by highlighting the scenario assumptions and finally, showing numerical results for a short term and a long term study. Both studies are calculated with a perfect foresight and a stochastic optimization solved by Stochastic Dual Dynamic Programming (SDDP) comparing the results of both approaches. The code and input data to model, create and interpret the results are published openly [74, 82].

The aim of the chapter are new insights into investment and dispatch decision based on uncertainty and to guide an energy system modeler through the steps of how to work with results in the context of uncertainty.

4.1 Model validation

At the beginning of this chapter, the model for Germany is validated for its accuracy. Of course, as a linear optimization problem is used to approximate a complex system like a power system, certain inaccuracies are expected.

The main output commodities of both case studies are electricity and CO_2. For validation, statistics of both regions are taken and compared to the modeling results. The results are created by using the input data (c.f. chapter 2) and optimizing the dispatch by not allowing capacity expansion of any technology.

For the comparison of power output, the statistic of “Länderarbeitskreis Energiebilanzen (LAE)” [73], which is a joint project of all states in Germany to collect their energy data, is used. The sum of power per used commodity is published on their website and is therefore suitable for a comparison on a country and state-wide level.

The German model validation is illustrated in figure 4.1. It can be seen that the *urbs* model produces 0.5 % (= 3 TWh) more power than stated in the reference. This is due to two facts: On the one hand, imports and exports are not considered, meaning that disadvantageous

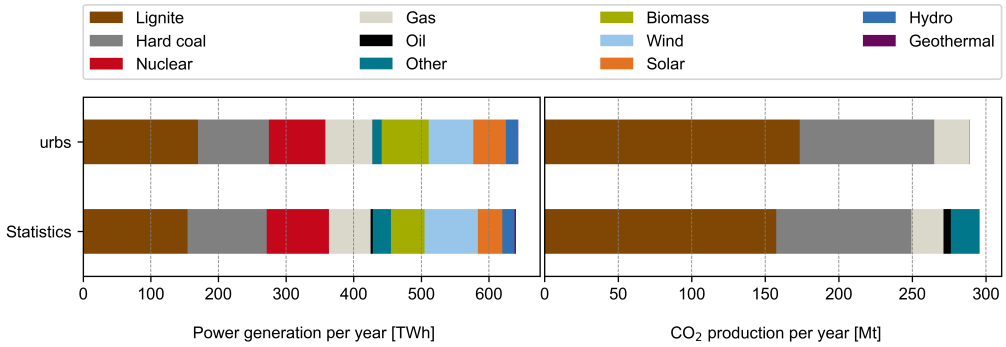


Figure 4.1: Comparison of German case study (urbs) with LAE [73] and Federal Ministry for Economic Affairs and Energy (FMEAE) [72]

times of renewable energy source availability cannot be equalized easily by imports from other countries and the other way around. On the other hand, as transmission and storage losses are already included by the demand time series, they are accounted for twice. As the distributions of these losses over time are not known by the statistics, this error is knowingly taken into account.

Power plant generation from biomass and combined cycle plants, which are a part of the gas quantity, diverge from the statistics (20 TWh and 8 TWh more than in the reference case) due to the fact that these plants are not only driven by electricity need but heat demand as well. As only the electricity sector is taken into account, subsidies for using these plants are included in the model as negative variable costs, however, the times of their real dispatch cannot be modeled accordingly.

In case of “other” (mostly waste), reasons for the difference are similar. As the main driver for choosing a power plant is the fuel and/or variable price, additional incentives such as usage of heat or mandatory run times are not taken into account. If there is data about such run times, they could be added to the model as additional constraints.

Another difference results from lignite and hard coal power plants. While the model for Germany produces 15 TWh more out of lignite, it produces 12 TWh less power from hard coal. In this case, regional distribution, conversion factors and prices might be the reason for this deviation.

Additionally, there is a mismatch of renewable power production in the Germany model. As mentioned in chapter 2, each of the 17 regions (16 states + one offshore region) has its own capacity factor for wind, solar and hydro. As there are no time series available, which represent the whole region, only one arbitrary point per region is taken as representation. For the solar time series, the orientation can be chosen. The time series used consists of a weighted sum of east, south and west orientations. For all regions, the same weighting (with south as main part) is used, which might also not be accurate, hence there is a overproduction of 32 % compared to the solar production of LAE statistic. For wind, it is the other way around and 16 % less than actually produced are calculated by the optimization model.

Regarding CO_2, all states of Germany have a set maximum bound on how much CO_2 they are allowed to produce during the time span of one year. These values are taken from the LAE statistic as well. This data set includes only the produced sum per state. It is not divided by

production commodity. In case of FMEAE's Energy Data, CO_2 statistics include output per commodity, but instead have no regional distribution. Due to this fact, both produced CO_2 sums diverge: the model of Germany has 2.25 % less CO_2 emissions than the statistics (equaling 6.6 Mt CO_2). As both statistics diverge for no obvious reasons, this result is difficult to validate. Due to the fact that the distribution of CO_2 by the commodities fits to the electricity production, it is taken as acceptable.

In general, it can be said that the linear model derived from mixed sources of open data gives a satisfying result under the highlighted circumstances. The mentioned problems regarding data inconsistency underline again results and requests from several institutes and project groups (c.f. [125] or [126]) for open data publishing.

4.2 Model sensitivity

> Models are often used in high polarized contexts and uncertainty may be used instrumentally.
>
> Saltelli et al. [1]

As Saltelli et al. [1] expressed in their quote, uncertainty, which is always apparent in science, has to be dealt with in a conscious manner. In their book about sensitivity analysis, they provide methods on how to deal with uncertainty in models. This thesis does not incorporate the proposed methods like variance based methods, as the underlying optimization problem is linear. Instead, it focuses on derivative methods based on dual variables [31].

Before going into detail on how to incorporate uncertainty in the presented model of Germany, one has to understand what to choose as uncertain parameter: nearly every parameter is – in some sense – uncertain. Costs, for example, are highly dependent on many factors and cannot be predicted easily. In chapter 2, it was already shown, how uncertain even power plant capacities can be.

For this thesis, the volatility of the renewable energy sources is taken for the sensitivity analysis. There are many works, which also focused onto this specific topic, e.g. the book chapter by Wozabal, Graf, and Hirschmann [122]. They discuss the impact of integrating renewable energy sources to the power market. The choice of focusing on these sources is justified by their volatile behavior and the difficulties of forecasting especially solar and wind supply.

As an extensive analysis for a similar linear optimization problem and its effect on storage usage and expansion has already been conducted by Kühne [29], only the motivation for choosing wind as uncertain parameter is highlighted.

As explained by Deif [31, p. 177], summarizing the work of Saaty and Gass [127, 128, 129], the dual variable of a constraint includes information about how a deviation of parameters of the constraint affects the solution of the optimization problem, both in terms of objective function and optimization variables. This can explained with help of the formulation of a linear

optimization problem, with primal variable $\boldsymbol{\chi}$ and dual variable $\boldsymbol{\lambda}$:

$$\min_{\boldsymbol{\chi}} \boldsymbol{c}^{\mathsf{T}}\boldsymbol{\chi} \tag{4.1}$$

$$\text{s.t. } \boldsymbol{A}\boldsymbol{\chi} = \boldsymbol{b} \ : \boldsymbol{\lambda} \tag{4.2}$$

$$\boldsymbol{\chi} \geq \boldsymbol{0} \tag{4.3}$$

The sensitivity of the objective function with respect to the changes of parameters $a_{ij} = [\boldsymbol{A}]_{ij}$ and $b_i = [\boldsymbol{b}]_i$ is given by:

$$\frac{\partial \boldsymbol{c}^{\mathsf{T}}\boldsymbol{x}}{\partial a_{ij}} = \lambda_i^* \chi_j^* \tag{4.4}$$

$$\frac{\partial \boldsymbol{c}^{\mathsf{T}}\boldsymbol{x}}{\partial b_i} = -\lambda_i^* \tag{4.5}$$

A '$*$' denotes the value of the variable in the optimum. Equation (4.4) represents the sensitivity of the cost function with respect to a deviation of the ij-th entry of matrix $\boldsymbol{A}$. For equation (4.5), it is similar, but the deviation lies within the right-hand-side parameter $\boldsymbol{b}$.

In case of renewable energy sources represented in the *urbs* framework, the power output of fluctuating sources is given by:

$$\epsilon_{vpct}^{\text{out}} = \kappa_{vp} s_{vct} r_{pc} \quad \forall v \in V, p \in P, c \in C_{\text{sup}}, t \in T_m \tag{4.6}$$

The capacity in site v of process p is given by the variable κ_{vp}. s_{vct} denotes the percentage of available capacity of an intermittent resource in the respective time step t, also called *capacity factor*. The ratio r_{pc} is the multiplication of input and output ratio of the process for commodity c. Sets V, P, C_{sup} and T_m denote the modeled sites, processes, fluctuating sources and time steps, respectively.

Reformulation of constraint (4.6) for separation of variables and parameters yields:

$$\begin{bmatrix} 1 & -s_{vct} r_{pc} \end{bmatrix} \begin{bmatrix} \epsilon_{vpct}^{\text{out}} \\ \kappa_{vp} \end{bmatrix} = 0 \ : \lambda_{vct} \tag{4.7}$$

For both case studies, the efficiency and losses of renewable energy sources are already inside parameter s_{vct}. Therefore, ratios r_{pc} are set to one. The sensitivity of the cost function ξ with respect to the deviation of the availability s_{vct} is then given by:

$$\frac{\partial \xi}{\partial s_{vct}} = \lambda_{vct}^* \kappa_{vp}^* \tag{4.8}$$

The presented models of Bavaria and Germany have three commodities which are modeled in such a way: wind, solar and hydro. Figure 4.2 shows the sensitivity of the cost function with respect to deviations of all three time series for a whole year.

The histogram shows how often a specific value of sensitivity occurs for all three commodities. Thin bars indicate a smaller range of costs as all graphs are scaled to the same axes. In total there are $\|V\| \cdot \|T\|_m = 16 \cdot 8760 = 140160$ values for hydro and solar and $17 \cdot 8760 = 148920$ values for wind (due to the offshore region).

The high price difference comes from the fact that it is dependent on the installed capacity of each commodity as given in equation (4.8). Hence, as especially in the southern states of

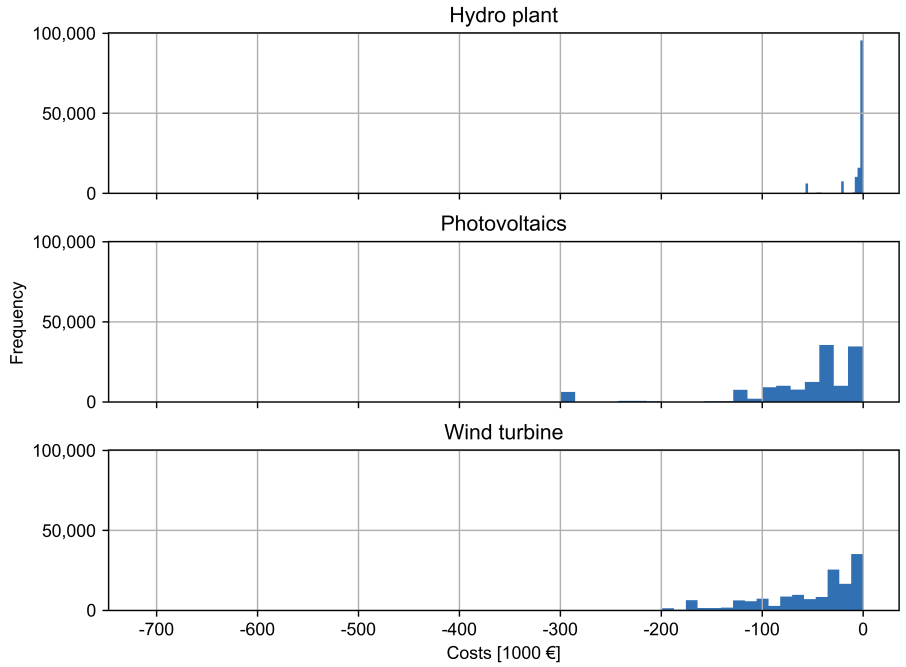


Figure 4.2: Sensitivity of cost function with respect to deviations of capacity factor for Germany

Germany the number of PV installations are particularly high (approximately 10GW in Bavaria), the sensitivity to these values is high. In case of hydro, most values lie near zero and only some are near –70,000€, also because the installed capacity of hydro power is lower than wind turbines and PV. Hence, for further analysis, hydro is not considered more closely.

Figure 4.3 shows the two subplots as from figure 4.2 for solar and wind, with a different scaling. The histograms indicate that, while the range for solar has a bigger spread and even values around –700,000€, cost sensitivity due to deviations of the wind capacity factor is of valid interest. Especially, as solar capacity factors can be seen as more predictable than wind, due to a fixed and limited availability during the day in case of solar power.

A time and region dependent plot shows how the sensitivity of the cost function with respect to the capacity factors is distributed and dependent on the installed capacity. For PV this is represented in the upper plot of figure 4.4. Similarly, this is done for wind in the lower diagram.

As Bavaria has a high amount of installed PV, the sensitivity to PV is high, especially in this region. As the installed wind power is more distributed over the northern states of Germany, the prices do not get as high as for Bavarian PV, but still reach considerable values.

It is valid to continue with taking the wind capacity factor as uncertain parameter while leaving the other input data fixed.

Before getting into the different scenarios and stochastic calculations, the virtual node or marginal prices are illustrated to indicate sensitive time intervals and regions. Figure 4.5 shows the dual variable $-\lambda_{vct}$ from equation (4.7). The dual variable in this context can be interpreted as such that one MW power demand is added to the system in one time step. From

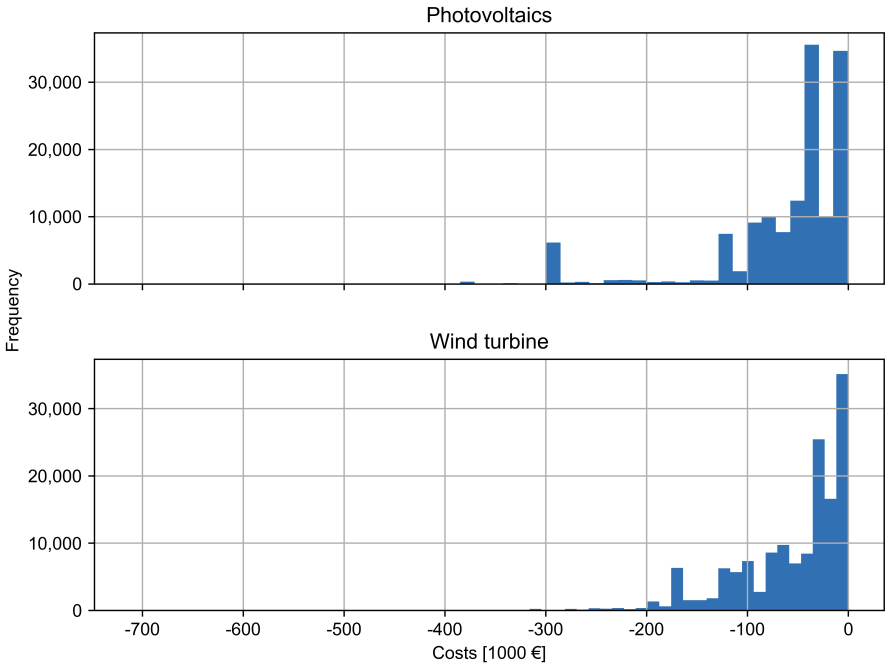


Figure 4.3: Sensitivity of cost function with respect to deviations of capacity factor of solar and wind for Germany

the upper graph in figure 4.5, it becomes obvious that especially during the winter, prices for electricity rise. Hence, the question arises, if more available renewable sources could help in these hours.

In case of solar, this is only the case, if the high price phase lies within possible sun light hours. These hours are plotted in lower plot of figure 4.5. It can be seen that during March and April, fewer hours can be eased by additional sunlight, meaning that the critical hours are within the night. There are also some points during summer and late winter, which occur during the night. However, the most critical points still lie within day time.

4.3 Case study definitions

Wozabal, Graf, and Hirschmann [122, p. 306] emphasize in their study on the impact of renewable energy sources on power markets the importance of looking at both the short term and long term production. On the one hand, short term analysis is important for operational decisions. On the other hand, investment decisions can only be driven by a long term study.

Due to this, two main studies are taken into consideration for the studies: a short term analysis running two days and a whole year optimization, for contrast noted as long term study. For the short term study, the first 12 hours are calculated as certain. The remaining 36 hours are packed in 3×4 and 3×8 hours, resulting in six subproblems per wind realization with different lengths, representing $3^6 = 729$ possible paths in a scenario tree approach. This

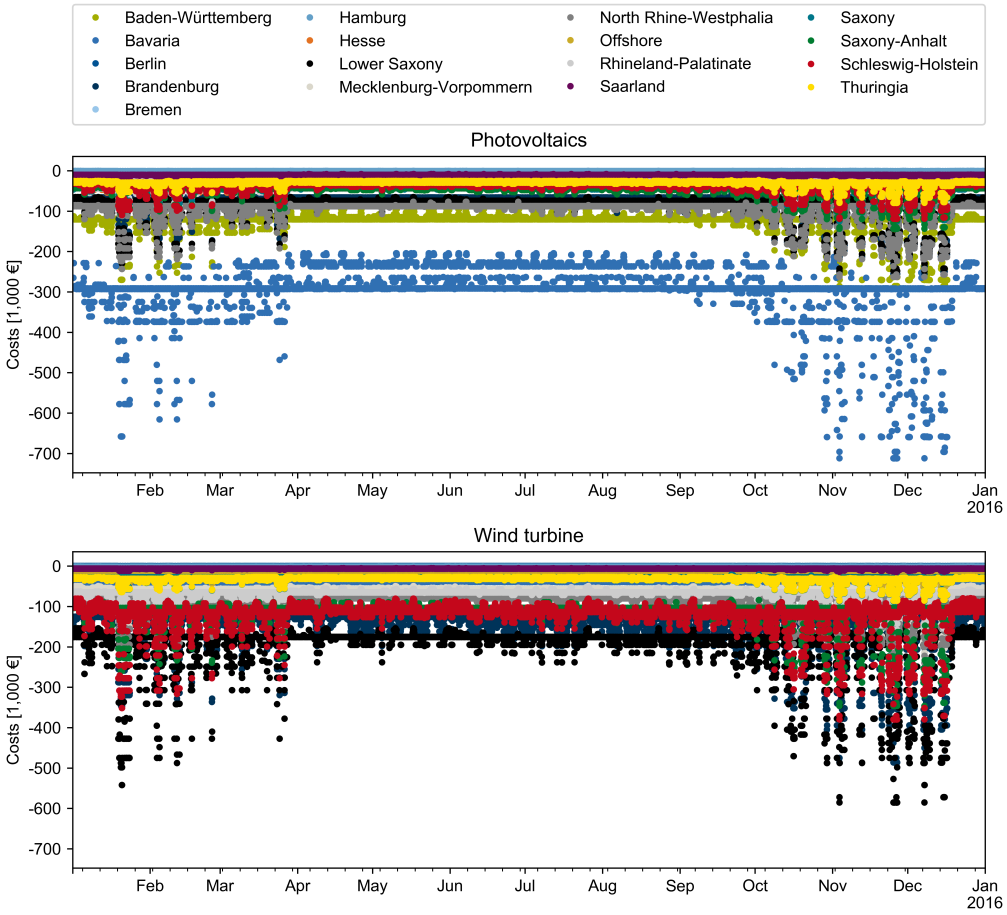


Figure 4.4: Sensitivity of cost function with respect to deviations of capacity factor of solar and wind per time step for Germany

partitioning is chosen to simulate a more detailed need for near events. Due to the short time frame, capacity expansion is not part of the optimization probe, but only dispatch. The installed capacities in the short term study are given by the *base* scenario and the *large storage* scenario, respectively.

The long term study is conducted with a seasonal uncertainty of three possible wind realizations, resulting in 15 subproblems compared to having $3^5 = 243$ possible paths in a classical stochastic approach (winter is splitted into two parts). In this case study, storage expansion but also dispatch is considered as the main point of interest.

In this context, four main scenarios are considered: the *base* case, where all input parameters stay the same and no expansion is allowed (c.f. section 4.1 Model validation). The second scenario is mainly for the dispatch/short term case, where a *large storage* is assumed, meaning that three times the storage content and power capacity are installed. For the long term case study, *storage expansion* is allowed up to three times the original storage capacity. An additional scenario for the long term study is conducted by scenario *long term dispatch*: The optimal storage investments of scenario *storage expansion* calculated by a perfect foresight approach are taken as input for a stochastic dispatch analysis. Table 4.1 summarizes the

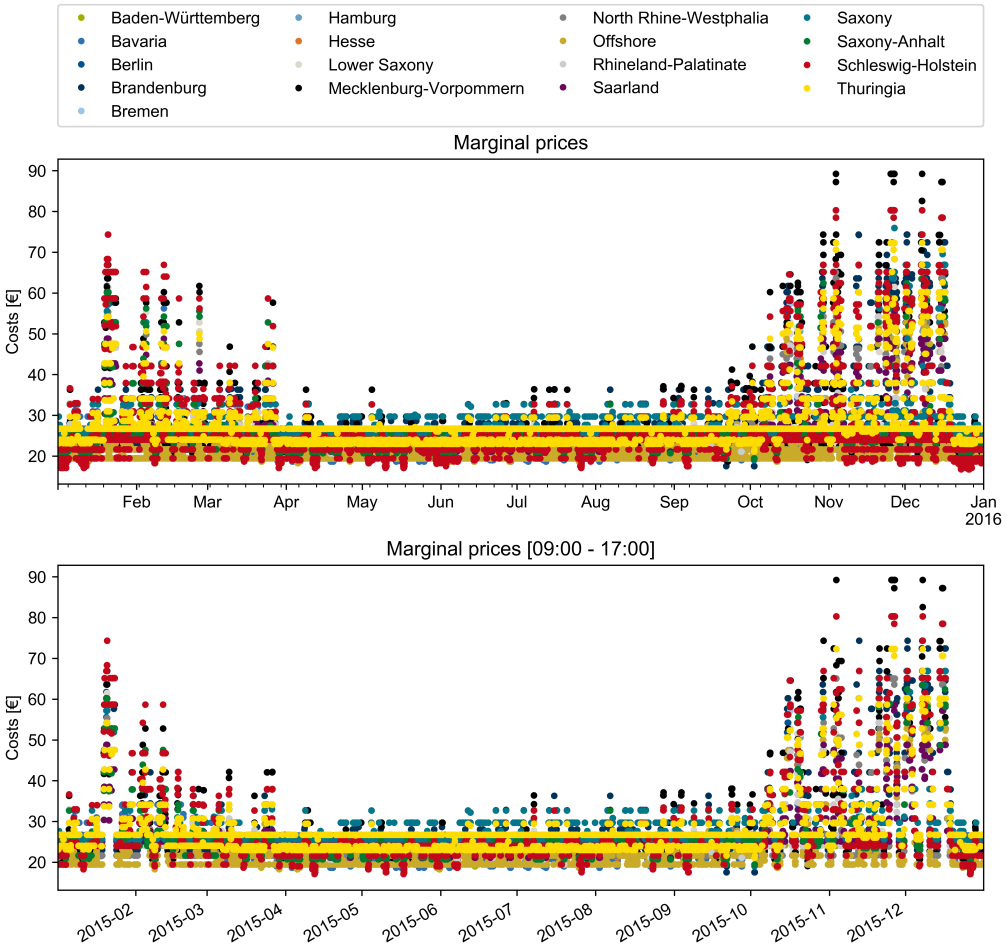


Figure 4.5: Sensitivity of cost function with respect to increasing the demand for one unit per time step for Germany for the whole year and just possible sun hours (09:00 - 17:00)

	base	large storage	storage expansion	long term dispatch
Installed storage capacity	original	3·original	original	optimal
Storage expansion	-	-	max. 3·original	-

Table 4.1: Assumptions for the scenarios

assumptions taken for the scenarios.

4.3.1 Comparison of stochastic and deterministic solution

When modeling and calculating with uncertain input parameters, the question on how to compare the gained results with a perfect foresight or deterministic approach becomes important. How can a solution be validated and compared to a perfect foresight model? Which criteria can be used to measure the usefulness of the solution?

Realization	Deviation	Value	Probability	Value
low	δ_{low}	–0.3	p_{low}	0.3
mid	δ_{mid}	0	p_{mid}	0.4
high	δ_{high}	0.3	p_{high}	0.3

Table 4.2: Wind uncertainty parameters: δ_r represents the deviation factor, p_r the probability of realization r.

Literature gives two main terms in comparing and analyzing solutions of stochastic approaches: Value of stochastic solution (VSS) and Expected value of perfect information (EVPI) [25, 150f].

VSS describes the difference between a stochastic approach and a deterministic approach using the mean of the uncertain parameter as input. An example would be to have an uncertain demand, which has to be optimally fulfilled. The stochastic approach would take several realizations for the demand and optimize over the weighted realizations. The deterministic approach would take the mean of the demand as fixed and solve the problem for this input. The difference in the objectives is then the VSS.

EVPI equals the deviation of the stochastic approach to the perfect information solution. For the demand example, this means the optimization problem has the precise demand as input. Hence, EVPI is also understood as the cost an operator would pay to achieve perfect knowledge about the uncertain parameter.

The two measures give the difference of the objective function between a deterministic and a stochastic model a value. In classic power system analysis, perfect foresight is used as input parameters instead of a mean. Hence, the second measure is the first choice for comparing both approaches. As the presented model for Germany is based on 2015 data, the problem is based on perfect foresight.

For both studies (short and long term) not only the objective function values will be compared but also the respective motivations for conducting these studies. As mentioned before, the short term study focuses more on the dispatch and how to schedule the respective power plants and storages. The long term study can give a deeper insight into a robust expansion of the country's power system.

4.3.2 Wind uncertainty modeling

For the stochastic approach input data about the uncertainty is needed. As described at section 3.3.2.2, the wind capacity factor s is changed by a deviation factor δ_r for every realization r. The case study includes three realizations: *low*, *mid* and *high* wind. The *mid* realization consists of the original 2015 time series and can be seen as base. To conduct the deviation factors for both *low* and *high* realizations, historical wind data is analyzed.

Figure 4.6 shows the deviation of wind capacity factor from 1980-2016 in comparison to the 2015 time series. The deviation is mainly close to zero and its frequency of occurrence is decreasing symmetrically in both directions. For the *low* and *high* realizations, deviations

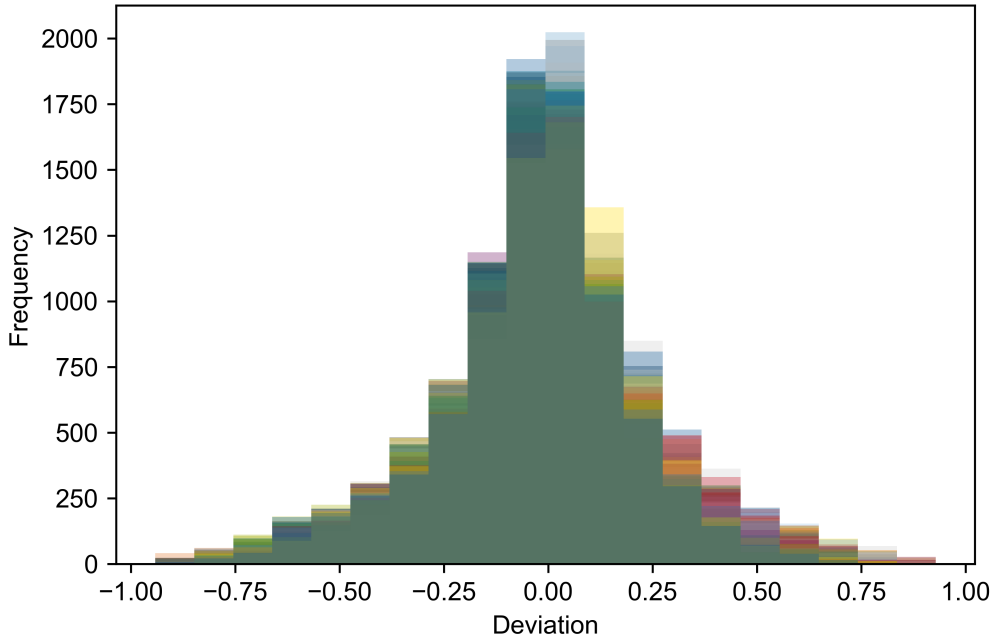


Figure 4.6: Deviation of wind time series 1980-2016 to 2015 time series

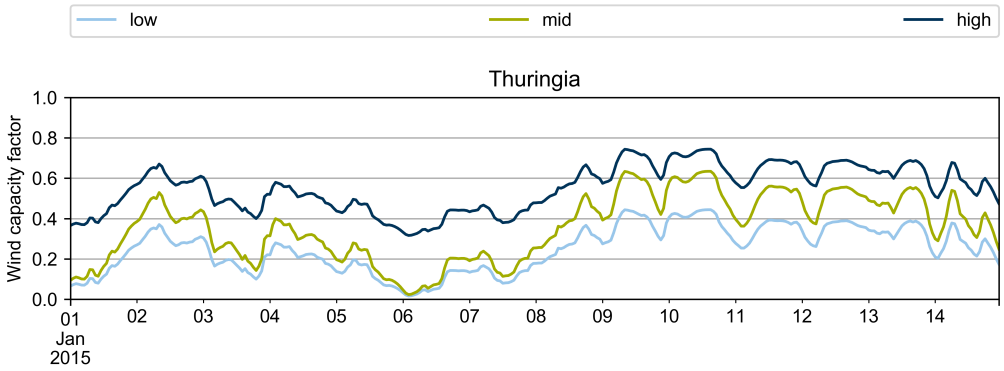


Figure 4.7: Wind capacity factor for realizations *low*, *mid* and *high* for two weeks in Thuringia

smaller than –0.1 and greater than 0.1 are considered, and their frequency analyzed. This results in the parameters, given in table 4.2.

The resulting time series for each realization is exemplary visualized in figure 4.7. The nearer the 2015 time series *mid* is to 0, the smaller the difference to the *low* realization and the bigger the difference to the *high* realization gets.

4.4 Results

For both, the short term and long term study, the presented model for Germany has been used (c.f. chapter 2). The framework *urbs* [60] is used in its original form for the deterministic approaches and in its adapted form for the stochastic approach [82] (c.f. chapter 3). As

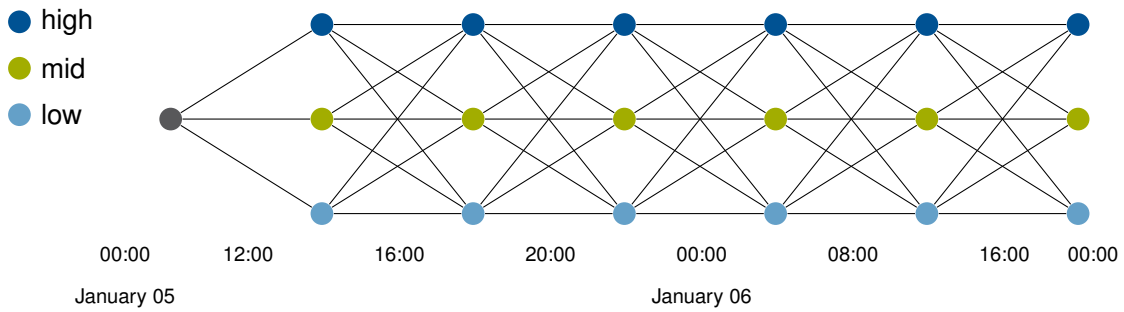


Figure 4.8: Master and subproblem structure in SDDP approach in short-term analysis with three different wind realization time series per subproblem

Scenario	Perfect foresight	SDDP	EVPI
base	14.17	14.46	0.29
large storage	14.33	14.62	0.29

Table 4.3: Costs for short term study scaled to one year in billion €

especially the long term study takes several iterations to get conclusive results, the solver gurobi [130] with an academic license is used. In comparison to the open source solver GLPK, gurobi does not only use the simplex method to optimize, but an interior point method as well. For larger problems this makes a significant difference in the solving time. The interested reader is referred to the book by Boyd and Vandenberghe [131] for an introduction into different solution algorithms for convex optimization problems like the simplex or interior point method.

4.4.1 Short term analysis

By analyzing the results of a SDDP approach, one has to be careful on what to draw as conclusion. For the short term case study, as there are no investments allowed, mainly the dispatch decision and how to operate power plants and storages with uncertain future events are of interest. Hence, only the scenarios *base* and *large storage* are taken into account for the short term analysis, as the expansion decision is not as interesting for the short term case as it is for the long term study.

As described, the short term study includes 48 hours of cost-optimal dispatch. The first twelve hours are taken as certain. The subproblems start at 12:00, 16:00, 20:00 on January 5th (first Monday of the year) and at 00:00, 08:00 and 16:00 on January 6th. The separation of the respective hours in the subproblems is visualized in figure 4.8. The two days have been optimized with a perfect foresight model and a varying wind time series with the three realizations *low*, *mid* and *high*. All subproblems overlap in one hour in the beginning and the end. These overlapping time steps are needed for exchanging the information between the subproblems and creating cuts out of it. While the last time step of each problem is optimized, the first time step is fixed to the value of the previous problem. If a subproblem wants its predecessor to change, e.g. the state of charge, the information is transported by the cuts.

The resulting objective values are weighted in such a way to get the price for a whole year.

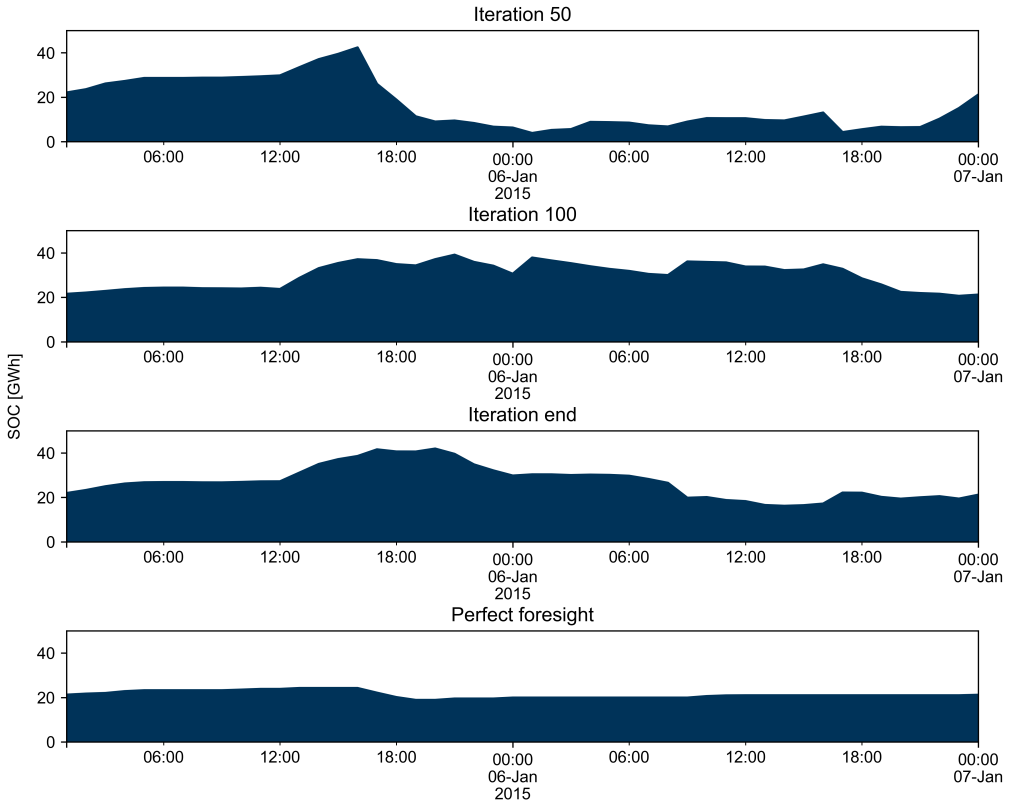


Figure 4.9: Storage content for two weeks over iterations for realization *mid* in scenario *base*

This weighting is a standard feature of *urbs* to ease the comparison between different time horizons. Table 4.3 summarizes the costs for each approach and their difference (EVPI). As expected, the stochastic solution is more expensive than the perfect foresight approach. The differences lie within the different dispatch. The *large storage* scenario has a slightly smaller EVPI than the *base* case as the increased storage capacity leads to a higher flexibility.

For the perfect foresight model, the costs for the *large storage* scenario rise due to higher fix costs for the storage. As the large storage is not completely needed during the modeled hours, the larger storage only leads to higher costs. This is similar in the stochastic approach.

For a better comprehension of the SDDP approach and how it comes to its solution, the dispatch and state of charge (SOC) for the two days are shown for different iterations. An exemplary result for the SOC is shown in figure 4.9. For all five subproblems, the chosen realization is *mid*. The jumps in iteration 10 between 12:00 and 00:00 and at 08:00 on January 6th, indicate a relaxed storage content constraint (λ). While the iterations proceed, the storage content is shifted to the time between 12:00 on January 5th and 06:00 at January 6th. As it is in the beginning not rewarding for the former subproblems to store energy, the information of the last subproblems, that costs can be reduced by a higher storage content, needs to be transported to the previous subproblems. In an SDDP algorithm, the information is transported via the cuts from back to forth (backward iteration, c.f. chapter 3).

Considering the dispatch, the difference between the stochastic and the perfect foresight approach is less visible, as illustrated in figure 4.10. During the night (00:00 - 06:00 January

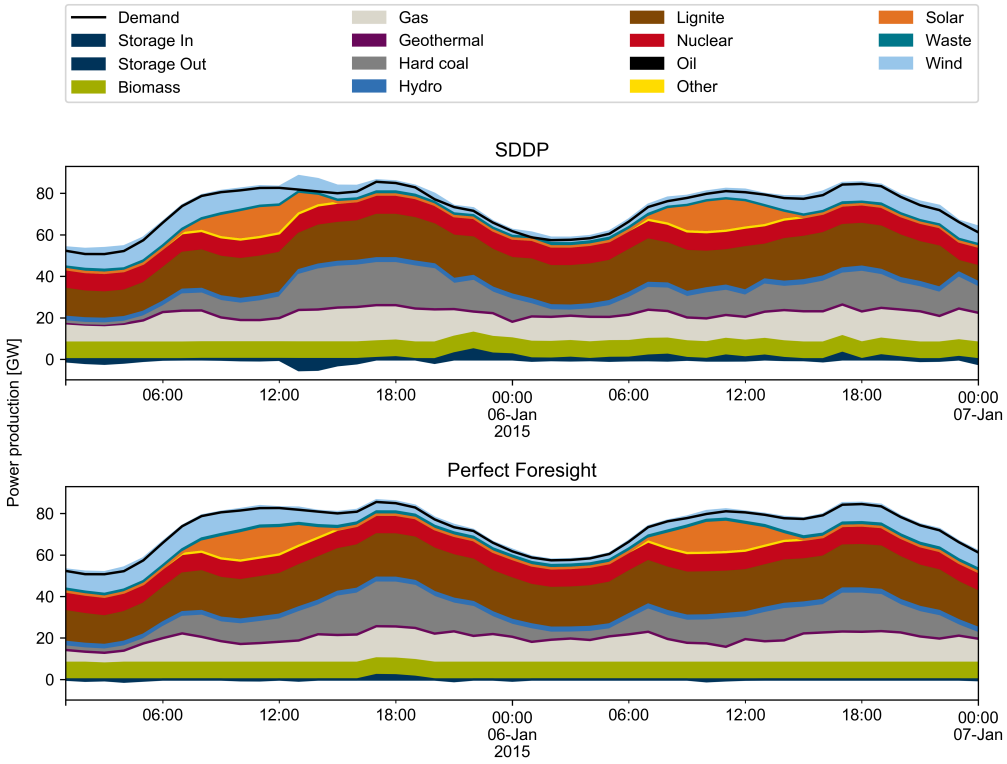


Figure 4.10: Dispatch for short term case study for realization *mid* in scenario *base*

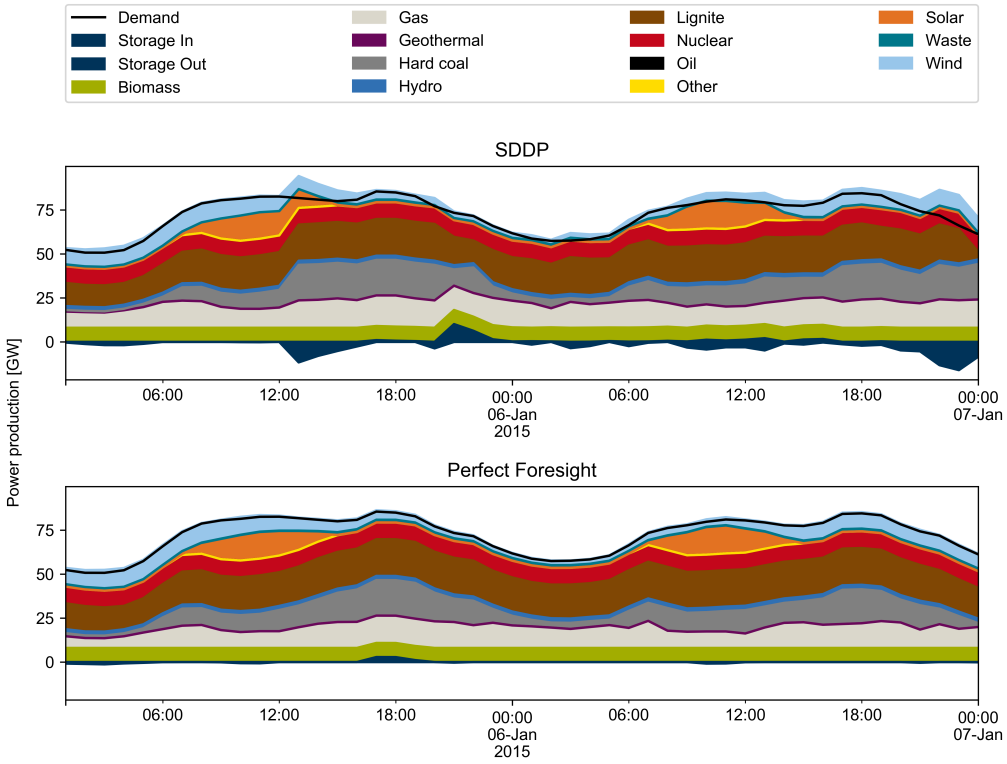


Figure 4.11: Dispatch for short term case study for realization *mid* in scenario *large storage*

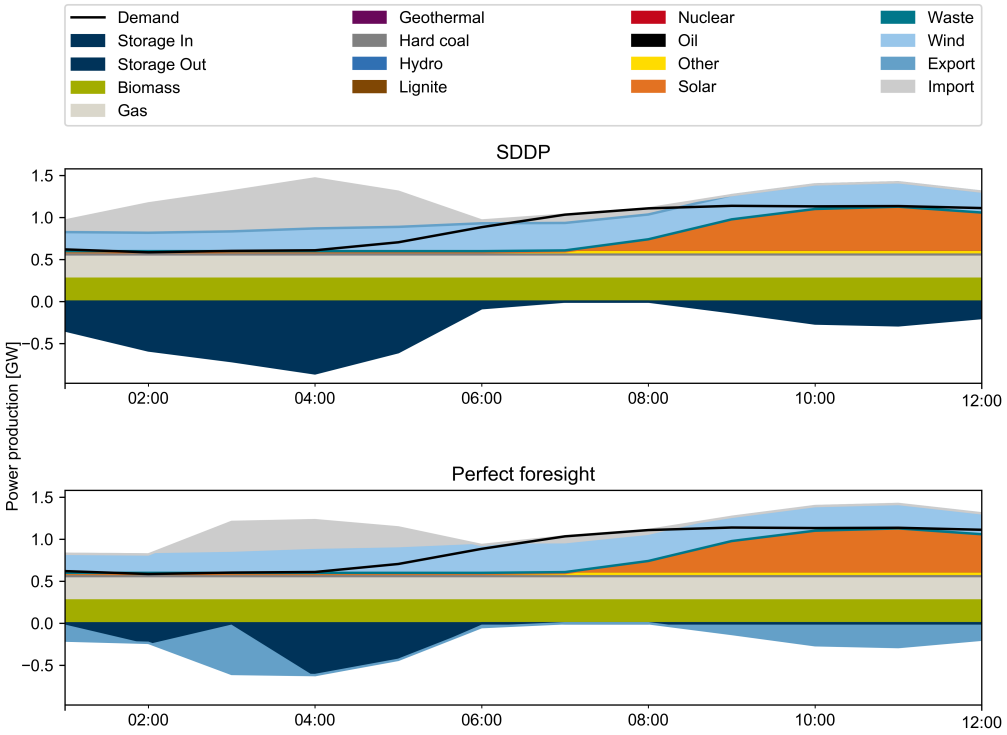


Figure 4.12: Dispatch for short term case study in scenario *base* in Thuringia

5th), the storage is slightly more filled than in the perfect foresight case. For gaining more flexibility, more power is stored in the afternoon of January 5th. In general, one can see more usage of the storage in case of the stochastic approach. This is also visible in the *large storage* scenario in figure 4.11.

By looking more closely on the first 12 hours, which are fixed in both cases, and on a regional level, the differences between both approaches become more visible. Figure 4.12 shows the dispatch in the morning hours of January 5th. It can be seen that the import, export and the loading of the storage are different in both approaches. The stochastic solution stores more in the early morning, especially from imports.

The regional differences are also clearly visible in the CO_2 emissions per region as presented in figure 4.13. It can be seen that especially in Brandenburg, North Rhine-Westphalia and Saxony, the CO_2 emissions in the first twelve hours diverge from the perfect foresight case. The budget of emitting CO_2 has to be saved to be able to fulfill the requirements of the power system. A similar behavior is also visible in the *large storage* scenario (c.f. figure B.1).

As a general observation, the classic merit order based on price is not the main driver in case of SDDP. An uncertain future is resulting in a more cautious dispatch of power plants, leading to more generation of gas compared to the perfect foresight case.

4.4.2 Long term analysis

For the long term case study, especially the *storage expansion* and *long term dispatch* scenarios are of interest. The *storage expansion* scenario will be analyzed with four different

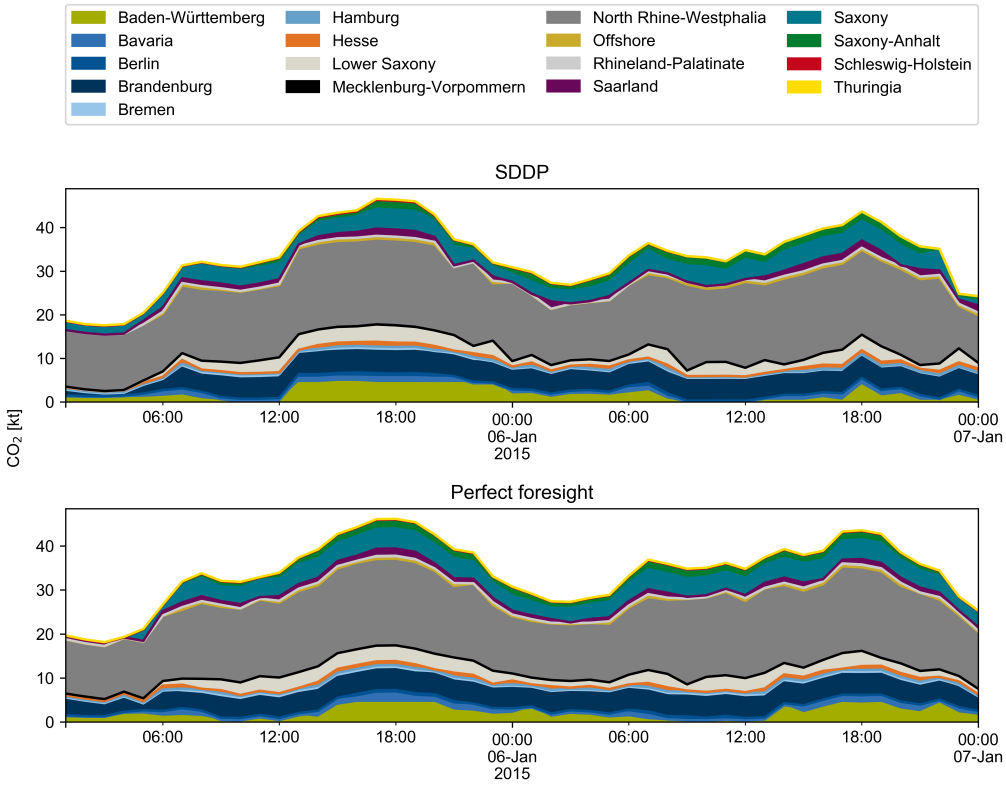


Figure 4.13: CO_2 emissions for short term case study for realization *mid* in scenario *base*

	Perfect foresight	Mean time series	DDP	SDDP
base	•	•	-	-
large storage	•	•	-	-
storage expansion	•	•	•	•
long term dispatch	•	-	-	•

Table 4.4: Used optimization approaches for the long term scenarios

Scenario	Perfect foresight
base	13.750
large storage	13.892
storage expansion	13.749

Table 4.5: Costs for long term study with perfect foresight approach in billion €

approaches: perfect foresight, with an averaged wind time series, with Dual Dynamic Programming (DDP) and finally with SDDP.

In case of perfect foresight, not only the expansion and dispatch scenario, but also *base*, which has already been used for the validation, and *large storage* are taken into account. The idea is to provide insights with a deterministic approach, before comparing it with the stochastic approach.

The deterministic optimization with an averaged wind time series comes from the concept of VSS. So, instead of using the wind time series for one specific year (like 2015), historic data is used to get one averaged time series per modeled region. This will show the dependence of expansion on the underlying time series.

Similar to chapter 3, where SDDP is derived from DDP, scenario *storage expansion* is optimized with a DDP approach. Again, the 2015 wind time series is used, hence, the input is the same as for the perfect foresight case. The main insight provided by this approach lies within the development of the solution, as it can be seen how the information between the subproblems is exchanged.

Finally, scenario *storage expansion* and *long term dispatch* are optimized by the SDDP approach to be more independent from single time series and specific weather phenomena. For an overview, all three (perfect foresight, average mean and stochastic) approaches are compared with respect to installed capacities as well as costs.

Table 4.4 summarizes the optimization approaches used to optimize the different scenarios.

4.4.2.1 Perfect foresight approach

Firstly, the scenarios *base*, *large storage* and *storage expansion* are evaluated with a yearly optimization assuming perfect foresight without uncertainty is drawn to point out some details.

The resulting costs for all three scenarios are stated in table 4.5. Similar to the short term analysis, the *large storage* scenario is the most expensive one because of the resulting fix costs. As the *storage expansion* scenario allows more freedom than the *base* scenario (by allowing cost beneficial investments), this scenario is the cheapest.

Figures 4.14 and 4.15 show the dispatch of one week in summer and winter, respectively. It can be seen that in summer next to a lower demand (black curve), PV has a considerable amount of dispatch. The storage usage is low as the installed capacities are enough to buffer even low wind/PV times. This is different in winter, where the storage is used more often. In summer, the difference between the three scenarios *base*, *large storage* and *storage expansion* is not as visible as during winter. Especially in the *large storage* scenario, storage use is clearly visible.

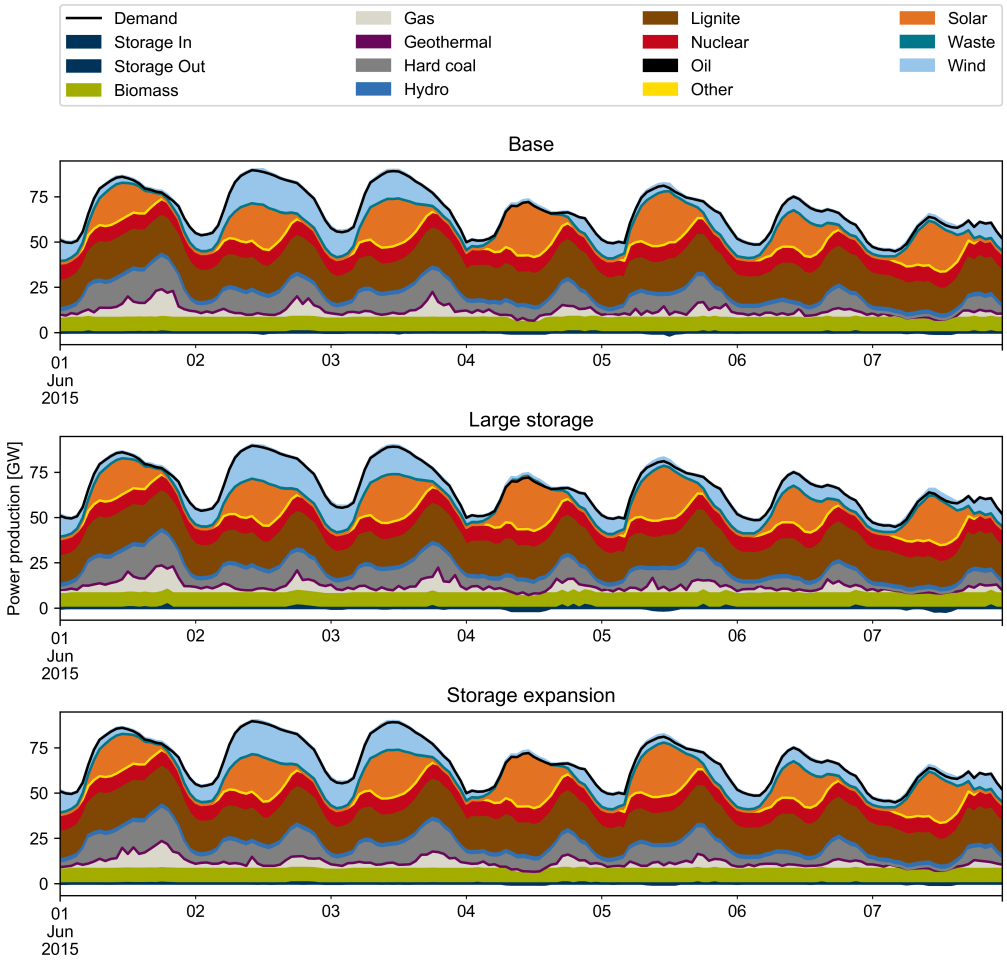


Figure 4.14: Dispatch of one week in June for scenario *base*, *large storage* and *storage expansion*

At the evening of the 7th of December, there is not much wind production (and of course no solar radiation) while the demand is relatively high. This day is one of the points which lead to a high price as shown in figure 4.16. It shows the marginal prices for all three scenarios for one year. Again, it is clearly visible that especially during the winter time, where demand is higher and renewable availability is more sparse, that these effects lead to higher prices for electricity.

Compared to the *base* case, the *large storage* scenario has in general lower prices as the storages help easen times of high demand. In the *storage expansion* case, the prices also go down a bit, but also can be higher. This becomes evident, when the differences to the *base* case are plotted separately. Figure 4.17 represents the differences during winter months at the end of the year. The higher costs in the *storage expansion* scenario compared to *base* (negative values) come from the fact that the annualized investment costs and higher fixed costs have to be taken into account.

The effect of how the storage size affects the marignal prices per season is clearly visible in figure 4.18. While all curves get damped by a large storage content, autumn and winter have the highest peaks in the evening. The higher peak for autumn comes from the fact that

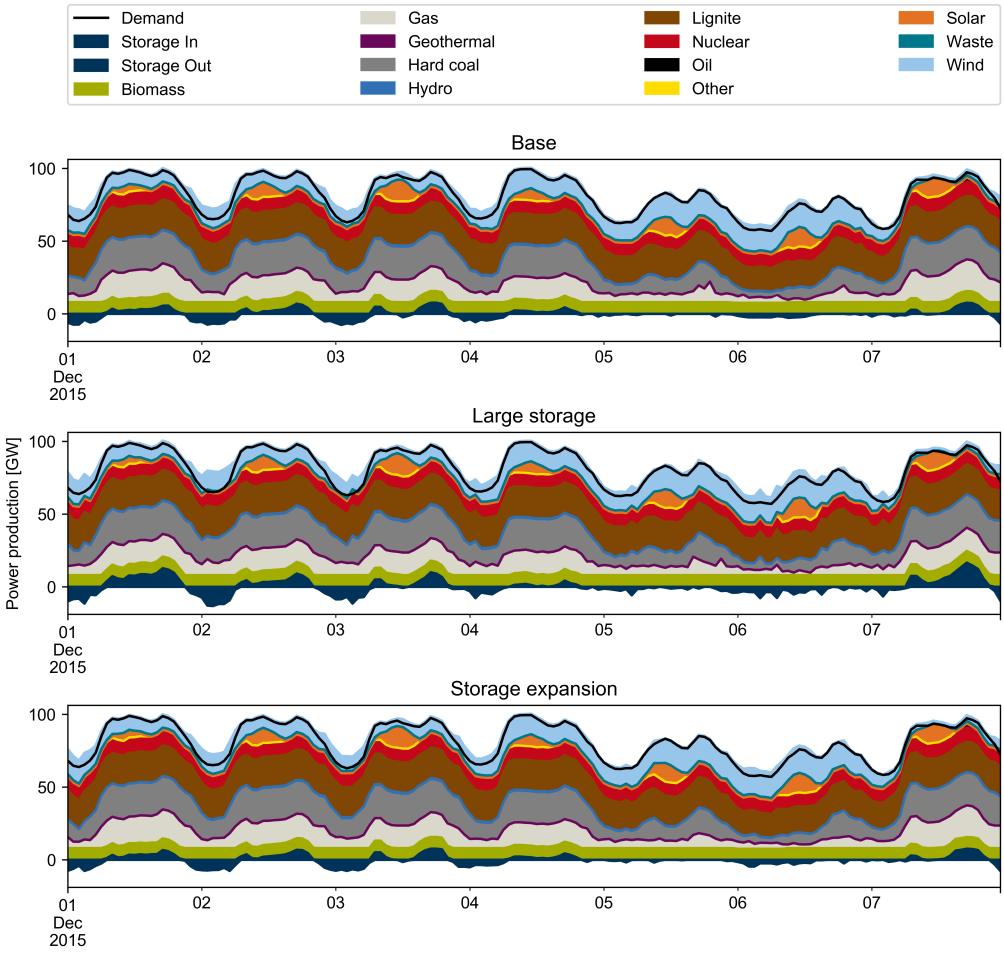


Figure 4.15: Dispatch of one week in December for scenario *base*, *large storage* and *storage expansion*

November is counted as autumn (next to September and October). Winter is represented by the months January, February and December. Another interesting effect which can be drawn from the curves is the visibility of installed PV. Bavaria, which has by far the most installed PV capacities, has especially in summer a higher decrease about noon.

Incentives for building storage comes from time steps which have high marginal prices as discussed. But storage expansion is not only favored in high price areas, it is also interesting for hours of overproduction. These hours can be illustrated by the residual load, which is given by the demand minus all renewable production. The residual load curve per state is visualized in figure 4.19. The hours, which are below the zero line, indicate overproduction. States like Saxony-Anhalt or Mecklenburg-Vorpommern have more than 160 hours of overproduction. For a clearer representation of these hours, the interested reader is referred to the detailed plot B.5 in the appendix. As transmission plays a vital role in these considerations, not all investment decisions (for storages) can be anticipated upfront. However, a tendency towards investing in regions with overproduction is visible.

From table 4.6 it can be seen that new capacity is especially installed in Saxony. Looking

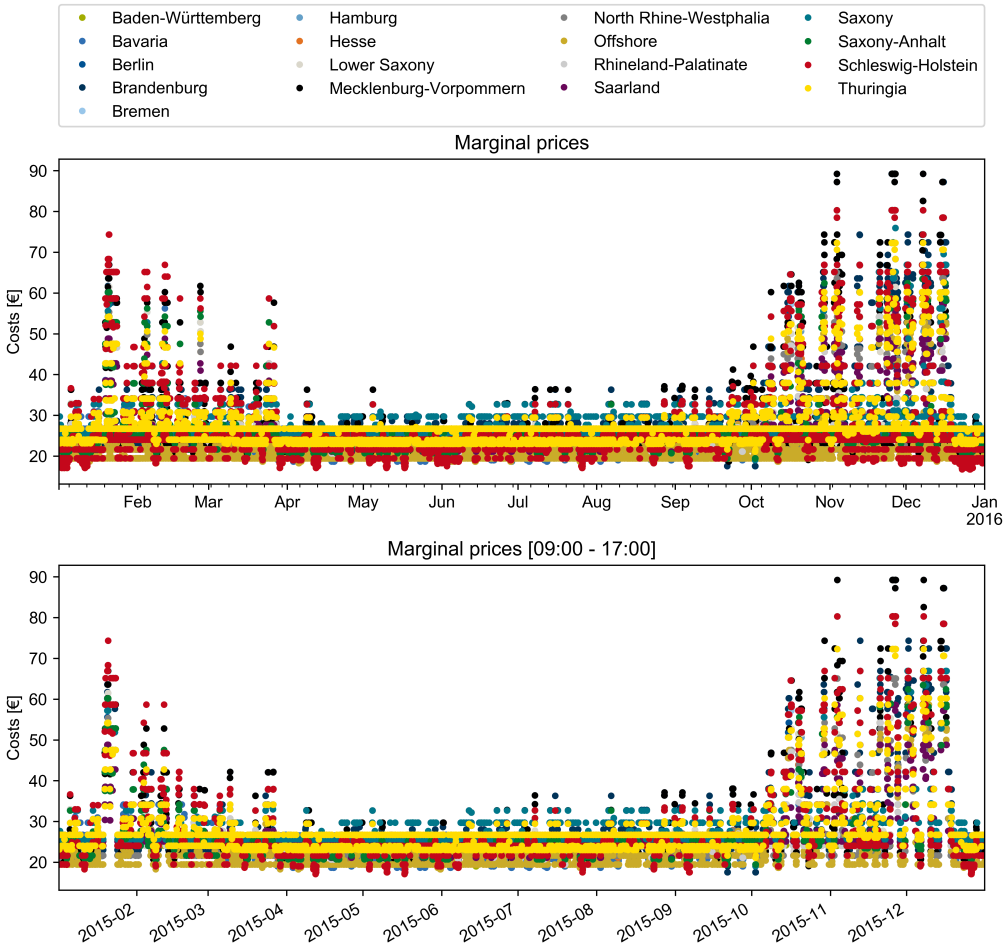


Figure 4.16: Marginal prices for one year for scenario *base*, *large storage* and *storage expansion*

State	Perfect foresight
Lower Saxony	820
North Rhine-Westphalia	582
Rhineland-Palatinate	98
Saxony	2.335
Saxony-Anhalt	115
Schlewsig-Holstein	829

Table 4.6: Newly installed capacity of storages in [MWh] for long term study with perfect foresight approach

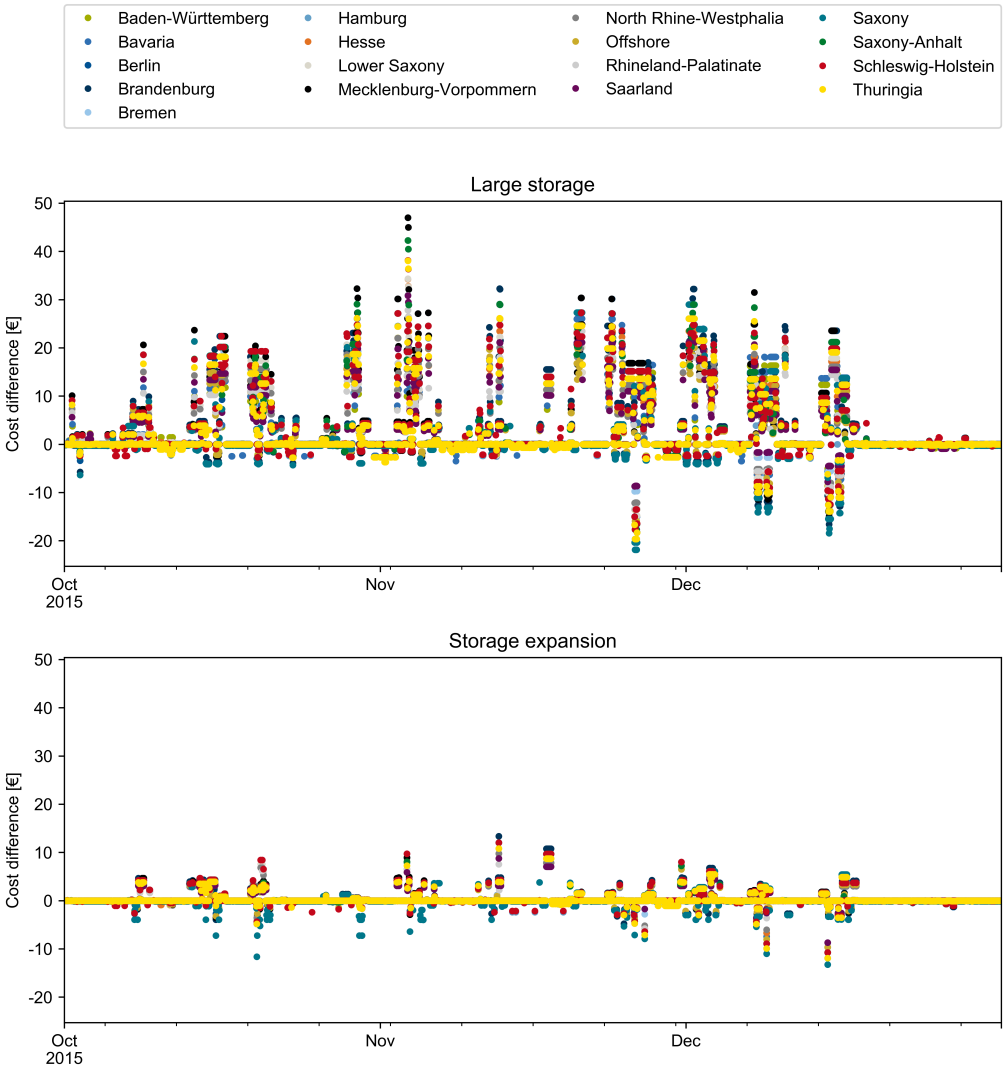


Figure 4.17: Marginal price differences between *base* case and the two other scenarios in the winter months

back at figure 4.17, mainly Saxony has higher prices than in the base case. As Thuringia is a neighboring state of Saxony, prices stay high there as well. However, in general over the whole year, marginal prices reduce as the capacity expansion restriction gets relaxed.

Coming back to the dispatch plot 4.15, upcoming high prices lead to higher production *before* this high price event, to ease this price difference. This can be seen clearly in the time span from the 5th-6th December, in which conventional plants like lignite, coal and gas are still producing power to store even if they are not needed anymore. However, the high price of the 7th of December does not lead to such a high price that the storages are fully expanded until their limit, which can be seen by comparing the storage usage of both scenarios *large storage* and *storage expansion*.

The question arises whether and how these expansion decisions change with the following approaches.

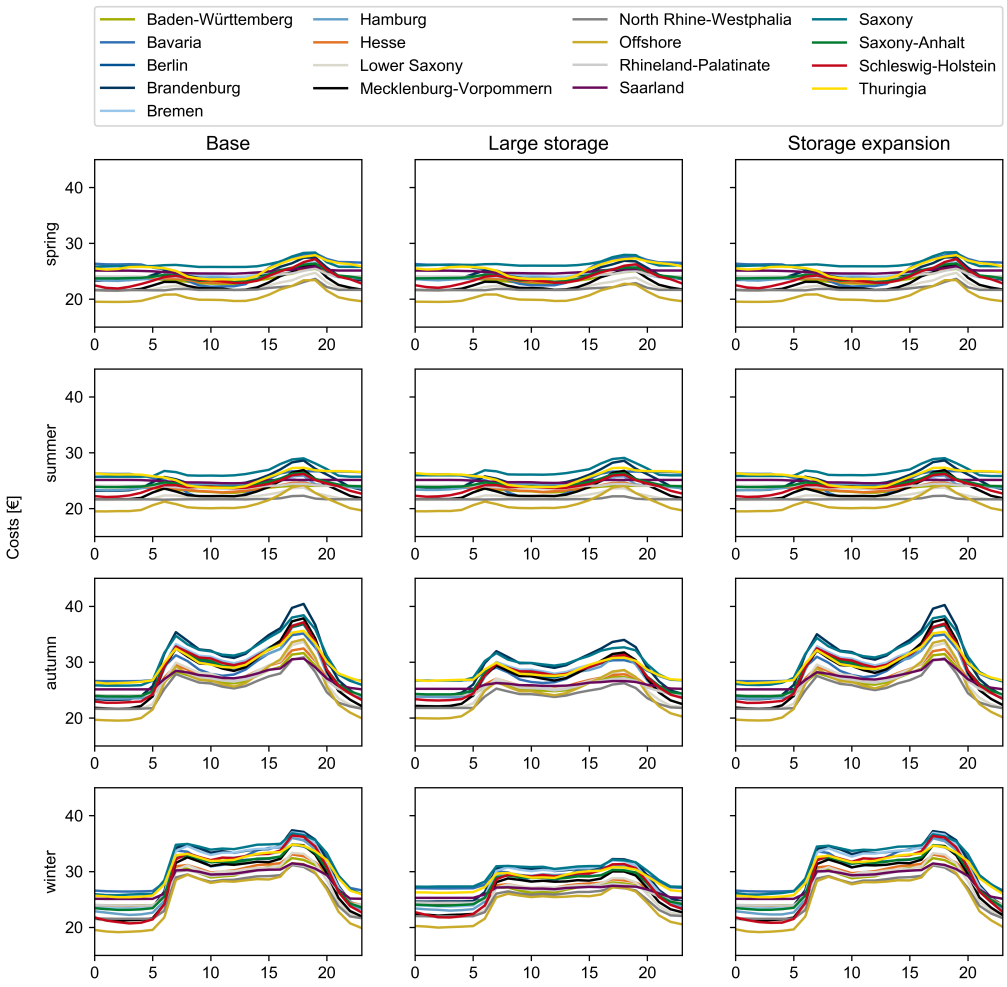


Figure 4.18: Average marginal prices for one day per season for scenario *base*, *large storage* and *storage expansion*

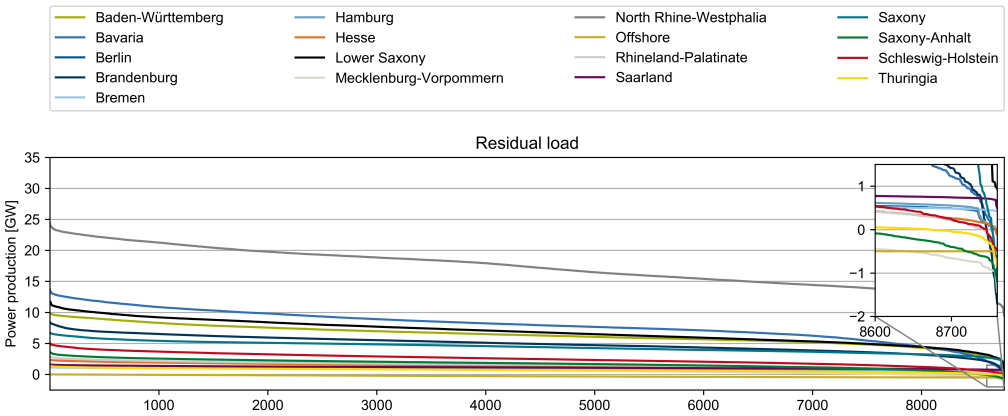


Figure 4.19: Residual load for perfect foresight approach

Scenario	Perfect foresight	Mean
base	13.750	13.821
large storage	13.89	13.964
storage expansion	13.749	13.820

Table 4.7: Costs for long term study with perfect foresight and mean wind time series approach in billion €

State	Perfect foresight	Mean
Baden-Württemberg	0	97
Bavaria	0	121
Lower Saxony	820	644
North Rhine-Westphalia	582	582
Rhineland-Palatinate	98	0
Saxony	2335	2007
Saxony-Anhalt	115	51
Schleswig-Holstein	829	353
Σ	4780	3855

Table 4.8: Newly installed capacity of storages in [MWh] for long term study with perfect foresight and mean wind time series approach

4.4.2.2 Mean wind time series

For an additional comparison, the mean of the wind time series from 2000-2017 (due to regional availability on renewables.ninja [77]) is taken per state and used as input. The resulting costs are stated in table 4.7. For all three scenarios, the costs for the energy system in one year are higher than in the perfect foresight case. For the mean wind time series approach, the *large storage* scenarios is again the most expensive. The difference between scenarios *storage expansion* and *base* is again smaller.

Figure 4.20 represents the marginal prices for a whole year simulation with the same input data as before but with the mean wind time series. Compared with the 2015 data in figure 4.16, especially the prices at the beginning and at the end of the year are lower. It is also worth noting that instead of Mecklenburg-Vorpommern (perfect foresight) Brandenburg has a high variety in prices. Two days at the end of September and beginning of October (for both time series) are lower than the average in Brandenburg, but the highest prices are also located in Brandenburg for the mean wind time series.

An extensive analysis with the residual load can be conducted by working with the figures B.6 and B.7 in the appendix. In general, the mean wind time series has less overproduction than the perfect foresight time series. This is supported by the resulting installed capacities, presented in table 4.8. While the mean wind time series approach has a broader distribution of newly installed capacities, the perfect foresight approach has a nearly 1 GWh more installed capacities.

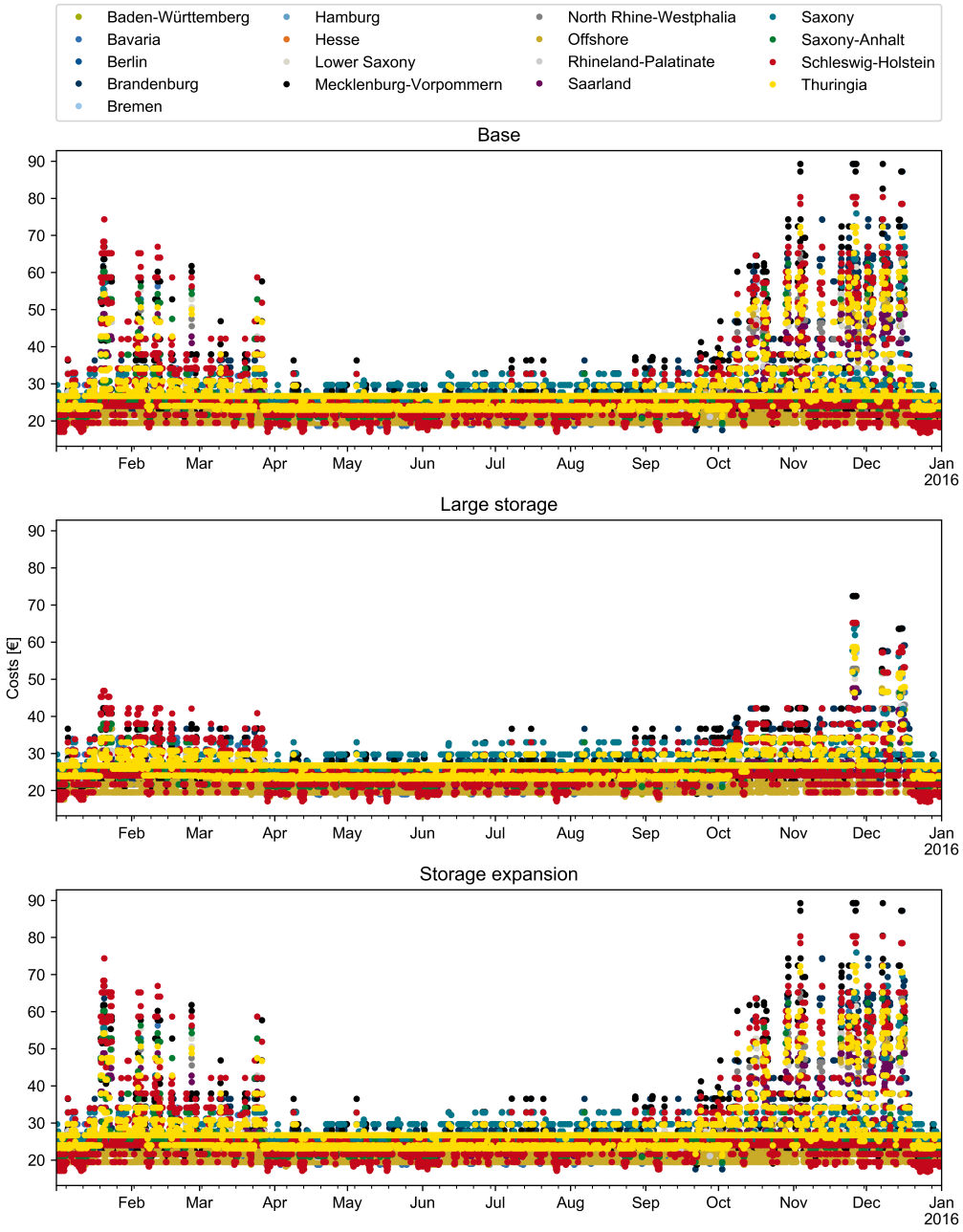


Figure 4.20: Marginal prices for one year for scenario *base*, *large storage* and *storage expansion* for mean wind time series

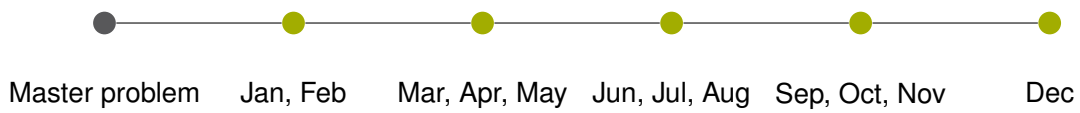


Figure 4.21: Master and subproblem structure in DDP approach: Green circles indicate subproblems with the 2015 wind time series

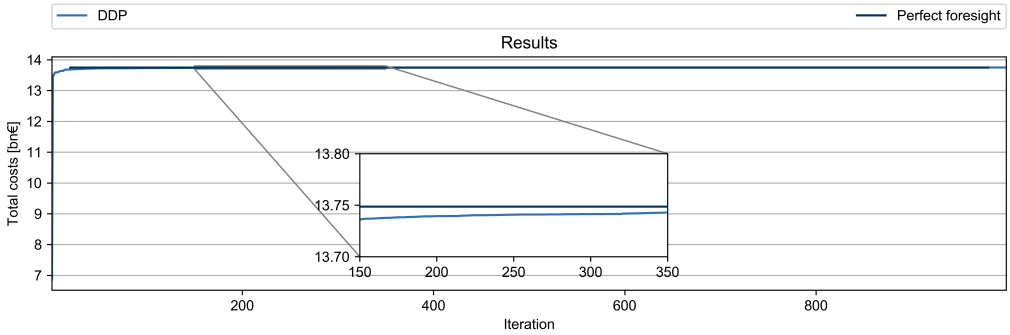


Figure 4.22: Resulting costs over iteration for DDP approach

4.4.2.3 DDP analysis

Before finally analyzing the problem with a stochastic approach, the general working of a DDP approach is explained more closely. The problem is separated into one master problem, which is just responsible for investment decisions, and several subproblems, which include the power production. The year is divided into four seasons, but as winter is both in the beginning and the end of the year, this results in five subproblems with different lengths, as indicated by figure 4.21. The wind time series is again set to the 2015 time series, as it is done in the perfect foresight approach.

For this specific model the variables, which have to be given to the next subproblem, are the storage content and CO_2 emissions. Especially the CO_2 emissions lead to a long run of the DDP approach as the global as well as the state bounds are often violated.

As the input is the same as for the perfect foresight approach, the resulting costs are converging to the same costs, which is visualized in figure 4.22. It can be seen, that the objective costs converges fast to the perfect foresight optimum. After 150 iterations, there is only a gap of 12.67 M€, which equals 0.09 % of the total costs of the perfect foresight approach.

The costs after the first iteration equal 6.86 bn€, which are exactly the fixed costs of the power system for one year. This result is due to the fact, that the master problem does not include dispatch in the long term study and hence, has not yet any information about how much the costs for fulfilling the demand will be. The cut for the master problem after the first iteration of the algorithm is given by:

$$\alpha_0 \geq 6.64 \cdot 10^9.$$

Hence, the first approximation of the future costs for the power system is given by a hyperplane without slope but an intercept of $6.64 \cdot 10^9$. The future costs increase over the iterations as more and more information is collected by calculating the paths forward and backward. While

the cuts in the first iteration only contain a bound for the future costs α_i of the i-th stage, cuts in later stages contain information about the storage content (e_sto_con) or already emitted CO_2 (e_co_stock_state) as well, adding hyperplanes with slopes. As an example, the second cut of the master problem is given by:

$$0.06\alpha_0 \geq 0.36 \cdot 10^9 - \sum_{v \in V} 0.47 \epsilon^{\text{con}}_{vs0} \quad V = \{\text{Baden-Württemberg, Thuringia}\}$$
$$s \in S = \{\text{Pumped storage}\}$$

The storage content $\sigma^{\text{con}}_{vs0}$ for the master problem has an indirect impact on the needed expansion, as the storage is initialized by a certain state of charge. The cut shows that enlarging the storage and hence the initial storage content decreases the future costs.

After some iterations where only cuts limiting the future costs are created, the capacity κ^{c}_{vs} itself is part of a cut:

$$0.01\alpha_0 \geq 91 \cdot 10^6 - 0.34\epsilon^{\text{con}}_{v_1 s0} - 0.37\epsilon^{\text{con}}_{v_2 s0} - 0.12\kappa^{\text{c}}_{v_1 s} - 0.16\kappa^{\text{c}}_{v_2 s}$$
$$v_1 : \text{Baden-Württemberg}, v_2 : \text{Thuringia}, s : \text{Pumped storage}$$

If the costs for investing in more storage (including resulting fixed costs) in Baden-Württemberg and Thuringia are now less than what can be saved, storage capacity will be increased there. It is interesting to see that the capacities of the storages and Baden-Württemberg and Thuringia are the first ones included in a cut, while on the other hand capacity in Schleswig-Holstein and Saxony are increased first. Both regions Baden-Württemberg and Thuringia are the first ones where the relaxation of the storage content constraint by λ is completely used, hence they appear first in the cut.

In case of subproblems a cut containing the storage content $\epsilon^{\text{con}}_{vst}$ has a different interpretation: As subproblems cannot increase storage capacity, the storage content they have at the end of their time span depends only on the power production. Hence, instead of increasing storage capacity, the subproblems might produce more power, if the costs are lower than what can be saved by it in the future. The fifth cut for the first subproblem (consisting of January and February) involves the CO_2 emissions:

$$0.33\alpha_1 \geq 1.88 \cdot 10^9 + \sum_{v \in V} 0.86 \rho^{\text{state}}_{vc1417} \quad V = \{\text{North Rhine-Westphalia}\}$$
$$c \in C = \{\text{CO}_2\}$$

If the total emitted CO_2 emissions after 1417 time steps (hours of January and February) $\rho^{\text{state}}_{vc1417}$ rise in North Rhine-Westphalia, the future costs (α_1) will increase as well. Other cuts include not one variable type at a time, but several and especially with different regions involved. With these cuts, more knowledge about the dependencies of costs with respect to emissions, storages and so on could be drawn in further analysis.

While the costs are fast converging to the optimum, the optimal expansion of pumped storages does not show a clear convergence at first, as visualized in figure 4.23. Due to the high amount of data, only every 20th iteration is saved for further analysis. As only states which already have a storage are allowed to expand at all, only some states are visible in the figure. Saxony, which has the most newly installed capacities in both former analysis, has the

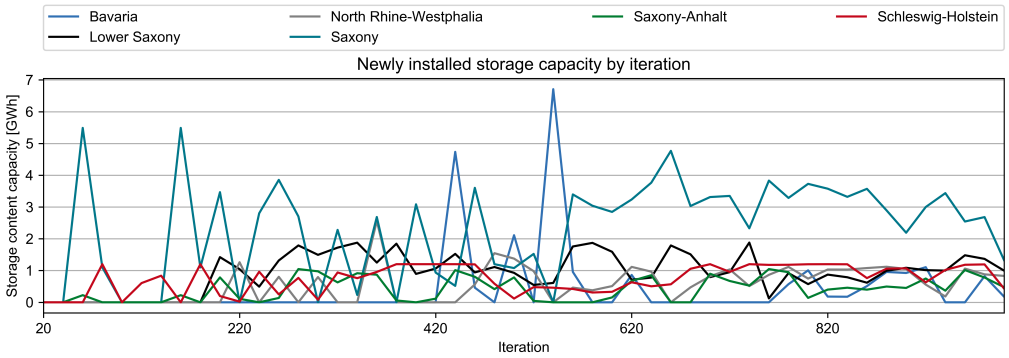


Figure 4.23: Newly installed storage capacity over iteration for DDP approach

first peak. The states Baden-Württemberg, Hesse, Rhineland-Palatinate and Thuringia are not investing at all even if they are allowed to. Baden-Württemberg, Hesse and Thuringia are also not considered in the perfect foresight approach. Bavaria, however, which is not expanded in the perfect foresight approach, has a high peak in expansion after several iterations. This might be caused by the high share of solar power and the high sensitivities due to solar power (c.f. figure 4.4).

One conclusion from both figures lies within the sensitivity of the model on expansion capacities (and hence the integration of renewables): While the total costs of the system are not sensitive to the storage expansion, the storage expansion itself is quite sensitive to the underlying time series of wind which is determined by the chosen path. The first part is evident from comparing the asymptotic behavior of the costs (figure 4.22) in comparison to the unsteady nature of the newly installed storage capacity (figure 4.23).

The fluctuating expansion of storage also hints at different expansion strategies which lead to similar costs, meaning that there are many points in the solution space having similar costs as the optimal solution.

The main problem of both the DDP and SDDP approach lies within the exchange of information. As a high investment in storage capacity needs more power generation near the end of the year to refill (as the storage has to be filled with the same amount as it started with), higher costs in the late winter months have to give their feedback via cuts through all previous subproblems until the information arrives at the master problem.

Another problem lies within the numerical issues. As the input file for the *urbs* framework contains a high range of parameters (e.g. millions for investment, but only cents for variable costs), the solver can face numerical problems. While the iterations proceed, these issues can accumulate and hence, affect the convergence of the algorithm.

4.4.2.4 SDDP analysis

The underlying structure of the SDDP approach is similar to that of DDP. The year 2015 is divided into five stages with three realizations for the wind time series per stage. The master problem is again responsible for expansion of storage capacity. The resulting structure is depicted in figure 4.24. As indicated, three different wind time series realizations are taken into account as described in section 4.3.2. At first, the scenario with storage expansion is

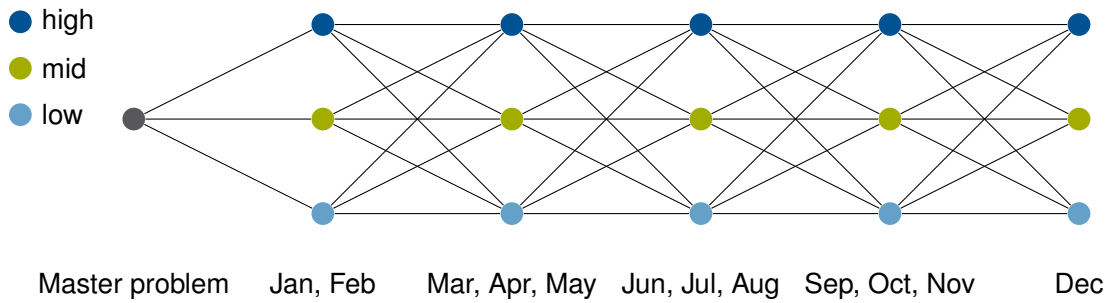


Figure 4.24: Master and subproblem structure in SDDP approach with three different wind realization time series per subproblem

Scenario	Perfect foresight	Mean	SDDP	EVPI	VSS
storage expansion	13.749	13.820	14.235	0.485	0.414

Table 4.9: Costs for long term expansion study with perfect foresight, mean wind time series and SDDP approach in billion €

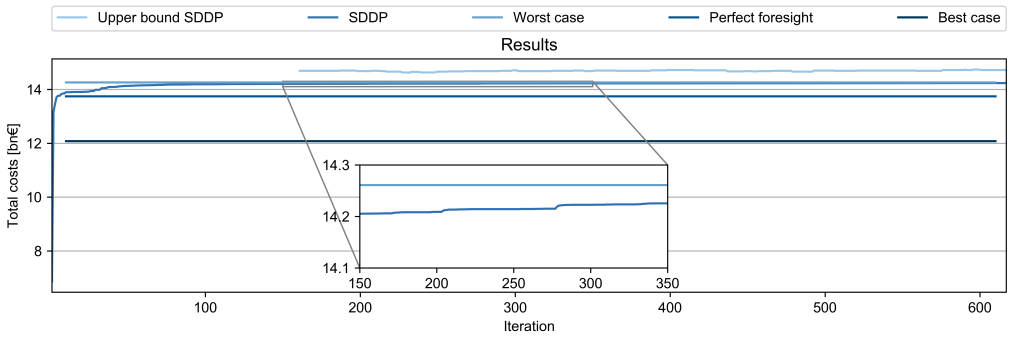


Figure 4.25: Resulting costs over iteration for SDDP approach

elaborated. Afterwards, the long term dispatch scenario is analyzed with the optimal storage expansion from the perfect foresight case taken as input.

Scenario storage expansion With an uncertain wind time series, the costs for fulfilling a constant energy demand increase, c.f. table 4.9. The system has to overcome disadvantageous wind constellations and still be able to fulfill CO_2 targets. The costs rise around 3.5 % compared to the perfect foresight approach with a stochastic approach. The difference between perfect foresight and SDDP is given by the EVPI. The VSS expresses the difference between the approach with a mean wind time series and SDDP. The rise of the costs compared to the mean wind time series are with 3 % also higher.

Looking at the convergence of costs over iterations (figure 4.25), one can observe similar behavior as for the DDP case: At first, the costs are increasing fast (due to cuts only restricting the future costs), then the progress is a bit slower. The different lines in the plot indicate next to the perfect foresight solution also two additional optimizations: The worst case line is calculated by a complete low wind time series put into a perfect foresight model. The best case boundary

is the same but with a complete high wind time series. While the best case leads to no storage expansion at all, the worst case installs way more storage than all other approaches (c.f. table 4.10). The light blue line indicates the mean of the upper bound of the last ten valid upper bounds. It can be seen, that it is not helpful for this kind of optimization. The calculation in case of investment possibilities should be improved in further studies.

By analyzing the cuts, the structure is similar to the deterministic DDP case. In the master problem, mainly cuts restricting only the future costs are added in the beginning. After some iterations, cuts containing the state of charge and the installed capacity are added. An exemplary cut is shown by the ninth cut added to the master problem:

$$0.004\alpha_0 \geq 0.27 \cdot 10^9 - 0.09\epsilon^{\text{con}}_{v_1 s0} - 0.08\epsilon^{\text{con}}_{v_2 s0} - 0.05\kappa^{\text{c}}_{v_1 s} - 0.07\kappa^{\text{c}}_{v_2 s}$$

v_1 : Baden-Württemberg, v_2 : Thuringia, s : Pumped storage

The cut shows that increasing both, the storage content $\epsilon^{\text{con}}_{v0}$ and the pumped storage capacity κ^{c}_{v} of Baden-Württemberg and Thuringia, leads to a reduction of the future costs. If the costs for investing in storage capacity are smaller than the expected savings, the model chooses to expand storage capacity. It is not necessarily the case that if the storage content of one state is included in the cut that its respective storage capacity is included as well. The appearance of the state of charge is due to the fact that the relaxed storage content constraint setting the state of charge at the beginning of the subproblem to the end state of the master problem $\epsilon^{\text{m, con}}_{vst}$ is active, i.e. fulfilled with equality instead of inequality:

$$\epsilon^{\text{con}}_{vst} \leq \epsilon^{\text{m, con}}_{vst} + \lambda\omega$$

The storage capacity on the other hand is restricted by another, also relaxed constraint:

$$\kappa^{\text{c}}_{vs} \leq \hat{\kappa}^{\text{c}}_{vs} + K^{\text{c}}_{vs} + \lambda\omega$$

The storage capacity κ^{c}_{vs} is at maximum the sum of given installed capacity K^{c}_{vs}, newly built capacity $\hat{\kappa}^{\text{c}}_{vs}$ and relaxation term $\lambda\omega$. Only if the constraint is active, the dual variable of the constraint will be nonzero and hence, the variable κ^{c}_{vs} will be included in the cut. This indicates that the main driver for the relaxation are these active constraints.

For the subproblems, the cuts contain the already emitted CO_2 emissions for varying states, similar to the DDP case. Hence, as the subproblems are more focused on the dispatch, the relaxation of the CO_2 constraints represents a more important facet than relaxing storage constraints.

When looking at the installed capacities over iterations, the expansion is slowly diverging as shown in figure 4.26. Similar to DDP only every 10th iteration is saved to create not too much data for analysis. In contrast to DDP, the algorithm does not start with Saxony but mostly considers Schleswig-Holstein for newly installed capacity. After more than 300 iterations, other states are included in the expansion as well. Saxony is one of the main interesting spots as well as Hesse, which was not considered by both, perfect foresight and mean wind time series.

As not every path might be in need of new storage capacity, the algorithm converges slowly.

Comparison of storage expansion The main question of the long term expansion study lies within the insight about storage investments gained by using an approach which incorporates

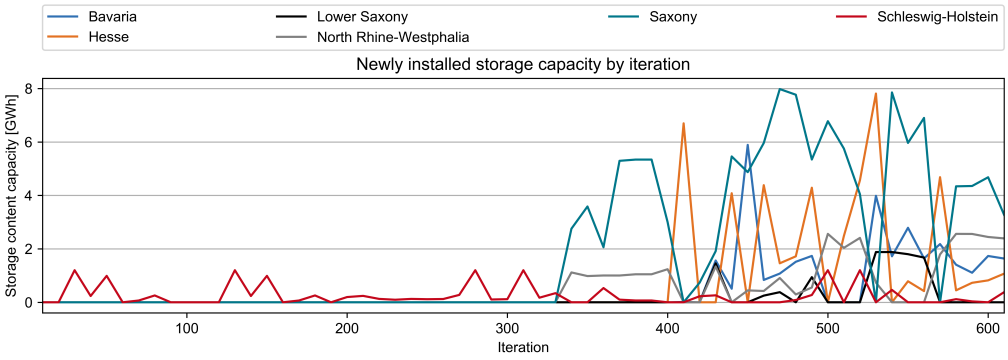


Figure 4.26: Newly installed storage capacity over iteration for SDDP approach

State	Perfect foresight	Mean	SDDP	Worst case
Baden-Württemberg	0	97	0	97
Bavaria	0	121	1637	556
Hesse	0	0	1071	253
Lower Saxony	820	644	0	1880
North Rhine-Westphalia	582	582	2390	1048
Rhineland-Palatinate	98	0	0	2599
Saxony	2335	2007	3267	5781
Saxony-Anhalt	115	51	0	513
Schleswig-Holstein	829	353	380	948
Σ	4780	3855	8745	13 675

Table 4.10: Newly installed capacity of storages in [MWh] for long term study with perfect foresight, mean and SDDP approach

uncertainty. Table 4.10 summarizes the newly installed power plant capacities per state for the different modeling techniques. For a regional analysis, figure 4.27 shows the same capacities in their respective state. It can be seen that especially in the northern states, new capacity is built. A first impression might be that these are the states where most wind capacity is installed, but Saxony, where most capacity is built, has neither great wind nor solar potential. The reasons for why so much capacity is built in Saxony mainly lies within its transmission lines to Brandenburg, where much installed wind power leads to overproduction.

The results for SDDP and DDP might seem a bit arbitrary, as the storage expansion is varying a lot over the iterations. So it seems to be difficult to draw concrete policy implications from these results. Comparing it with the behavior of the cost function, one can draw another conclusion from these results: mainly that there are many expansion solutions which lead to nearly the same costs. Hence, all similar solutions can be taken and analyzed further with respect to e.g. social aspects or regional distinctions, which have not been considered for this analysis.

Scenario long term dispatch For the *long term dispatch* scenario, the optimal storage capacity is taken from the perfect foresight approach (scenario *storage expansion*) as input.

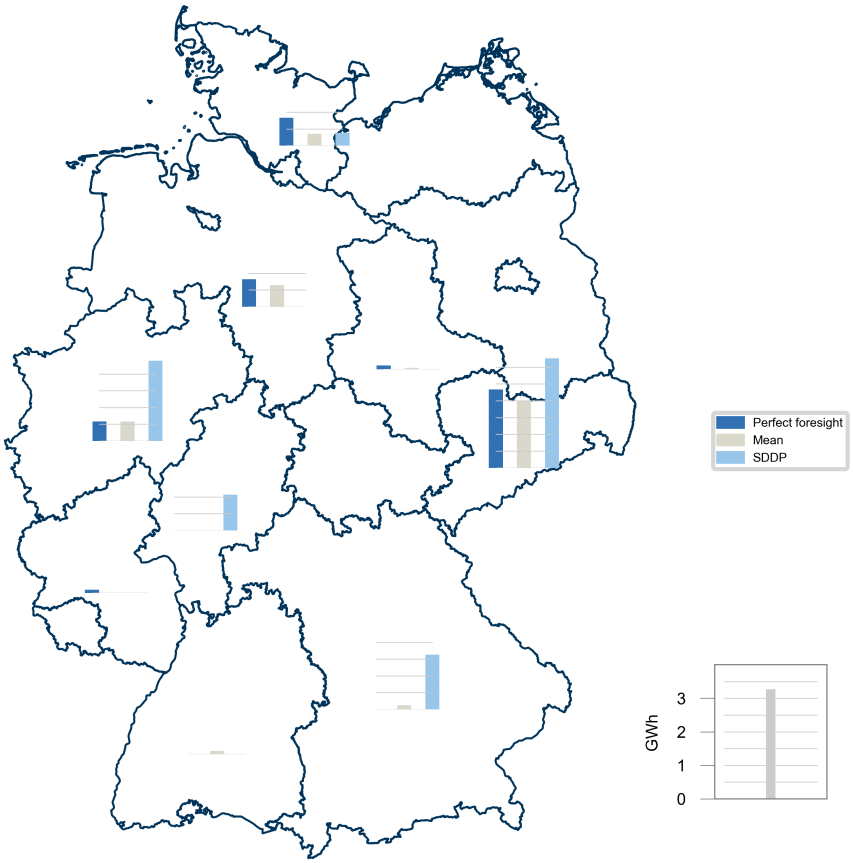


Figure 4.27: Comparison of additional storage content capacity results in scenario *storage expansion*

Scenario	Perfect foresight	SDDP	EVPI
long term dispatch	13.745	14.241	0.496

Table 4.11: Costs for long term dispatch study with perfect foresight and SDDP approach in billion €

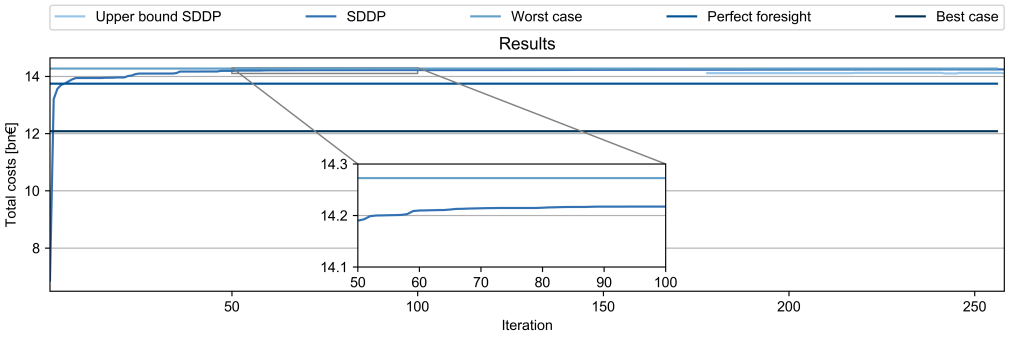


Figure 4.28: Resulting costs over iteration for SDDP approach for dispatch optimization

Only the dispatch is optimized for one year. The partitioning into five subproblems and the realizations are taken as before. The results can be compared with the perfect foresight approach for scenario *storage expansion* subtracting the investment costs. Table 4.11 shows the costs for one year dispatch. Comparing the costs for the perfect foresight case with and without expansion, one can see that annualized investment costs only account for 0.03 % of the total costs for one year. The costs for one year of dispatch calculated by the SDDP approach rise by 3.6 % compared to the perfect foresight approach.

The convergence of the costs compared to the perfect foresight approach and two perfect foresight approaches with the lowest possible wind and the highest possible wind time series are visualized in figure 4.28. Similar to the expansion scenario, the costs for the stochastic approach are converging towards the worst-case costs. However, there is still a gap between both solutions of 496 M€ as expected. The main problem in this figure is again the upper bound calculation. The tradeoff between calculating an approximate value and having none is visible again: the mean upper bound of the last ten iterations, which is only updated if the forward simulation does not violate the complicating constraints (storage constraints or CO_2 emission boundaries). As this is mostly the case for paths including *high* or sometimes *mid* realizations, the value can be lower, too.

Looking at the cuts, the main difference to the DDP case lies within that more cuts only containing the future costs are created. This is due to the fact that there are many possible paths in the stochastic case compared to the deterministic case. These cuts on the future costs are visible in all types of problems (master and sub). As the iterations proceed, cuts containing the CO_2 emissions are included in the subproblems. This shows that mainly the relaxation of CO_2 constraints has the biggest influence on the specific problem.

Comparison of dispatch Figure 4.29 shows the produced energy for one year with different approaches: high wind, low wind and perfect foresight account again for a deterministic approach with different wind time series (only high, only low wind and the original 2015 time series). Row SDDP represents the dispatch for the path *mid*, i.e. if the 2015 wind time series would indeed be the one which takes place. Therefore, the bars for wind and solar production are equal to the perfect foresight approach as the time series is the same. Compared to the perfect foresight approach more gas is used than lignite or coal in the SDDP case. Hence, the emitted CO_2 emissions are lower than in the perfect foresight case. This is due to the fact

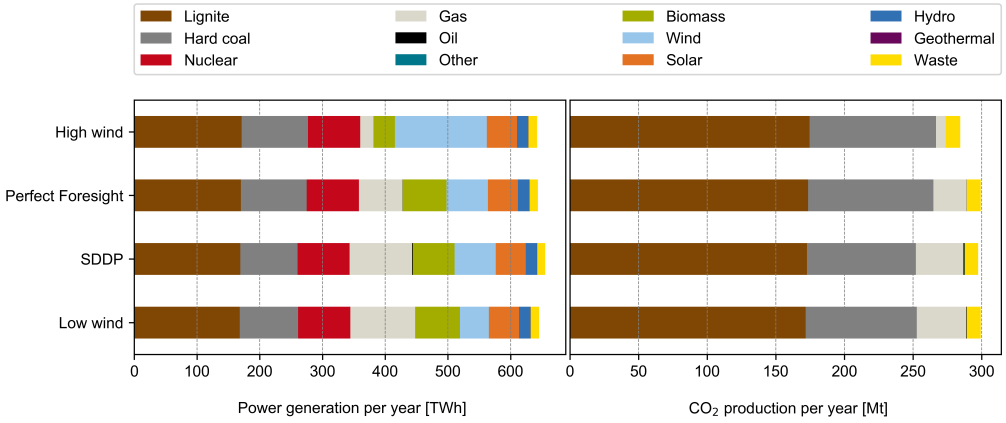


Figure 4.29: Comparison of produced energy and CO_2 per energy carrier for long term dispatch scenario

that in the stochastic approach, all paths have to fulfill the CO_2 target, so CO_2 emissions are "saved" for those times with worse wind conditions.

From comparing the low wind with the high wind approach, one can see, that less lignite and coal are again used in the low wind case as to fulfill the given CO_2 target. With all three deterministic approaches one can draw the conclusion that the first step of saving CO_2 emissions lies within replacing lignite and coal production by gas.

4.5 Remarks and summary

In general, it can be observed that the SDDP algorithm is highly dependent on the chosen paths. Hence, a convergence criterion has to be set carefully to ensure that the algorithm does not stop too early or too late. If enough information about the optimization problem is collected, it might be beneficial to use these worst paths for generating cuts. The convergence does not necessarily profit from it, but additional insight can be generated through it.

The upper bound calculation for SDDP with the relaxed optimization problem formulation (3.70)-(3.73) is especially difficult if only times of high marginal prices are considered. In that case, different wind realizations lead to large differences between possible paths. If a path with high wind realizations is already solvable without realization, and hence, leads to an upper bound calculation, every other path might not be feasible, yet. The upper bound is only set, if a path is feasible, i.e. the relaxation is not necessary except for additional production, transmission or storage capacities. The additional costs for these capacities can be calculated via the investment costs (c.f. [113]). For a faster stopping criterion it was tried during the thesis to approximate the violation costs to faster obtain upper bounds. The problem with these calculations lies in the difficulty to know how much additional power would cost at that moment, as capacity, environmental and other factors have to be taken into account. Too high approximations can lead to unrealistically high upper bounds, too low approximations might lead to a too early stop of the algorithm with an assumed convergence. Therefore, the calculation of the upper bound and the convergence criterion has to be analyzed carefully and is open for improvement.

For future analysis, it might be beneficial to look into the clustering of the regions more closely. A preliminary study with a one node model of Germany shows promising results as convergence is faster and computation times in general decrease. Due to the aggregation of the regions, cuts which only contain information about some regions do not lead to an overcompensation at another region.

All in all, the chapter shows how an extensive result analysis might look like and how to work with results from an approach incorporating uncertainty. Next to having insights about the optimal dispatch with an uncertain future, such an analysis can provide additional information about robust investment decisions for storage capacities. However, the chosen setup does not seem suitable for an explicit investment recommendation, but a more general direction in which regions an investment is appropriate due to high fluctuations of newly installed capacities. This problem is mainly due to the low annualized investment costs compared to fuel and variable costs. However, the investment results, that storage expansion in states situated in the middle of Germany like Saxony and Rhineland-Palatinate have a beneficial influence for a future energy system, can be used as basis for a further analysis based upon other factors like social aspects, which have not been modeled for this case study. For both, the short term and long term dispatch, the method is equally suitable and gives more insight into the general operation of an energy system.

Chapter 5

Conclusion

The thesis consists of three main pillars playing a key role to handle uncertainty in power system optimization: data processing, stochastic optimization and model result analysis. This chapter provides a short summary on each of them, as well as an outlook on where further work could yield most benefits.

5.1 Summary

Data processing is an essential part of every power system study. The need for transparency and openness, which go hand in hand, has been motivated and mentioned throughout this thesis. Multiple examples from several branches in daily life (economic, administrative, etc.) in which the usage of open data does not only ensure transparency but also leads to business cases and economic growth are given. But not only the advantages have been presented, but issues arising from opening up data are illustrated as well, mainly the problem of licensing. The thesis has introduced licenses, which guarantee open data, and has discussed the common problem of contradicting licenses (share alike) in data sets. Next to this theoretical background of data gathering, the concepts of openness and transparency have been applied to gather, transform and aggregate data for a power system representation of Bavaria and Germany. The processing itself has been published for reuse in an easily accessible form.

After preparing the data, the modeler is faced with choosing the appropriate mathematical way to express the power system. This choice is, of course, highly dependent on the research question at hand, e.g. a stability analysis of a microgrid might need more refined ways than a holistic optimization of whole Europe. As the focus in this thesis lies on uncertainty regarding renewable energy sources, the chosen method has been a stochastic approach with an implementation of Stochastic Dual Dynamic Programming (SDDP) in contrast to a deterministic representation with help of a scenario tree. The algorithm has been developed step-by-step with a non-stochastic Dual Dynamic Programming (DDP) procedure before elaborating further into the specifics of the stochastic approach. The thesis has extended the algorithm by a unified cut procedure known from Benders decomposition. The specific adaptations which have to be done to fit a power system model have been elaborated and the resulting framework is openly published as well.

The last step of modeling is the result analysis containing validation and interpretation of

the gained insights. The presented case studies focus on the power system model of Germany. After a brief validation of the results, a sensitivity analysis on the capacity factors of renewable energy sources has yielded that wind is the most unpredictable of the three sources (wind, solar and hydro), and has next to solar the biggest impact on the result. Hence, the capacity factor of the wind time series is taken as the uncertain parameter for the stochastic analysis. Three case studies have demonstrated the different possibilities of using SDDP: a short term (two day) study for dispatch decisions and a long term study for robust investment decisions of storage expansion and long term dispatch. For all studies, the important criteria have been explained and the post processing scripts have been published for reproducibility.

5.2 Outlook

The thesis has tackled several problems of handling uncertainty in power system optimization. However, there are several topics which could not be explored. These are:

Regarding the data processing for input preparation, working with GitHub releases (with versionable DOIs) and Jupyter notebooks is a viable way. As Jupyter notebooks are not easily trackable via git due to their output, one might need to look for alternatives which incorporate both advantages: trackability and easily accessible output graphs and pictures. For the input data itself, it might be beneficial to incorporate the usage of the Open Energy Database (oedb) [132]. The database provides an API, which could load the data from the framework itself. While the underlying data can be accessed by browser or API, the main feature of the database lies within its metadata potential. Each set can be tagged, licensed, versionized and altered by anybody. It is yet in construction, but might be a good starting point to increase transparency and reproducibility within studies.

Next to stochastic optimization, there are several other ways on to express uncertainty in optimization problems, which have been briefly mentioned in the introduction. Especially the combination of robust optimization and stochastic optimization is more and more common in the literature. Most hybrid methods work with some kind of weighted objective function to incorporate worst-case robustness and stochastic information [133, 94, 95]. Another way is to ease the conservatism of the robust approach by partioning the uncertainty set with help of stochastic information [96].

The method cannot only be extended by including robust optimization, but the algorithm itself could be improved further. As it has been already mentioned in chapter 3 and 4, the calculation of the upper bound can be further refined [134]. Additionally, the optimization of the weighting coefficients for the relaxation terms in the cut generation can be improved [113]. Other methods to improve the computation time might be achieved with a scenario reduction [123] or with quadratic regularization [111]. It might also be of interest to look into methods on how to expand the linear formulation into a mixed-integer unit commitment problem and therefore expand the SDDP method to an integer method called Stochastic Dual Dynamic Integer Programming (SDDiP) [135].

Other interesting case studies would include a further analysis of the detail of regional clustering. If less regions are used within the presented algorithm, faster convergence could be reached by both less iteration time as well as less overcompensating effects due to the cuts.

Hence, while the thesis provides valuable insight into the modeling of power systems with uncertainty, there are still many open research questions to dive into the topic further. Hopefully, this thesis can motivate new modelers to pick up some ideas from it, work on them and create the next step for a better understanding of energy systems, which is needed for the challenges we face regarding the energy transition.

Appendix A

Input Data

This appendix includes the input data for both models – Germany and Bavaria except time series due to space limitations. As described in chapter 2, the scripts to create these input files [74] and the files itself [82] can be found online.

Bavaria	Germany
7,898	302,854

Table A.1: CO_2 [kt] upper bounds used for Bavaria and Germany [73]

Commodity	Type	Price [€/MWh]
Hydro	Supim	-
Solar	Supim	-
Wind	Supim	-
Elec	Demand	-
CO_2	Env	0
Biomass	Stock	21
Coal	Stock	7
Gas	Stock	20
Geothermal	Stock	0
Lignite	Stock	4
Nuclear	Stock	2
Oil	Stock	21
Other	Stock	20
Slack	Stock	999
Waste	Stock	4

Table A.2: Used commodities and prices for all sites of Bavaria and Germany [136, p. 12]

Process / Site	LB	LF	MF	SW	UB	UF	UP
Biomass plant	190	115	220	294	406	100	207
CC plant	118	269	1,001	0	2,178	0	0
Geothermal plant	0	0	0	0	25	0	0
Hard coal plant	0	24	17	0	804	0	0
Hydro plant	527	85	17	310	1,047	33	68
Natural gas plant	16	57	18	51	349	0	13
Nuclear plant	1,410	0	0	1,288	0	0	0
Oil plant	0	0	0	57	1,325	0	0
Other plant	0	0	0	0	6	0	0
Solar plant	2,507	1,060	1,123	2,032	2,294	748	1,175
Waste plant	0	45	18	10	87	0	54
Wind plant	12	383	318	80	111	345	212

Table A.3: Installed process capacities in Bavaria per administrative region in MW [61]

Site Process	BW	BY	BE	BB	HB	HH	HE	NI	MV
Biomass	1,022	1,535	61	539	9	61	294	1,476	379
CC	434	3,568	444	282	15	127	345	2,174	183
Geothermal	1	25	0	0	0	0	0	0	0
Hard coal	4,667	847	777	0	772	194	753	2,162	514
Hydro	864	2,091	0	4	20	0	81	58	3
Lignite	0	0	0	4,409	0	0	34	352	0
Natural gas	518	487	467	301	0	22	1,114	701	136
Nuclear	2,712	2,698	0	0	0	0	0	2,696	0
Oil	702	1,384	218	334	86	0	25	56	0
Other	0	0	0	0	0	0	0	19	0
Solar	4,985	10,943	78	2,861	38	36	1,761	3,429	1,270
Waste	98	214	36	118	91	24	112	73	17
Wind	605	1,465	4	5,445	151	51	1,080	8,095	2,594

Site Process	NW	OF	RP	SL	SN	ST	SH	TH
Biomass	866	0	176	19	282	589	441	272
CC	4,968	0	1,728	75	440	449	0	291
Geothermal	0	0	8	0	0	0	0	0
Hard coal	7,827	0	13	1,822	0	0	680	0
Hydro	159	0	232	11	210	26	5	33
Lignite	10,618	0	0	0	4,325	1,136	0	0
Natural gas	1,714	0	185	58	113	238	22	177
Nuclear	0	0	0	0	0	0	1,410	0
Oil	332	0	0	0	17	212	320	0
Other	0	0	0	0	0	0	0	0
Solar	4,039	0	1,845	404	1,558	1,785	1,474	1,095
Waste	432	0	83	28	16	185	17	12
Wind	3,684	776	2,694	239	1,082	4,306	4,726	1,185

Table A.4: Installed process capacities in Germany per state in MW [61] [65]

Process	Commodity	Direction	ratio
Biomass plant	Biomass	In	2.86
	CO_2	Out	0.00
	Elec	Out	1.00
CC plant	CO_2	Out	0.34
	Elec	Out	1.00
	Gas	In	1.71
Geothermal plant	Elec	Out	1.00
	Geothermie	In	10.00
Hard coal plant	CO_2	Out	0.87
	Coal	In	2.56
	Elec	Out	1.00
Hydro plant	Elec	Out	1.00
	Hydro	In	1.00
Lignite plant	CO_2	Out	1.02
	Elec	Out	1.00
	Lignite	In	2.56
Natural gas plant	CO_2	Out	0.51
	Elec	Out	1.00
	Gas	In	2.53
Nuclear plant	Elec	Out	1.00
	Nuclear	In	3.03
Oil plant	CO_2	Out	0.68
	Elec	Out	1.00
	Oil	In	2.50
Other plant	CO_2	Out	0.69
	Elec	Out	1.00
	Other	In	2.56
Solar plant	Elec	Out	1.00
	Solar	In	1.00
Waste plant	CO_2	Out	0.77
	Elec	Out	1.00
	Waste	In	2.56
Wind plant	Elec	Out	1.00
	Wind	In	1.00

Table A.5: Relations between input and output for processes in Bavaria and Germany [137] [138]

Site In	Site Out	inv-cost	fix-cost	inst-cap
Lower Bavaria	Middle Franconia	206,779	113	349
	Upper Bavaria	199,196	109	4,973
	Upper Palatinate	208,330	114	824
Lower Franconia	Middle Franconia	193,552	106	608
	Upper Franconia	213,460	117	2,897
Middle Franconia	Upper Bavaria	338,188	186	1,128
	Upper Franconia	170,299	93	2,478
	Upper Palatinate	193,214	106	371
Swabia	Upper Bavaria	206,779	113	2,970
Upper Franconia	Upper Palatinate	175,124	96	732

Table A.6: Transmission investment [€/MW], fixed costs [€/MW] and capacities [MW] for Bavaria

Site In	Site Out	inv-cost	fix-cost	inst-cap
Baden-Württemberg	Bavaria	333,375	183	5,393
	Hesse	495,933	272	2,763
	Rhineland-Palatinate	425,807	234	593
	Saarland	372,365	204	852
Bavaria	Hesse	557,899	306	6,374
	Thuringia	521,140	286	2,141
Berlin	Brandenburg	105,444	57	7,105
	Mecklenburg-Vorpommern	299,992	164	583
	Saxony-Anhalt	258,062	141	764
Brandenburg	Mecklenburg-Vorpommern	275,790	151	773
	Saxony	375,369	206	6,685
	Saxony-Anhalt	171,744	94	4,500
Bremen	Lower Saxony	67,012	36	2,961
Hamburg	Lower Saxony	202,751	111	3,538
	North Rhine-Westphalia	568,071	312	800
	Schleswig-Holstein	144,964	79	5,014
Hesse	Lower Saxony	489,655	269	800
	North Rhine-Westphalia	281,246	154	9,521
	Rhineland-Palatinate	287,624	158	5,503
	Saarland	406,584	223	1,016
	Thuringia	287,410	158	1,803
Lower Saxony	North Rhine-Westphalia	365,399	200	11,769
	Offshore	635,027	349	5,309
	Saxony-Anhalt	400,229	220	2,783
	Schleswig-Holstein	318,636	175	4,356
Mecklenburg-Vorpommern	Offshore	216,129	118	600
	Saxony-Anhalt	410,100	225	1,021
	Schleswig-Holstein	373,320	205	2,279
North Rhine-Westphalia	Rhineland-Palatinate	340,136	187	12,518
Offshore	Schleswig-Holstein	386,595	212	0
Rhineland-Palatinate	Saarland	137,063	75	3,203
Saxony	Thuringia	342,088	188	5,658
Saxony-Anhalt	Thuringia	263,381	144	2,517

Table A.7: Transmission investment [€/MW], fixed costs [€/MW] and capacities [MW] for Germany

Parameter	Value
eff-in	0.80
eff-out	1
inv-cost-p [€/MW]	450,000
inv-cost-c [€/MWh]	6,500
fix-cost-p [€/MW]	11,000
fix-cost-c [€/MWh]	0
var-cost-p [€/MW]	0.3
var-cost-c [€/MWh]	0
depreciation [a]	70
wacc	0.07

Table A.8: Technical pump storage parameters for Bavaria and Germany [138]

Site	inst-cap-c	inst-cap-p
Lower Bavaria	100	10
Lower Franconia	950	164
Middle Franconia	900	160
Swabia	0	0
Upper Bavaria	1,100	92
Upper Franconia	0	0
Upper Palatinate	404	127

Table A.9: Installed capacities of pumped-storage capacity [MWh] and power [MW] in Bavaria

Site	inst-cap-c	inst-cap-p
Baden-Württemberg	10,392	1,873
Bavaria	3,354	543
Berlin	0	0
Brandenburg	0	0
Bremen	0	0
Hamburg	0	0
Hesse	3,906	623
Lower Saxony	940	220
Mecklenburg-Vorpommern	0	0
North Rhine-Westphalia	1,280	291
Offshore	0	0
Rhineland-Palatinate	4,630	1,291
Saarland	0	0
Saxony	4,609	1,085
Saxony-Anhalt	523	79
Schleswig-Holstein	600	119
Thuringia	12,115	1,509

Table A.10: Installed capacities of pumped-storage capacity [MWh] and power [MW] in Germany

Appendix B

Additional graphs

This appendix contains additional graphs for the case studies presented in chapter 4. The graphs include plots of CO_2 emissions, residual loads and marginal prices.

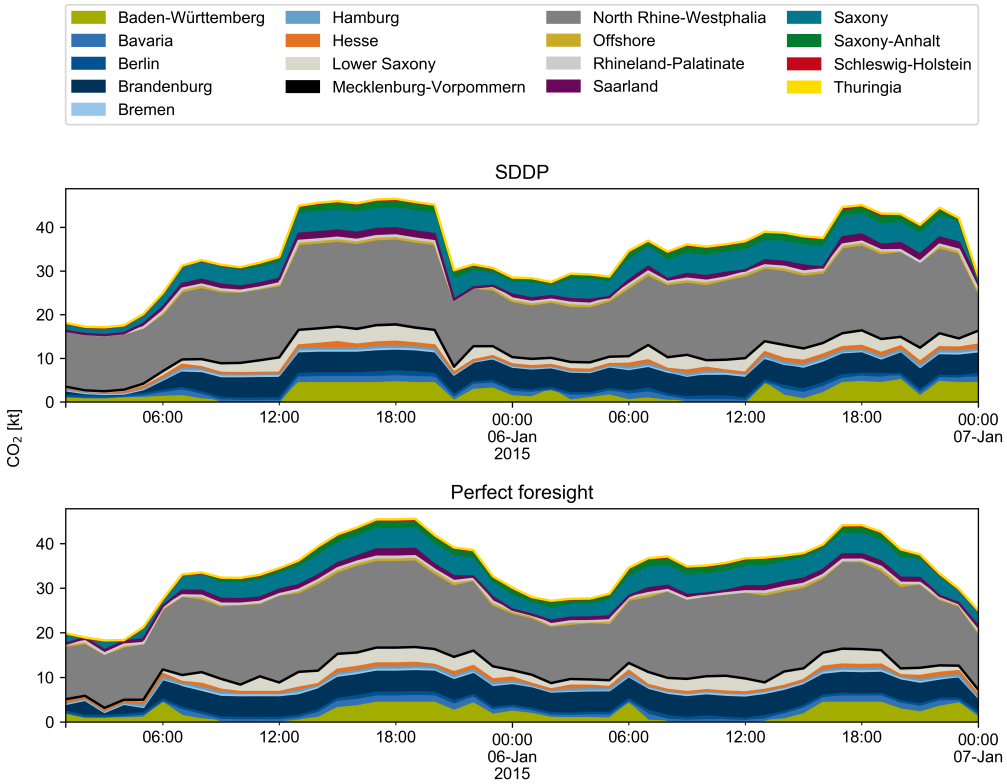


Figure B.1: CO_2 emissions for short term case study for realization *mid* in scenario *large storage*

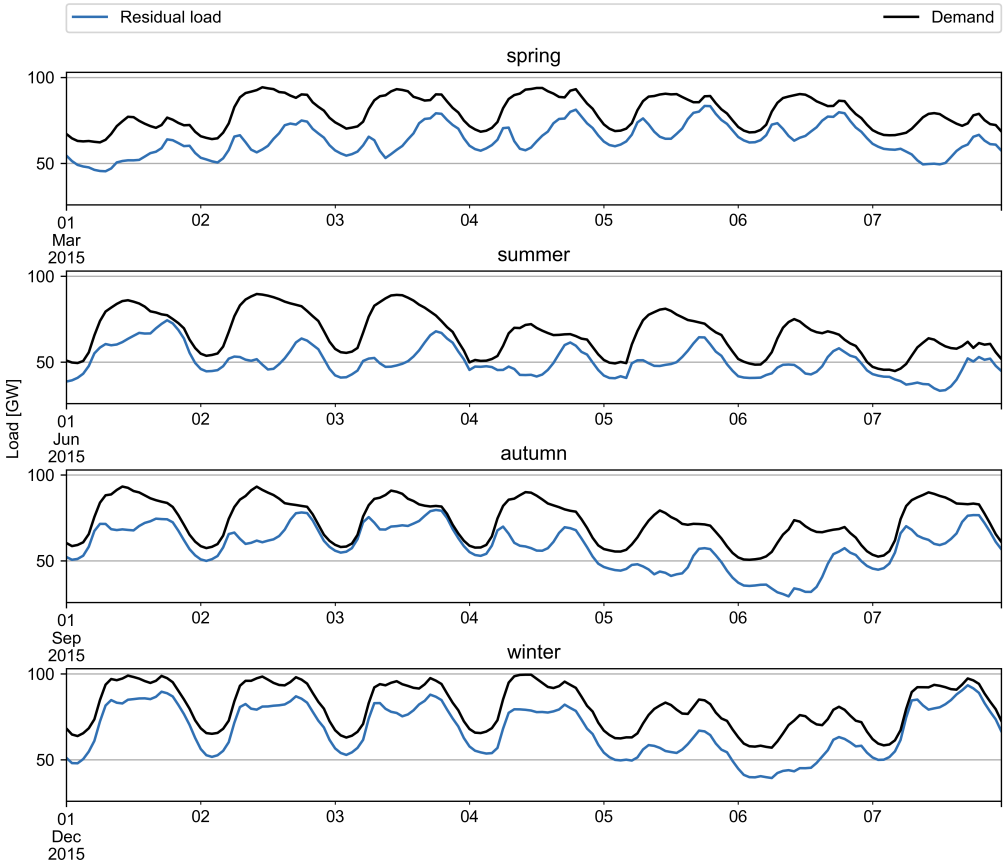


Figure B.2: Residual load in spring, summer, autumn and winter

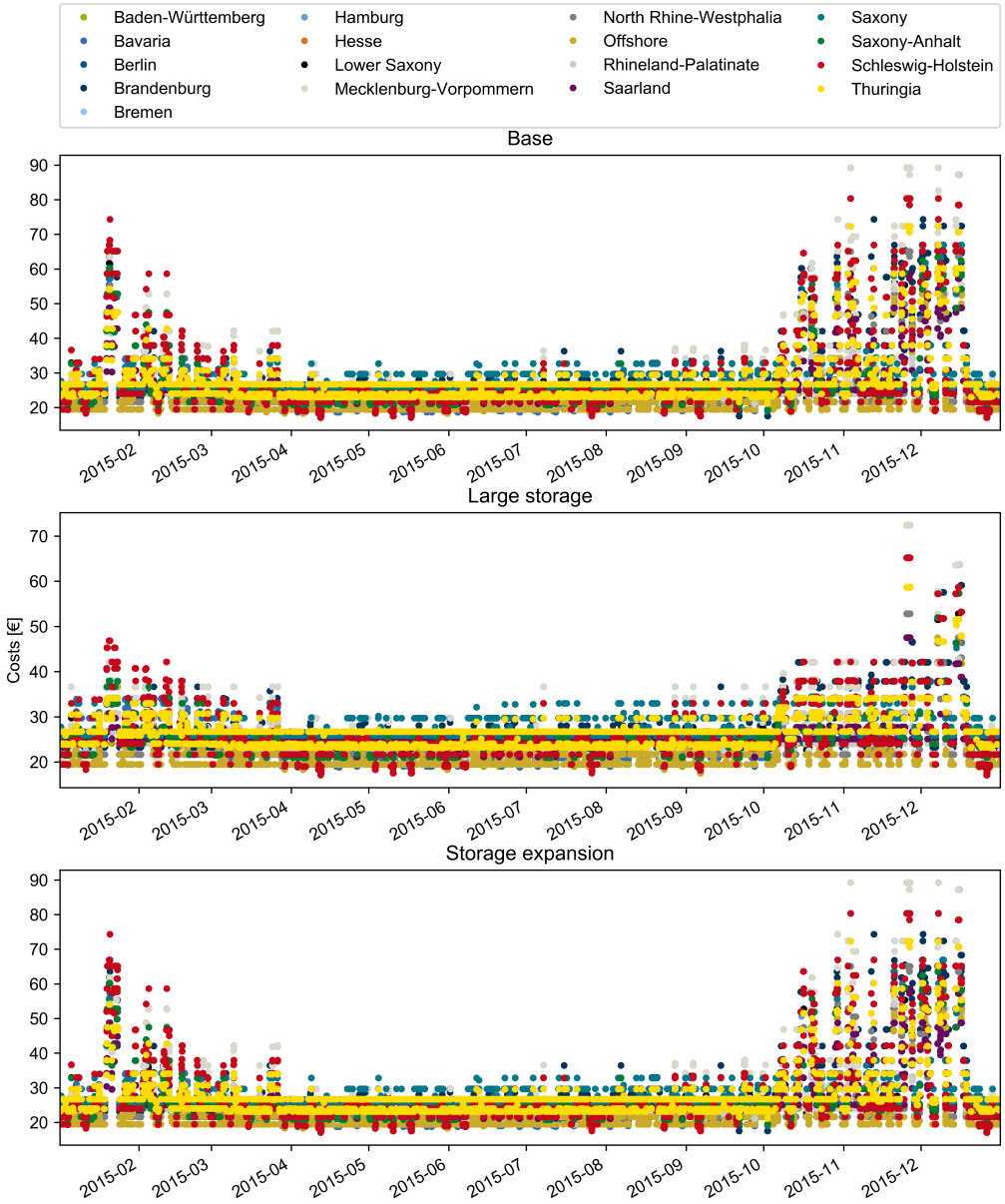


Figure B.3: Marginal prices for one year for scenario *base*, *large storage* and *large storage expansion* during day time (09:00 - 17:00)

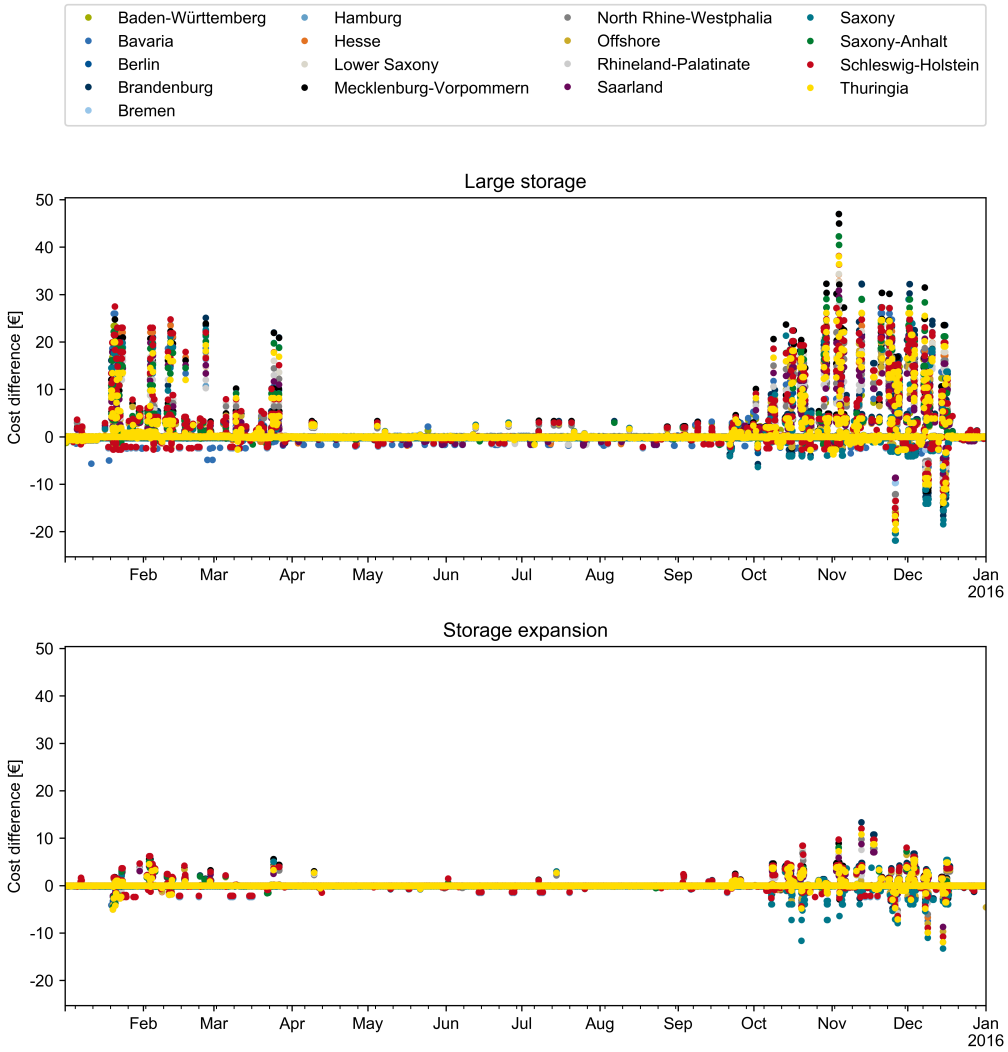


Figure B.4: Marginal price differences between *base* case and the two other scenarios

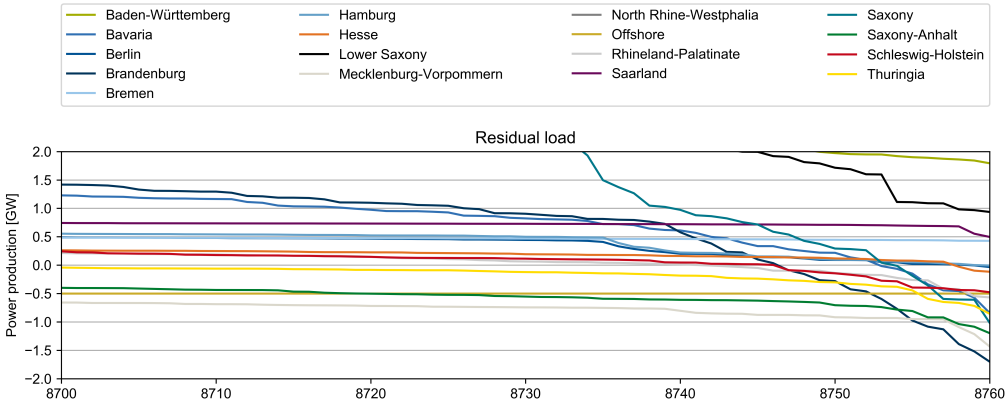


Figure B.5: Residual load for perfect foresight approach (last 60 hours)

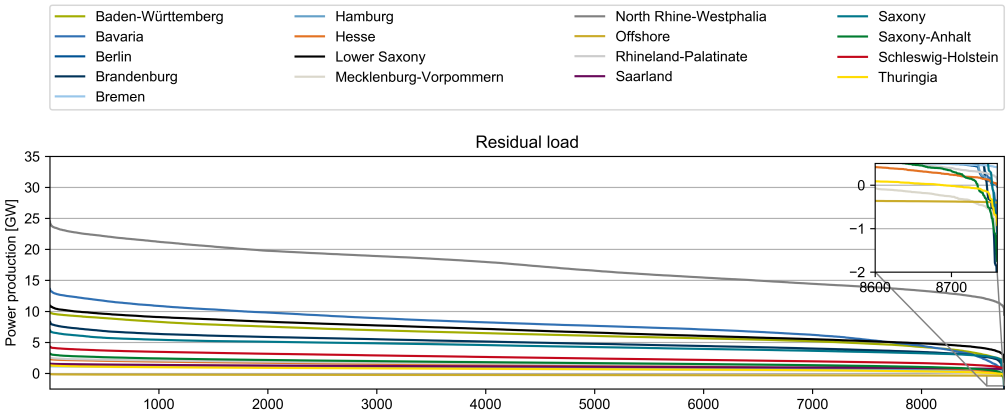


Figure B.6: Residual load for mean wind time series approach

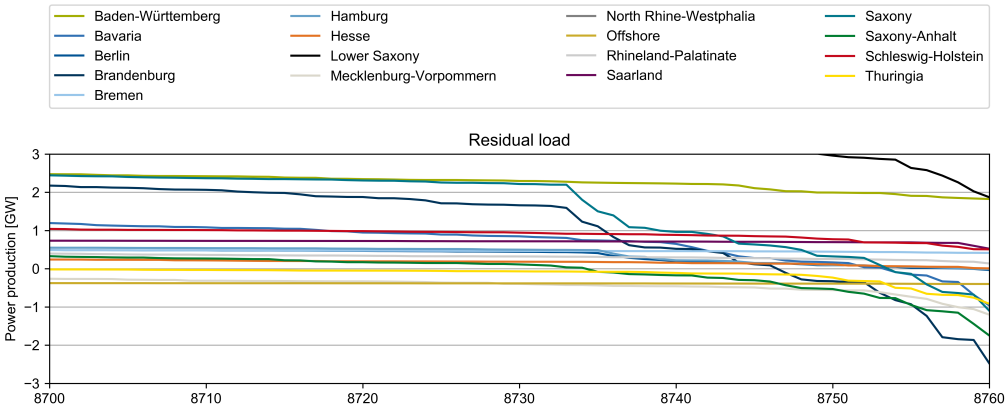


Figure B.7: Residual load for mean wind time series approach (last 60 hours)

Glossary

BNetzA Bundesnetzagentur (German Federal Network Agency) 23, 25, 26

CC Creative Commons 10, 20, 21, 22

DDP Dual Dynamic Programming 1, 3, 4, 5, 7, 13, 44, 47, 46, 55, 56, 57, 58, 60, 62, 63, 64, 66, 77, 96, 98, 104, 106, 104, 106, 107, 108, 109, 110, 111, 113, 117

DIKW Data-Information-Knowledge-Wisdom 4, 15, 16

EVPI Expected value of perfect information 90, 91, 93, 109, 111

FAIR findable, accessible, interoperable and reusable 21

FMEAE Federal Ministry for Economic Affairs and Energy 30, 83, 84

FOSS Free and Open Source Software 21

GDP Gross Domestic Product 19

GIT Generalized Information Theory 9

LAE Länderarbeitskreis Energiebilanzen 83, 84

NC non-commercial 21

OGD Open Government Data 18, 19

OPSD Open Power System Data 11, 23, 26

PDF Probability Density Function 39

SA share alike 22

SAA Sample Average Approximation 56

SDDiP Stochastic Dual Dynamic Integer Programming 118

SDDP Stochastic Dual Dynamic Programming 1, 3, 4, 5, 6, 7, 13, 14, 35, 39, 41, 42, 56, 57, 58, 57, 60, 61, 62, 66, 68, 70, 72, 77, 79, 77, 79, 80, 81, 83, 93, 94, 96, 98, 108, 109, 110, 111, 113, 111, 113, 114, 117, 118

SOC state of charge 94

TSO Transmission System Operator 23

VSS Value of stochastic solution 90, 91, 98, 109

Bibliography

[1] Andrea Saltelli et al. *Global Sensitivity Analysis. The Primer*. John Wiley & Sons, Ltd, Jan. 2008. ISBN: 9780470059975. DOI: 10.1002/9780470725184 (cit. on pp. 9, 14, 85).

[2] William Briggs. *Uncertainty*. Springer International Publishing, 2016. DOI: 10.1007/978-3-319-39756-6 (cit. on p. 9).

[3] Bilal M Ayyub and George J Klir. *Uncertainty Modeling and Analysis in Engineering and the Sciences*. Chapman and Hall/CRC, May 25, 2006. 400 pp. ISBN: 1584886447 (cit. on p. 9).

[4] George J. Klir and Mark J. Wierman. *Uncertainty-Based Information: Elements of Generalized Information Theory (Studies in Fuzziness and Soft Computing)*. Physica Verlag, 1998. ISBN: 3-7908-1073-8. DOI: 10.1007/978-3-7908-1869-7 (cit. on pp. 9, 12).

[5] George J. Klir. "Generalized information theory". In: *Fuzzy Sets and Systems* 40.1 (Mar. 1991), pp. 127–142. DOI: 10.1016/0165-0114(91)90049-v (cit. on p. 9).

[6] Deutsche Forschungsgemeinschaft (DFG). "Safeguarding Good Scientific Practice". In: *Safeguarding Good Scientific Practice*. Wiley-VCH Verlag GmbH & Co. KGaA, Oct. 2013, pp. 1–109. ISBN: 9783527679188. DOI: 10.1002/9783527679188.oth1 (cit. on p. 9).

[7] James Speight and Russell Foote. *Ethics in science and engineering*. Hoboken, NJ ; Salem, Mass.: Wiley ; Scrivener, 2011. ISBN: 9780470626023 (cit. on p. 9).

[8] Francesco Sylos Labini. *Science and the Economic Crisis*. Springer International Publishing, 2016. DOI: 10.1007/978-3-319-29528-2 (cit. on pp. 9, 11, 14).

[9] Robbie Morrison. "Energy system modeling: Public transparency, scientific reproducibility, and open development". In: *Energy Strategy Reviews* 20 (Apr. 2018), pp. 49–63. DOI: 10.1016/j.esr.2017.12.010 (cit. on p. 10).

[10] European Commission's Central IP Service. "Reuse policy: a study on available reuse implementing instruments and licensing considerations". In: EUR 29685 EN (Mar. 28, 2019). Ed. by Publications Office of the European Union. Luxembourg. DOI: 10.2760/95373 (cit. on p. 11).

[11] Stefan Pfenninger et al. "Opening the black box of energy modelling: Strategies and lessons learned". In: *Energy Strategy Reviews* 19 (Jan. 2018), pp. 63–71. DOI: 10.1016/j.esr.2017.12.002 (cit. on p. 11).

[12] Karl-Kiên Cao et al. “Raising awareness in model-based energy scenario studies — a transparency checklist”. In: *Energy, Sustainability and Society* 6.1 (Sept. 28, 2016), p. 1. ISSN: 2192-0567. DOI: 10.1186/s13705-016-0090-z (cit. on p. 11).

[13] Stefan Pfenninger, Adam Hawkes, and James Keirstead. “Energy systems modeling for twenty-first century energy challenges”. In: *Renewable and Sustainable Energy Reviews* 33 (May 2014), pp. 74–86. DOI: 10.1016/j.rser.2014.02.003 (cit. on p. 11).

[14] Georgios Savvidis et al. “The gap between energy policy challenges and model capabilities”. In: *Energy Policy* 125 (Feb. 2019), pp. 503–520. DOI: 10.1016/j.enpol.2018.10.033 (cit. on p. 11).

[15] Hans-Kristian Ringkjøb, Peter M. Haugan, and Ida Marie Solbrekke. “A review of modelling tools for energy and electricity systems with large shares of variable renewables”. In: *Renewable and Sustainable Energy Reviews* 96 (Nov. 2018), pp. 440–459. DOI: 10.1016/j.rser.2018.08.002 (cit. on p. 11).

[16] Markus Groissböck. “Are open source energy system optimization tools mature enough for serious use?” In: *Renewable and Sustainable Energy Reviews* 102 (Mar. 2019), pp. 234–248. DOI: 10.1016/j.rser.2018.11.020 (cit. on p. 11).

[17] Stefan Pfenninger et al. “The importance of open data and software: Is energy research lagging behind?” In: *Energy Policy* 101 (Feb. 2017), pp. 211–215. DOI: 10.1016/j.eneco.2013.07.014 (cit. on p. 11).

[18] Frauke Wiese et al. “Open Power System Data – Frictionless data for electricity system modelling”. In: *Applied Energy* 236 (Feb. 2019), pp. 401–409. DOI: 10.1016/j.apenergy.2018.11.097 (cit. on p. 12).

[19] Raul Banos et al. “Optimization methods applied to renewable and sustainable energy: A review”. In: *Renewable and Sustainable Energy Reviews* 15.4 (May 2011), pp. 1753–1766. DOI: 10.1016/j.rser.2010.12.008 (cit. on p. 12).

[20] A. Rezaee Jordehi. “How to deal with uncertainties in electric power systems? A review”. In: *Renewable and Sustainable Energy Reviews* 96 (Nov. 2018), pp. 145–155. DOI: 10.1016/j.rser.2018.07.056 (cit. on p. 12).

[21] Weldon A. Lodwick and Phantipa Thipwiwatpotjana. *Flexible and Generalized Uncertainty Optimization*. Springer International Publishing, 2017. DOI: 10.1007/978-3-319-51107-8 (cit. on p. 12).

[22] Alireza Soroudi and Turaj Amraee. “Decision making under uncertainty in energy systems: State of the art”. In: *Renewable and Sustainable Energy Reviews* 28 (Dec. 2013), pp. 376–384. DOI: 10.1016/j.rser.2013.08.039 (cit. on pp. 12, 39).

[23] Morteza Aien, Ali Hajebrahimi, and Mahmud Fotuhi-Firuzabad. “A comprehensive review on uncertainty modeling techniques in power system studies”. In: *Renewable and Sustainable Energy Reviews* 57 (May 2016), pp. 1077–1089. DOI: 10.1016/j.rser.2015.12.070 (cit. on pp. 12, 39).

[24] Urmila Diwekar. “Optimization under Uncertainty”. In: *Optimization under Uncertainty* 13.1 (2002), pp. 1–8. URL: http://www.mcs.anl.gov/~leyffer/views/13-1.pdf (visited on 12/05/2018) (cit. on p. 13).

[25] Urmila M. Diwekar. *Introduction to Applied Optimization*. Springer US, 2003. ISBN: 978-1-4757-3745-5. DOI: 10.1007/978-1-4757-3745-5 (cit. on pp. 13, 14, 91).

[26] Nikolaos E. Koltsaklis, Pei Liu, and Michael C. Georgiadis. “An integrated stochastic multi-regional long-term energy planning model incorporating autonomous power systems and demand response”. In: *Energy* 82 (Mar. 2015), pp. 865–888. ISSN: 0360-5442. DOI: 10.1016/j.energy.2015.01.097 (cit. on p. 13).

[27] Mario Pereira and Leontina Pinto. “Multi-stage stochastic optimization applied to energy planning”. In: *Mathematical programming* 52.1-3 (May 1991), pp. 359–375. DOI: 10.1007/bf01582895 (cit. on pp. 13, 42, 46).

[28] Steffen Rebennack. “Combining sampling-based and scenario-based nested Benders decomposition methods: application to stochastic dual dynamic programming”. In: *Mathematical Programming* 156.1-2 (Mar. 2015), pp. 343–389. DOI: 10.1007/s10107-015-0884-3 (cit. on pp. 13, 40, 42).

[29] Max Kühne. “Drivers of energy storage demand in the German power system: an analysis of the influence of methodology and parameters on modelling results”. PhD thesis. Department of Electrical and Computer Engineering, Technical University of Munich, June 13, 2016. URL: http://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:bvb:91-diss-20160613-1280982-1-7 (visited on 01/03/2019) (cit. on pp. 14, 85).

[30] Haris Doukas, Alexandros Flamos, and Jenny Lieu, eds. *Understanding Risks and Uncertainties in Energy and Climate Policy*. Springer International Publishing, 2019. DOI: 10.1007/978-3-030-03152-7 (cit. on p. 14).

[31] Assem Deif. *Sensitivity Analysis in Linear Systems*. Springer Berlin Heidelberg, 1986. ISBN: 978-3-642-82741-9. DOI: 10.1007/978-3-642-82739-6 (cit. on pp. 14, 85).

[32] Jennifer Rowley. “The wisdom hierarchy: representations of the DIKW hierarchy”. In: *Journal of Information Science* 33.2 (Feb. 2007), pp. 163–180. DOI: 10.1177/0165551506070706 (cit. on pp. 15, 16).

[33] Chaim Zins. “Conceptual approaches for defining data, information, and knowledge”. In: *Journal of the American Society for Information Science and Technology* 58.4 (2007), pp. 479–493. DOI: 10.1002/asi.20508 (cit. on p. 16).

[34] Open Knowledge International. *What is Open?* Nov. 13, 2018. URL: https://okfn.org/opendata/ (visited on 11/13/2018) (cit. on pp. 16, 17).

[35] Bridgette Wessels et al. *Open Data and the Knowledge Society*. 2017. DOI: 10.5117/9789462980181 (cit. on p. 17).

[36] Cristiano Castelfranchi. “Comment: Six critical remarks on science and the construction of the knowledge society”. In: *Journal of Science Communication*. International School for Advanced Studies 6 (4 2007), pp. 1–3. DOI: 10.22323/2.06040303 (cit. on p. 17).

[37] Nico Stehr. *Knowledge societies*. Sage Publications, 1994. 304 pp. ISBN: 0-8039-7892-8. URL: https://www.ebook.de/de/product/3834593/nico_stehr_knowledge_societies.html (visited on 10/22/2018) (cit. on pp. 17, 23).

[38] Sarah Cummings et al. “Critical discourse analysis of perspectives on knowledge and the knowledge society within the Sustainable Development Goals”. In: *Development Policy Review* 36.6 (Aug. 2018), pp. 727–742. DOI: 10.1111/dpr.12296 (cit. on pp. 17, 18).

[39] United Nations Educational, Scientific and Cultural Organization. *Towards knowledge societies*. UNESCO, 2005, p. 220. ISBN: 92-3-204000-X. URL: http://unesdoc.unesco.org/images/0014/001418/141843e.pdf (visited on 11/09/2018) (cit. on p. 18).

[40] United Nations. *Transforming our world: The 2030 agenda for sustainable development*. Oct. 21, 2015. URL: http://www.un.org/en/development/desa/population/migration/generalassembly/docs/globalcompact/A_RES_70_1_E.pdf (cit. on p. 18).

[41] Abraham Taherivand. “Das freie Internet ist nicht tot. Wie können wir gemeinsam dafür kämpfen?” German. In: *Das ist Netzpolitik*. Talk. Sept. 22, 2018. URL: https://media.ccc.de/v/np14-5-das_freie_internet_ist_nicht_tot_wie_koennen_wir_gemeinsam_dafuer_kaempfen (cit. on p. 18).

[42] S&P Global Platts. *An Historical Perspective*. Tech. rep. 2009. URL: https://www.platts.com/IM.Platts.Content/AboutPlatts/Platts%20History-Full%20Summary_5-08-09Final_pdf.pdf (cit. on p. 18).

[43] Richard Scott. *IEA The First 20 Years: The History of the International Energy Agency 1974-1994 – Origins and Structure*. Volume One. ISBN: 92-64-14059-X. 2, Rue André-Pascal, 75775 Paris Cedex 16, France, 1994. DOI: 10.1787/9789264020931-en. URL: https://web.archive.org/web/20070415071141/http://www.iea.org/Textbase/nppdf/free/1990/1-ieahistory.pdf (cit. on p. 18).

[44] International Energy Agency. *Data Services: Electricity Information (2018 Edition)*. IEA – International Energy Agency. 2018. URL: http://data.iea.org/payment/products/102-electricity-information-2018-edition.aspx (visited on 08/28/2018) (cit. on p. 18).

[45] Barbara Ubalbi. “Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives”. In: *OECD Working Papers on Public Governance* (May 2013). DOI: 10.1787/5k46bj4f03s7-en (cit. on p. 19).

[46] Bret Goldstein and Lauren Dyson, eds. *Beyond transparency: Open Data and the Future of Civic Innovation*. San Francisco, CA: Code for America press, 2013. URL: https://beyondtransparency.org/ (visited on 03/04/2019) (cit. on p. 19).

[47] Leonhard Dobusch et al. “Open Government Data: eine Initiative der Open-Commons Region Linz”. In: *Open Source – Konzepte, Risiken, Trends* 283 (2012). ISSN: 1436-3011 (cit. on p. 19).

[48] European Commission. “Commission notice: guidelines on recommended standard licenses, datasets and charging for the reuse of documents”. In: *Official Journal of the European Union* (July 24, 2014). URL: http://ec.europa.eu/newsroom/dae/document.cfm?action=display&doc_id=6421 (visited on 11/26/2018) (cit. on p. 19).

[49] Jan Kucera and Dusan Chlapek. “Benefits and Risks of Open Government Data”. In: *Journal of Systems Integration* 5.1 (2014), pp. 30–41. DOI: 10.20470/jsi.v5i1.185 (cit. on p. 19).

[50] Mar Negreiro. “Re-use of public sector information”. In: *EU Legislation in Progress*. Nov. 2018. URL: http://www.europarl.europa.eu/RegData/etudes/BRIE/2018/628312/EPRS_BRI(2018)628312_EN.pdf (visited on 11/14/2018) (cit. on p. 19).

[51] Logan Byers et al. *A Global Database of Power Plants*. Tech. rep. Washington, DC: World Resources Institute, 2018. URL: https://www.wri.org/publication/global-power-plant-database (cit. on p. 20).

[52] Till Kreutzer. *Open Content. A Practical Guide to Using Creative Commons Licences*. Deutsche UNESCO-Kommission, Hochschulbibliothekszentrum des Landes Nordrhein-Westfalen, Wikimedia Deutschland - Gesellschaft zur Förderung des Freien Wissens (2014), 2016. ISBN: 978-3-940785-57-2. URL: https://www.unesco.de/sites/default/files/2018-06/Open_Content_A_Practical_Guide_to_Using_Open_Content_Licences_web.pdf (visited on 11/15/2018) (cit. on pp. 20–22).

[53] Daniel Dietrich et al. *Open Data Handbook – Guide*. Open Knowledge International. Aug. 30, 2018. URL: http://opendatahandbook.org/guide/en (cit. on p. 20).

[54] Sandra Collins et al. *Turning FAIR data into reality. Final report and action plan from the European Commission expert group on FAIR data*. Tech. rep. Luxembourg: Directorate-General for Research and Innovation (European Commission), Publications Office of the European Union, Nov. 26, 2018. DOI: 10.2777/1524. (Visited on 11/28/2018) (cit. on p. 21).

[55] Open Knowledge International. *Open Definition – Conformant Licenses*. Nov. 13, 2018. URL: http://opendefinition.org/licenses/ (visited on 11/13/2018) (cit. on p. 21).

[56] Matěj Myška et al. “Creative Commons and Grand Challenge to Make Legal Language Simple”. In: *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2012, pp. 271–285. DOI: 10.1007/978-3-642-35731-2_19 (cit. on p. 22).

[57] Monica Adriana Lupascu. “Legal Implications of Open Licenses”. In: *Lex et Scientia* 23.2 (2016), p. 64. URL: http://lexetscientia.univnt.ro/en/article/LEGAL-IMPLICATIONS-OF-OPEN-LICENSES~575.html (cit. on p. 22).

[58] Georgia M. Kapitsaki and Frederik Kramer. “Open Source License Violation Check for SPDX Files”. In: *Lecture Notes in Computer Science. Software Reuse for Dynamic Systems in the Cloud and Beyond*. 14th International Conference on Software Reuse. (Jan. 4, 2015). Miami, USA: Springer International Publishing, 2014, pp. 90–105. DOI: 10.1007/978-3-319-14130-5_7 (cit. on p. 22).

[59] P. De Filippi and L. Maurel. “The paradoxes of open data and how to get rid of it? Analysing the interplay between open data and sui-generis rights on databases”. In: *International Journal of Law and Information Technology* 23.1 (Oct. 2014), pp. 1–22. DOI: 10.1093/ijlit/eau008 (cit. on p. 22).

[60] Johannes Dorfner. “Open Source Modelling and Optimisation of Energy Infrastructure at Urban Scale”. PhD thesis. Department of Electrical and Computer Engineering, Technical University of Munich, 2016. URL: http://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:bvb:91-diss-20161206-1285570-1-6 (visited on 01/03/2019) (cit. on pp. 23, 77, 92).

[61] Bundesnetzagentur für Elektrizität, Gas, Telekommunikation, Post und Eisenbahnen. *Kraftwerksliste*. Data Sheet, dl-de/by-2-0. Nov. 2016. URL: https://www.bundesnetzagentur.de/SharedDocs/Downloads/DE/Sachgebiete/Energie/Unternehmen_Institutionen/Versorgungssicherheit/Erzeugungskapazitaeten/Kraftwerksliste/Kraftwerksliste_2016.html (visited on 02/02/2019) (cit. on pp. 25–27, 122, 123).

[62] Ingmar Schlecht. *Renewable power plants*. 2018. DOI: 10.25832/renewable_power_plants/2018-03-08. (Visited on 02/03/2019) (cit. on pp. 25–27).

[63] Jens Weibezahn et al. *Conventional Power Plants*. 2018. DOI: 10.25832/conventional_power_plants/2018-12-20. (Visited on 02/03/2019) (cit. on pp. 25–27).

[64] Jonathan Muehlenpfordt. *Time series. Load, wind and solar, prices in hourly resolution*. Data sheets, Licence: ENTSO-E Transparency. 2018. DOI: 10.25832/time_series/2018-06-30. URL: https://doi.org/10.25832/time_series/2018-06-30 (visited on 02/01/2019) (cit. on pp. 25, 30).

[65] Federal Ministry for Economic Affairs and Energy. *Übersicht Offshore-Netzanbindungen*. Website, CC BY-ND 3.0 DE. July 2018. URL: https://www.erneuerbare-energien.de / EE / Navigation / DE / Technologien / Windenergie - auf - See / Offshore - Projekte/Netzanbindungen/netzanbindungen.html (cit. on pp. 27, 123).

[66] Jonas Egerer. *Open source Electricity Model for Germany (ELMOD-DE)*. DIW Data Documentation 83. Tech. rep. German Institute for Economic Research (DIW Berlin), 2016. URL: https : / / www . econstor . eu / handle / 10419 / 129782 (visited on 02/03/2019) (cit. on p. 26).

[67] Fabian Hofmann, Jonas Hörsch, and Fabian Gotzens. *Fresna/Powerplantmatching: V0.2*. Aug. 2018. DOI: 10 . 5281 / zenodo . 1405595. URL: https : / / zenodo . org / record/1405595 (visited on 02/03/2019) (cit. on p. 26).

[68] Carsten Matke, Wided Medjroubi, and David Kleinhans. *SciGRID - An Open Source Reference Model for the European Transmission Network (v0.2)*. ODbL v1.0. July 2016. URL: http://www.scigrid.de (visited on 02/03/2019) (cit. on pp. 27, 29).

[69] H. P. St. Clair. “Practical Concepts in Capability and Performance of Transmission Lines”. In: *IEEE Transactions on Power Apparatus and Systems* 72.6 (Dec. 1953), pp. 1152–1157. DOI: 10.1109/AIEEPAS.1953.4498751 (cit. on p. 27).

[70] OpenStreetMap contributers. *OpenStreetMap project*. 2018. URL: https://openstreetmap.org (cit. on p. 27).

[71] Bundesnetzagentur für Elektrizität, Gas, Telekommunikation, Post und Eisenbahnen. *SMARD Strommarktdaten*. Website, CC BY-ND 3.0 DE. July 2018. URL: https://smard.de/ (cit. on p. 30).

[72] Federal Ministry for Economic Affairs and Energy. *Energy Data: Complete Edition*. Data Sheet. Jan. 2018. URL: https://www.bmwi.de/Redaktion/DE/Binaer/Energiedaten/energiedaten-gesamt-xls.xls (cit. on pp. 30, 84).

[73] Länderarbeitskreis Energiebilanzen. Bereitgestellt durch das Statistische Landesamt Bremen. 2017. URL: http://www.lak-energiebilanzen.de/ (cit. on pp. 30, 32, 34, 83, 84, 121).

[74] Magdalena Stüber and Leonhard Odersky. *Phd thesis release. Jupyter notebooks for pre- and post-processing of Bavaria and Germany*. Jupyter Notebook, License: GPL-3.0. Oct. 2019. DOI: 10.5281/zenodo.3522610 (cit. on pp. 31, 34, 83, 121).

[75] Stefan Pfenninger and Iain Staffell. "Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data". In: *Energy* 114 (Nov. 2016), pp. 1251–1265. DOI: 10.1016/j.energy.2016.08.060 (cit. on p. 31).

[76] Ronald Gelaro et al. "The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2)". In: *Journal of Climate* 30.14 (July 2017), pp. 5419–5454. DOI: 10.1175/jcli-d-16-0758.1 (cit. on p. 31).

[77] Iain Staffell and Stefan Pfenninger. "Using bias-corrected reanalysis to simulate current and future wind power output". In: *Energy* 114 (Nov. 2016), pp. 1224–1239. DOI: 10.1016/j.energy.2016.08.068 (cit. on pp. 31, 104).

[78] Dietrich Oeding and Bernd Oswald. *Elektrische Kraftwerke und Netze*. 8th ed. Springer, 2016. DOI: 10.1007/978-3-662-52703-0 (cit. on p. 32).

[79] Wasserstraßen und Schifffahrtsverwaltung des Bundes (WSV, bereitgestellt durch Bundesanstalt für Gewässerkunde (BfG). *Pegelonline*. Website, E-Mail request. Nov. 2019. URL: https://www.pegelonline.wsv.de/ (visited on 02/03/2019) (cit. on p. 32).

[80] Bayerisches Landesamt für Umwelt. *Gewässerkundlicher Dienst Bayern – Waterlevel Bavaria*. Website, Licence: CC BY 4.0. 2018. URL: https://www.gkd.bayern.de/en/rivers/waterlevel/ (cit. on p. 32).

[81] Bayerisches Landesamt für Umwelt. *Energieatlas Bayern – Kartenteil*. Website, Licence: CC BY 4.0. 2018. URL: https://www.energieatlas.bayern.de/ (cit. on p. 32).

[82] Magdalena Stüber. *Stochastic optimization with urbs for Germany and Bavaria*. 2019. DOI: 10.5281/ZENODO.3522730 (cit. on pp. 34, 77, 83, 92, 121).

[83] James K. Ho, Tak C. Lee, and R. P. Sundarraj. "Decomposition of linear programs using parallel computation". In: *Mathematical Programming* 42.1-3 (Apr. 1988), pp. 391–405. DOI: 10.1007/bf01589413 (cit. on p. 35).

[84] George B. Dantzig and Philip Wolfe. “Decomposition principle for linear programs”. In: *Operations research* 8.1 (Feb. 1960), pp. 101–111. DOI: 10.1287/opre.8.1.101 (cit. on pp. 35, 36).

[85] Jacques Benders. “Partitioning procedures for solving mixed-variables programming problems”. In: *Numerische Mathematik* 4.1 (1962), pp. 238–252. ISSN: 0945-3245. DOI: 10.1007/BF01386316 (cit. on pp. 35, 38).

[86] Antonio Conejo et al. *Decomposition techniques in mathematical programming: engineering and science applications*. Springer, 2006. DOI: 10.1007/3-540-27686-6 (cit. on pp. 35, 36).

[87] Ragheb Rahmaniani et al. “The Benders decomposition algorithm: A literature review”. In: *European Journal of Operational Research* 259.3 (June 2017), pp. 801–817. DOI: 10.1016/j.ejor.2016.12.005 (cit. on p. 38).

[88] George B. Dantzig. “Linear Programming Under Uncertainty”. In: *Management Science* 50.12_supplement (Dec. 2004), pp. 1764–1769. DOI: 10.1287/mnsc.1040.0261 (cit. on p. 39).

[89] Dimitris Bertsimas, David Brown, and Constantine Caramanis. “Theory and applications of robust optimization”. In: *SIAM review* 53.3 (2011), pp. 464–501. DOI: 10.1137/080734510 (cit. on p. 39).

[90] Virgine Gabrel, Cécile Murat, and Aurélie Thiele. “Recent advances in robust Optimization: An overview”. In: *European Journal of Operational Research* 235.3 (June 2014), pp. 471–483. DOI: 10.1016/j.ejor.2013.09.036 (cit. on p. 40).

[91] Bo Zeng. *Solving Two-Stage Robust Optimization Problem by a Constraint-and-Column Generation Method*. Tech. rep. University of South Florida: Departement of Industrial Management Systems Engineering, June 2011. URL: http://www.optimization-online.org/DB_FILE/2011/06/3065.pdf (cit. on p. 40).

[92] Anna Danandeh, Long Zhao, and Bo Zeng. “Job Scheduling With Uncertain Local Generation in Smart Buildings: Two-Stage Robust Approach”. In: *IEEE Transactions on Smart Grid* 5.5 (Sept. 2014), pp. 2273–2289. DOI: 10.1109/TSG.2014.2315166 (cit. on p. 40).

[93] Ruiwei Jiang et al. “Two-stage network constrained robust unit commitment problem”. In: *European Journal of Operational Research* 234.3 (May 2014), pp. 751–762. DOI: 10.1016/j.ejor.2013.09.028 (cit. on p. 40).

[94] Germán Morales-Espana et al. *Robust Unit Commitment with Dispatchable Wind: An LP Reformulation of the Second Stage*. Optimization Community. July 2015. URL: http://www.optimization-online.org/DB_FILE/2014/09/4542.pdf (cit. on pp. 40, 118).

[95] Germán Morales-España, Álvaro Lorca, and Mathijs M. de Weerdt. “Robust unit commitment with dispatchable wind power”. In: *Electric Power Systems Research* 155 (Feb. 2018), pp. 58–66. DOI: 10.1016/j.epsr.2017.10.002 (cit. on pp. 40, 118).

[96] Patrick Gögler, Magdalena Dorfner, and Thomas Hamacher. “Hybrid Robust/Stochastic Unit Commitment With Iterative Partitions of the Continuous Uncertainty Set”. In: *Frontiers in Energy Research* 6 (July 2018). DOI: 10.3389/fenrg.2018.00071 (cit. on pp. 40, 118).

[97] Chaoyue Zhao and Yongpei Guan. “Unified Stochastic and Robust Unit Commitment”. In: *IEEE Transactions on Power Systems* 28.3 (Aug. 2013), pp. 3353–3361. DOI: 10.1109/tpwrs.2013.2251916 (cit. on p. 40).

[98] Yu An and Bo Zeng. “Exploring the Modeling Capacity of Two-Stage Robust Optimization: Variants of Robust Unit Commitment Model”. In: *IEEE Transactions on Power Systems* 30.1 (Jan. 2015), pp. 109–122. DOI: 10.1109/TPWRS.2014.2320880 (cit. on p. 40).

[99] I. Blanco and J. M. Morales. “An Efficient Robust Solution to the Two-Stage Stochastic Unit Commitment Problem”. In: *IEEE Transactions on Power Systems* 32.6 (Nov. 2017), pp. 4477–4488. ISSN: 0885-8950. DOI: 10.1109/TPWRS.2017.2683263 (cit. on p. 40).

[100] A. B. Philpott, V. L. de Matos, and L. Kapelevich. “Distributionally robust SDDP”. In: *Computational Management Science* 15.3-4 (May 2018), pp. 431–454. DOI: 10.1007/s10287-018-0314-0 (cit. on p. 40).

[101] John R. Birge and François Louveaux. *Introduction to Stochastic Programming*. Springer New York, 2011. DOI: 10.1007/978-1-4614-0237-4 (cit. on pp. 40, 41, 56).

[102] Alexander Shapiro, Darinka Dentcheva, and Andrzej Ruszczyński. *Lectures on Stochastic Programming: Modeling and Theory*. Society for Industrial & Applied Mathematics (SIAM), Jan. 2009. DOI: 10.1137/1.9780898718751 (cit. on pp. 40, 56, 80).

[103] R. T. Rockafellar and R. J-B. Wets. “Stochastic Convex Programming: Relatively Complete Recourse and Induced Feasibility”. In: *SIAM Journal on Control and Optimization* 14.3 (May 1976), pp. 574–589. DOI: 10.1137/0314038 (cit. on p. 41).

[104] Werner Römisch and Rüdiger Schultz. “Stochastic programs with complete recourse: Stability and an application to power dispatch”. In: (1990), pp. 688–696. DOI: 10.1007/BFb0008424 (cit. on p. 41).

[105] James Murphy. *Benders, Nested Benders and Stochastic Programming: An Intuitive Introduction*. CUED/F-INFENG/TR.675. Tech. rep. Cambridge University Engineering Department, Dec. 2013. URL: http://www.optimization-online.org/DB_FILE/2013/12/4157.pdf (cit. on pp. 41, 42).

[106] George B Dantzig and Gerd Infanger. *Large-scale stochastic linear programs: Importance sampling and Benders decomposition*. Tech. rep. Departement of Operations Research, Stanford University, 1991. URL: https://apps.dtic.mil/dtic/tr/fulltext/u2/a234962.pdf (visited on 12/05/2018) (cit. on p. 42).

[107] George B. Dantzig and Gerd Infanger. “Multi-stage stochastic linear programs for portfolio optimization”. In: *Annals of Operations Research* 45.1 (Dec. 1993), pp. 59–76. DOI: 10.1007/bf02282041 (cit. on p. 42).

[108] Jesús Velásquez, Pedro J. Restrepo, and Rafael Campo. “Dual dynamic programing: A note on implementation”. In: *Water Resources Research* 35.7 (July 1999), pp. 2269–2271. DOI: 10.1029/1999wr900052 (cit. on pp. 54, 55).

[109] Alexander Shapiro. “Analysis of stochastic dual dynamic programming method”. In: *European Journal of Operational Research* 209.1 (Feb. 2011), pp. 63–72. DOI: 10.1016/j.ejor.2010.08.007 (cit. on p. 56).

[110] Martin L. Puterman. *Markov Decision Processes*. John Wiley & Sons, Inc., Feb. 25, 2005. 680 pp. ISBN: 0-471-72782-2. DOI: 10.1002/9780470316887 (cit. on p. 70).

[111] Tsvetan Asamov and Warren B. Powell. “Regularized Decomposition of High-Dimensional Multistage Stochastic Programs with Markov Uncertainty”. In: *SIAM Journal on Optimization* 28.1 (Jan. 2018), pp. 575–595. DOI: 10.1137/16m1072231 (cit. on pp. 70, 118).

[112] Tito Homem-de-Mello, Vitor L. de Matos, and Erlon C. Finardi. “Sampling strategies and stopping criteria for stochastic dual dynamic programming: a case study in long-term hydrothermal scheduling”. In: *Energy Systems* 2.1 (Jan. 2011), pp. 1–31. DOI: 10.1007/s12667-011-0024-y (cit. on pp. 70, 81).

[113] Paul Stursberg. “On the Mathematics of Energy System Optimization. Network Models, Decomposition, and Economic Incentives”. PhD thesis. Departement of Mathematics, Technical University of Munich, 2019 (cit. on pp. 70, 72–76, 80, 114, 118).

[114] Matteo Fischetti, Domenico Salvagnin, and Arrigo Zanette. “A note on the selection of Benders’ cuts”. In: *Mathematical Programming* 124.1-2 (May 2010), pp. 175–182. DOI: 10.1007/s10107-010-0365-7 (cit. on pp. 72, 73, 75, 76).

[115] G. Cornuéjols and C. Lemaréchal. “A convex-analysis perspective on disjunctive cuts”. In: *Mathematical Programming* 106.3 (Nov. 2005), pp. 567–586. DOI: 10.1007/s10107-005-0670-8 (cit. on pp. 72, 73, 76).

[116] Egon Balas. “Disjunctive Programming”. In: *Annals of Discrete Mathematics* 5 (1979), pp. 3–51. DOI: 10.1016/s0167-5060(08)70342-x (cit. on p. 72).

[117] Egon Balas. *Disjunctive Programming*. Springer International Publishing, 2018. DOI: 10.1007/978-3-030-00148-3 (cit. on pp. 72, 73).

[118] T. Magnanti and R. Wong. “Accelerating Benders Decomposition: Algorithmic Enhancement and Model Selection Criteria”. In: *Operations Research* 29 (3 1981), pp. 417–626. DOI: 10.1287/opre.29.3.464 (cit. on p. 73).

[119] Julius Farkas. “Theorie der einfachen Ungleichungen.” In: *Journal für die reine und angewandte Mathematik (Crelle’s Journal)* 1902.124 (1902), pp. 1–27. DOI: 10.1515/crll.1902.124.1 (cit. on p. 75).

[120] N. Dinh and V. Jeyakumar. “Farkas’ lemma: three decades of generalizations for mathematical optimization”. In: *TOP* 22.1 (Feb. 2014), pp. 1–22. ISSN: 1863-8279. DOI: 10.1007/s11750-014-0319-y (cit. on p. 75).

[121] Johannes Dorfner et al. *urbs; v0.7*. 2017. DOI: 10.5281/zenodo.242029 (cit. on p. 77).

[122] David Wozabal, Christoph Graf, and David Hirschmann. “Renewable Energy and Its Impact on Power Markets”. In: *International Series in Operations Research & Management Science*. Springer US, 2013, pp. 283–311. DOI: 10.1007/978-1-4614-9035-7_12 (cit. on pp. 79, 85, 88).

[123] Jitka Dupačová and Václav Kozmík. “SDDP for multistage stochastic programs: preprocessing via scenario reduction”. In: *Computational Management Science* 14.1 (June 2016), pp. 67–80. DOI: 10.1007/s10287-016-0261-6 (cit. on pp. 81, 118).

[124] Suvrajeet Sen and Zhihong Zhou. “Multistage Stochastic Decomposition: A Bridge between Stochastic Programming and Approximate Dynamic Programming”. In: *SIAM Journal on Optimization* 24.1 (Jan. 2014), pp. 127–153. DOI: 10.1137/120864854 (cit. on p. 81).

[125] Lion Hirth, Ingmar Schlecht, and Jonathan Mühlenpfordt. *Open Data for Electricity Modeling. An assessment of input data for modeling the European electricity system regarding legal and technical usability*. A report for the Federal Ministry for Economic Affairs and Energy, Germany. Tech. rep. neon neue energieökonomik, Nov. 6, 2018. URL: https://www.bmwi.de/Redaktion/EN/Publikationen/Studien/open-Data-for-electricity-modeling.pdf?__blob=publicationFile&v=5 (visited on 01/02/2019) (cit. on p. 85).

[126] Lion Hirth, Jonathan Mühlenpfordt, and Marisa Bulkeley. “The ENTSO-E Transparency Platform – A review of Europe’s most ambitious electricity data platform”. In: *Applied Energy* 225 (Sept. 2018), pp. 1054–1067. DOI: 10.1016/j.apenergy.2018.04.048 (cit. on p. 85).

[127] Thomas Saaty and Saul Gass. “Parametric Objective Function (Part 1)”. In: *Journal of the Operations Research Society of America* 2.3 (Aug. 1954), pp. 316–319. DOI: 10.1287/opre.2.3.316 (cit. on p. 85).

[128] Saul Gass and Thomas Saaty. “Parametric Objective Function (Part 2) - Generalization”. In: *Journal of the Operations Research Society of America* 3.4 (Nov. 1955), pp. 395–401. DOI: 10.1287/opre.3.4.395 (cit. on p. 85).

[129] Saul Gass and Thomas Saaty. “The computational algorithm for the parametric objective function”. In: *Naval Research Logistics Quarterly* 2.1-2 (Mar. 1955), pp. 39–45. DOI: 10.1002/nav.3800020106 (cit. on p. 85).

[130] Gurobi Optimization, LLC. *Gurobi Optimizer Reference Manual*. 2019. URL: http://www.gurobi.com (visited on 05/08/2019) (cit. on p. 93).

[131] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. 7th. 2009. URL: https://web.stanford.edu/~boyd/cvxbook/ (visited on 05/08/2019) (cit. on p. 93).

[132] Open Energy Platform. *Open Energy Database*. 2019. URL: https://openenergy-platform.org/dataedit/ (visited on 05/17/2019) (cit. on p. 118).

[133] Dimitris Bertsimas et al. “Adaptive Robust Optimization for the Security Constrained Unit Commitment Problem”. In: *IEEE Transactions on Power Systems* 28.1 (Feb. 2013), pp. 52–63. DOI: 10.1109/TPWRS.2012.2205021 (cit. on p. 118).

[134] Vincent Leclère et al. “Exact converging bounds for Stochastic Dual Dynamic Programming via Fenchel duality”. Version 2. In: *Archive ourverte HAL* (Apr. 18, 2018). URL: https://hal-enpc.archives-ouvertes.fr/hal-01744035v2 (visited on 10/15/2019) (cit. on p. 118).

[135] Jikai Zou, Shabbir Ahmed, and Xu Andy Sun. “Stochastic dual dynamic integer programming”. In: *Mathematical Programming* 175.1-2 (Mar. 2018), pp. 461–502. DOI: 10.1007/s10107-018-1249-5 (cit. on p. 118).

[136] S. Wissel et al. *Stromerzeugungskosten im Vergleich*. Technical Report 4. Institute of Energy Economics and the Rational Use of Energy, Feb. 2008. URL: http://www.ier.uni-stuttgart.de/publikationen/arbeitsberichte/downloads/Arbeitsbericht_04.pdf (cit. on p. 122).

[137] International Energy Agency. *World Energy Model Assumptions 2016– Investment in power generation*. Data Sheet, Licence: www.iea.org/t&c; as modified by Magdalena Dorfner. 2016. URL: https://www.iea.org/weo/weomodel/ (cit. on p. 124).

[138] Philipp Kuhn. “Iteratives Modell zur Optimierung von Speicherausbau und -betrieb in einem Stromsystem mit zunehmend fluktuierender Erzeugung”. PhD thesis. Department of Electrical and Computer Engineering, Technical University of Munich, 2012. URL: http://mediatum.ub.tum.de/?id=1271192 (cit. on pp. 124, 127).