University of Amsterdam

Master Thesis

Stochastic optimization in the power grid

Author:
T.S. Leenman
(Student nr. 10296522)

Supervisors:
Prof. dr. N.M. van Dijk
Drs. F. Phillipson

September 3, 2012

Abstract

In this thesis steps are described to determine the locations of new wind mills which minimize energy loss on the Dutch High Voltage power grid. A vindication of the used power grid model is provided; the simulation procedure for stochastic wind power is described; and the required mathematical optimization models are formally derived as well as implemented in the optimization software package Aimms. Results are shown and their relation to real life problems is discussed.

The goal of the thesis is to work out a case study of power grid design where stochasticity plays a major role, and where transportation losses should be minimized. By lack of data this has not been carried out for a (perhaps more relevant) more locally oriented and Medium Voltage case, but this thesis provides a description of all crucial steps that should be taken. This thesis contains the information on power systems and the explanation of the relevant mathematical techniques for any reader to find his way in the literature on power grid design, and to address problems of a similar nature as the one treated in this thesis.

Preface

The thesis presented here is the result of a seven month internship at TNO Communication and Information Technology in Delft, with which I conclude my Master in Operations Research at the University of Amsterdam. During this period I had the opportunity to occupy myself with the energy sector, to read scientific research for a broad variety of applications, and to get to know research practice in a business environment. There was much time and room to pursue my own interests, and to participate in other activities besides my main research. For that I would like to express my gratitude to TNO, which provided me with this opportunity.

I thank Nico van Dijk for the interesting and motivating conversations; Frank Phillipson for opening up to me the research area and giving structure to my work; and Carolien van Vliet-Hameeteman for the weekly updates and advices. I am grateful to Cees Duin for thinking along in the early stages; and Robin Swinkels, Miroslav Zivkovic and especially my father Jan Peter Leenman for reading and correcting the preliminary versions of this thesis.

Contents

1 Introduction

1.1 Motivation and subject

The subject of this Master thesis is the design of electrical power grids. It is a vast subject of enormous complexity. Every single aspect of the power system can be designed and operated in numerous ways, each different way having a different effect on the whole. All parts of it are liable to change and uncertainty. A large variety of different parties contribute to or make use of it, governed by their agreements on the liberalized market. Because power consumption is constantly growing, power production and power grids are constantly extended and modified; and at the same time, out of environmental concerns, there is a strong call for renewable energy, which by its changable or small scale nature, brings difficulties of its own.

These developments render it necessary that important decisions are made, and that a lot of research is done, at every level of the electricity system. In Dutch politics, it has been decided that liberalization of power companies was a good idea, and still there is the ever relevant issue of incentives for renewable energy; in international trade and politics, oil gas and coal prices should be monitored, and European electricity market integration is prepared; in national market design, the question is how to design the market so as to stimulate efficiency, and let usefulness for society coincide with the financial interests for every involved party; in power grid design goals are to minimize outage probability, energy losses due to transportation, and to keep up with the increased load of the grid. New research areas such as *Locational Marginal Pricing* and *Smart Grids* try to integrate the physical properties of the power network with the market structure of energy trade in order to achieve technical goals by economical means.

In the daily operation of the power grid by utilities and companies, regulation of the voltages to ensure system stability, forecasting the load of consumers, forecasting of uncertain production units, the cost-optimal scheduling of dispatchable power generation, asset maximization of power production in combination with CO_2 emission and heat generation, optimal trading strategies are all taken care of. Many aspects of the power grid are changing over time, and the one common thing in these developments is that everything tends to become more flexible. Trading possibilities increase, both for power companies (which can buy and sell now to and from neighbouring countries until an hour before delivery), as well as for consumers which increasingly have access to the market themselves. Of all renewable energy, wind and solar are the most succesful, but they are both of an uncontrolable and changable nature.

This thesis considers two aspects of the power system in order to study their effect on the design of the power grid. These aspects are stochasticity due to wind energy production, and energy losses due to electricity transportation. Stochastics in the power system is currently the subject of much research in the academic world[1], and of much worrying from the operator's side[2]. As known from [Energy in the Netherlands 2011] and [Croes 2011], in the Netherlands about 5% of generated electricity is lost during transportation, which comes down to an estimated 400 million euros per year in this country alone. So every slight increase in efficiency would yield enormous economic merits. In this thesis, the problem of *Generation Expansion Planning* is considered, by looking at the question where to build new wind mills. The criterion by which an answer is obtained, is the minimization of the transportation losses.

As a case example, the national power network of the Netherlands is taken. For this case, extensive wind data, the network properties of the High Voltage (HV) power grid, and load data are available online. The objective of the study is twofold: on the one hand, the research question is answered for this specific case study; and on the other hand, it provides a general orientation into the possibilities of stochastic modeling of the power grid.

The interesting relation between stochastic power generation and transportation losses could also be seen in the following way. Loss percentage for transportation of a certain amount of

[1] for example, [Toveren met Vermogen 2011] describes the research programs performed within IOP EMVT, several of which try to solve problems due to stochasticity in the grid.

[2] as became clear at the work meetings in TNO project FlexiQuest in cooperation with grid operator Alliander.

electricity is proportional to travelled distance. If all demand and production of power on the grid were deterministic and time-invariant (which they are most definitely not), transportation losses would be minimized by placing all generators as close to demand points as possible. But when demand and production are time-varying or stochastic, sometimes local production is higher than local demand. Local excess power in peripheral parts of the grid has to be transported further than if it were produced at some central point. So local placement of generators is advantageous when local demand is larger than local production, but disadvantageous when local demand is smaller than local production. In that last case central placement would be better. Therefore loss-minimizing placement of stochastic generators may be regarded as a kind of tradeoff between the advantages of local and central placement.

For wind power generation, also another aspect plays a role: the correlation between the stochastic productions of wind turbines reduces when they are placed further apart. So a certain spreading of the wind mills will result in smaller variance of total wind power production, and therefore may reduce the disadvantage of its stochastic nature. However, spreading of the wind mills may have to mean that some are placed in regions with less wind, which in itself is undesirable. Also in this respect a tradeoff has to be made.

1.2 Research question

The research question for this thesis is formulated as follows:

Given a normal load situation of the Dutch HV network, what would be the optimal locations to build a certain number of new wind mills, in order to minimize the expected energy transportation losses?

So in this way regarding a fixed load situation of the Dutch Power Grid, and then placing a number of windmills on their optimal locations, the result will show some distribution pattern of those wind mills over the Netherlands. This pattern will give an idea about the optimal way to distribute wind mills over the Netherlands (from a transmission efficiency point of view). We could compare this to figure 1, which shows the locations of the current large wind mills in the Netherlands (source: [Enslin et al. 2003]), to see if the results match the current construction practices.

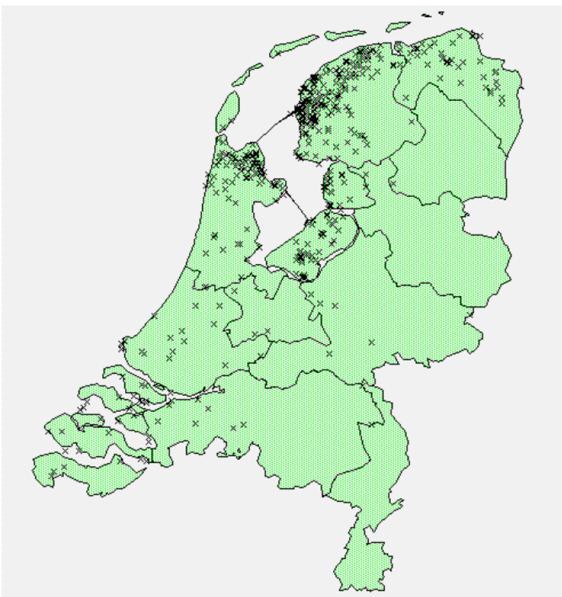


Figure 1: Large wind mills in the Netherlands

The research question is addressed in three steps: first, by modeling transportation losses and the study of power flow which is required for that (chapter 3); second, by modeling the stochastics of wind power production by means of simulation (chapter 6); and third, by combining the former two in a stochastic optimization model (chapter 4 and 5). An outline of the order in which these subjects are treated, is provided in the next paragraph.

1.3 Structure of the thesis

In chapter 2 an overview is given of the scientific literature connected with the topic of this thesis. It is intended to give an impression of the directions of current research in the field, and of the relative position of this thesis.

In chapter 3 a short introduction to power grids is provided. Then a simplified model is derived for the performance of energy transportation through the power grid. The principles of electronics are used to arrive at the technical description of power grid behaviour: the so called power flow equations. Some simplifications are applied in order to derive a model in terms of network flow and mathematical programming. The chapter was written such as to conform to the intuition of someone active in the field of optimization. However in order to be rigorous and to give a complete vindication of the simplifications, it was necessary to derive the model in power engineering terminology. Someone unfamiliar with it could skip section 3.3.

Chapter 4 first introduces the model to answer the research question in a general manner, discussing some relevant aspects and choices made therein. It then goes on to formulate the model in a precise mathematical manner, distinguishing two stages, and adopting terminology from graph theory and mathematical (stochastic) programming.

In chapter 5 the mathematical optimization theory is discussed by which the model formulated in chapter 4 can be brought into a form that is solvable by standard optimization software. This is done by treating the way in which the stochasticity of the problem can be handled, and how network flow duality is applied in order to formulate the problem as a standard optimization problem.

Chapter 6 is on wind power modeling. The optimization model derived in chapter 4 and 5 introduced some parameters which represent the uncertainty contained in wind power generation. The purpose of chapter 6 is to perform a simulation in order to estimate these parameters. A detailed account is given of how this is done, and justifications are provided by statistical testing and refering to scientific literature.

In chapter 7 the way is described how the model is implemented into the optimization software package Aimms. Because I regard knowledge of Aimms as very useful for a professional in operations research, and because it requires some time to get used to the software, I dedicate some space to the issues one is confronted with when implementing a practical optimization model of this scale.

Chapter 8 is devoted to heuristics. Because the rigorous optimization quickly becomes too large to perform, and because there is quite some structure in the problem that it does not make use of, two heuristics are presented that heavily rely on network structure and common sense.

Chapter 9 discusses the results. The first paragraph considers only wind, and shows which wind turbine locations would be favorable from this point of view. Paragraph 9.2 considers only the Dutch HV network, and shows which locations would be favorable for new production units in general (not depending on wind). Paragraph 9.3 answers the research question, which is shown to be a tradeoff between the results of paragraphs 9.1 and 9.2.

Chapter 10 discusses the relative significance of the thesis topic to problems that the energy sector copes with in practice. The research question on itself is not that crucial, but it is shown that there are several problems of a very comparable structure which may become relevant very soon.

In chapter 11 a summary of the thesis is provided, and conclusions are drawn. Some opportunities for further research are suggested.

2 Literature survey

In order to give an impression of the relation of this research thesis to the topics that are treated elsewhere in scientific literature the contents of some recent publications are discussed. These articles are clustered around the themes DC modeling of power flow (which is relevant for chapter 3); network design for loss minimization (which is the subject of the whole thesis); wind power modeling (which is relevant for chapter 6); the integration of wind power in the electricity system (chapter 4); and the increasing use of stochastic programming in general, and in the energy field in particular (chapter 5). Many technical terms which are used in this chapter will be explained only in later chapters (particularly chapter 3 and chapter 10) in order to keep technical explanation and discussion of literature separated.

2.1 DC power flow model

In this thesis, the so-called DC power flow model will be derived and used (chapter 3). Because in recent times this model has regained much of its old popularity, quite some new papers appear on the topic after a period of silence. In former times the full AC model was too hard to solve, and therefore the DC model was used instead. [Crevier 1972] treats the DC model and even more simplified models for power flow analysis. Then by increase of processor speed the AC models became solvable, and DC models were abandoned. But now several new developments in power grid operations have come up which demand such speed in solving for the power flows, that the DC model is coming into use again. A few of these applications are: Locational Marginal Pricing (LMP), where electricity prices on the market are influenced by network congestion; real time generation dispatch, that is the cost optimal way to generate the desired amount of power by a set of (different) power plants; day ahead unit commitment; and the use of stochastic programming techniques (which are more complex than deterministic programming) to address problems in network design and generation expansion planning under uncertainty.

[Purchala et al. 2005] provide a quality assessment of the DC power flow model. First they discuss the current uses of the DC model, and its relations to other power flow models; then they formulate the model and treat the assumptions under which it is assumed to be realistic. In the main part of the article they do both AC and DC power flow computations for the same load situation in the Belgian HV network. By comparing the DC solutions to their AC equivalent (which they assume exact), they try to come up with guidelines on how precise the initial assumptions of the DC model should be fullfilled, in order to remain within an 5% error bound. Their most significant result is that the quality of the DC model is especially sensitive to a flat voltage profile over the buses in the grid.

[Stott et al. 2009] give the same reasons as [Purchala et al. 2005] for studying DC power flow. They remark that although the DC model is very well known, and that several different versions of the model are being used, yet almost nowhere there is a classification of the different models. The main purpose of the publication is to classify and improve the DC model at a fundamental level. The authors give a very clear overview of the practical and theoretical advantages of the DC model, and discuss the historical background and the recently appeared scientific literature. Next they start to derive the model from the basics in a concise manner, and quickly arrive at two slight generalizations or improvements on the model most commonly in use. These improvements are based on better impedance parameter choice, and on the incorporation of the line losses. In the second part of the article they discuss the advantage of using a prior AC solution of the power flow problem to do additional computations based on the DC model. They call this a *hot-start model* in contrast with the *cold-start model* where no AC basic solution is used. They find that when possible, it is very advantageous to start from an AC solution. Also they discuss the DC modeling of phase shifting transformers, HVDC (High Voltage components which exploit Direct instead of Alternating Current) and FACTS devices (Flexible AC Transmission System devices in order to manipulate power flow distribution through the network), and security-constrained optimal power flow formulations. This paper is very well written and contains the state of the art in DC power flow modeling.

[Grijalva et al. 2003] are concerned with the DC power flow model because of "Transfer based electricity markets" and congestion analyses. Based on a current load scenario they need fast and accurate algorithms to determine how much more power can be transported between parties on the grid. The term for this analysis is Available Transfer Capacity (ATC) calculations. The ATC calculations based on the DC power flow model are called linear. The authors of this article observe that the DC model leaves out reactive power altogether, and they propose an improvement for that. Then they display a resourceful manipulation with the power flow equations and derive an upper bound for the reactive power flow over a transmission line. They summarize by saying that reactive power flow can be implicitly handled in linear ATC calculations by adjusting the heat capacities of the branches, and then performing the regular computations. The rest of the article is devoted to testing how large the practical improvements of this adjustment are. This article is of interest when the DC model is used for analyses in networks where congestion plays a role.

2.2 Network design for loss minimization

When searching the scientific literature for minimization of power losses in a network, it becomes clear that there exists a long history of research on this topic, and that it is now vivid as ever. The major part of the research papers is on very technical subjects such as the optimal placement of capacitors (which serve as correctors for reactive power flow) on a distribution line, or the optimal configuration of sectionalizing switches in a near-radial distribution system. Particular interest exists for the lower voltage networks, which show the highest power losses in the system. Often these networks contain no cycles (because they form the last link to the individual customers), and therefore many algorithms are proposed for so called *radial distribution systems*. Reactive power flow is very essential to these networks, so DC power flow models are not mentioned in these articles.

[Ababei and Kavasseri 2011] address the old question of network reconfiguration, and come up with an improvement on the classical method of [Baran and Wu 1989]. They give a faster algorithmic implementation on the loss estimating technique DistFlow by [Baran and Wu 1989], and propose a better heuristic for trying new configurations.

[Srinivasa Rao and Narasimham 2008] is an example of a recent article which shows that the research for loss-minimizing capacitor placement is still going on. [Wang and Nehrir 2004] and [Le et al. 2007] show an interest in using Distributed Generation (that is, small scale power generation units on the low voltage parts of the grids) to minimize tranportation losses. They motivate their topic by saying that an ever increasing power demand asks for capacity expansion on the distribution lines, which can be very expensive especially when long cables have to be replaced. Installing small local generators gives the opportunity to keep up with the demand without increasing the capacity of the transportation cables. The cited articles explore the possiblities to design and locate these small generating units such that line losses are as small as possible, and voltage quality as high as possible. [Quezada et al. 2006] show that these Distributed Generation (DG) units do not always have a positive effect on the energy losses when penetration becomes too high (because then of course some locally generated energy has to be transported back to the main grid).

[Le et al. 2007] note that traditionally this loss minimizing use of DG had nothing to do with small scale renewable energy generation units like wind turbines or photovoltaic cells. It was about local small fossil fuel generators in order to avoid capacity expansion, or transportation losses. Then they note that renewable energy sources are currently becoming more apt to be used for some of the same purposes (mainly reducing energy losses, but also increasing reliablity of supply). However such uses have the greatest relevance in large countries where long distances have to be covered at a lower voltage. In the Netherlands this is typically not the case. The costs of a DG unit would easily be higher than the costs for capacity expansion of the grid when distances are so small.

2.3 Wind power modeling

The field of renewable energy is constantly growing, and wind power is one of its largest contributors. A lot of research is done into wind power modeling under two main themes: firstly, *steady state analysis*, which is mainly used in long term planning, power grid design and generation expansion studies; and secondly, dynamic analysis (or, more commonly: *forecasting*). This is used for short term prediction of future wind power generation, in order to optimize intraday market biddings and short term power system control. Especially because of the reactive power behaviour of wind turbines, voltage stability in the system may depend crucially on their correct operation.

[Feijoo et al. 1999] mention the growth of wind energy as a reason for investigating good simulation methods. They cite the International Electrotechnical Commission to vindicate their Rayleigh distribution assumption of single site wind power. Next they discuss two methods of simulation for the joint behaviour of several wind farms on different locations, based on measurement data. The first method is based on the transformation of independent simulations by a Cholesky decomposed correlation matrix. The second method is based on sampling from measurement data, and does not rely on any prior assumption on wind speed distribution (it is *nonparametric*). They discuss the respective advantages and disadvantages of the methods, and treat very shortly in what manner such simulations should be used. Also they describe how to model the reactive power with respect to a wind farm: they advise to use the asynchronous machine model (which relates active power generation behaviour to a fixed reactive power generation behaviour). In chapter 6 of this thesis their first simulation method is described in greater detail when it is applied to Dutch wind measurement data.

[Papaefthymiou and Klöckl 2008] are concerned with the time dependent output behaviour of a single wind turbine. It is their aim to contribute to a "simple stochastic black-box model for wind generation". They first discuss the practical importance of having good wind forecasts. Then they go on to argue that a direct simulation of wind power output yields a result with better statistical properties than first simulating the wind speed, and then transform to wind power by the power curve of the wind turbine. This last method is currently the common one. Because many different states from the wind speed domain, are mapped into the same state in the wind power domain, fewer parameters have to be estimated when directly simulating wind power output. This results in faster convergence of the stochastic model parameters. After shortly discussing standard ARMA Time Series models of wind power, they describe the Markov Chain description of wind speed and power. States are defined, Markov Chain basics are treated. Because first order Markov Chains do not yield correct auto correlation functions (acf) of the resulting time series, higher order Markov Chains should be used. Then the authors describe their Markov Chain Monte Carlo (MCMC) simulation method of wind power. They show that fewer states and a lower order of the Markov Chain are needed to achieve the same quality in pdf and acf of the output (as opposed to transformed wind speed simulation).

2.4 Power system expansion

Above scientific literature on the separate topics of power flow, loss minimization and wind power modeling were considered. Below two publications are discussed which combine the above topics, namely the inclusion of wind turbines in an existing grid. This thesis is on the optimal placement of wind turbines with the sole purpose of loss minimization. On this specific topic there is no literature to be found (as to the reason why, is discussed in chapter 10), but it clearly bears similarity to these publications.

[Carpinelli et al. 2001] address the question of "distribution system planning in the face of a worldwide growth of DG penetration". The aim of their article is to present a "proper tool, able to find the siting and sizing of DG units which minimize generalized cost" as a help for the distribution system planner. They begin their paper by summarizing why they expect the increasing use of Distributed Generation to continue. They mention for example: reduction of transportation losses because DG is closer to customers; recent technology has succeeded in building efficient small scale generators; it is easier to find sites for small generators; DG plants require shorter installation times, so involve less investment risk by being more flexible; electricity market liberalization has increased opportunities for new power generating companies; by increasing load of the grid, transportation cost increase, whereas DG costs are dropping – so that the added value of DG is growing[3].

The authors then go on to say that they designed a three-step algorithm to find the optimal locations and sizes of several wind turbines, to be built within the area of some Medium Voltage (MV) distribution grid. The first step consists of an implementation of the method of [Feijoo et al. 1999] discussed above; the second step is a genetic algorithm which steadily improves the intermediate solution. The objective function (generalized cost) contains several terms, under which location dependent building costs, but also the cost of power losses. The third step is a method of determining the robustness of a solution under uncertain future conditions. Lastly, they apply the method to a particular Italian MV grid. This article is interesting because it discusses a problem similar to my research question, and because it places location selection for wind turbines so explicitly within the context of increasing DG penetration.

[Li Yang Chen 2010] wrote a PhD thesis on stochastic network design. He takes his applications from design in logistic networks and power systems. The subject of chapter 4 is Transmission and Generation Expansion Planning (TGEP) for wind farms in power systems. In the chapter introduction he discusses renewable portfolio standards in the United States, and the growth of wind energy as a consequence. He thinks it particularly characteristic for wind generation expansion that location selection is so dominant. By the "highly stochastic and location dependent" nature of wind speeds, a spreading of wind farms is attractive from a generation point of view. He shows that the power output of wind farms that are distributed over multiple sites has significantly lower variance than when the same capacity is built on one site. But because the best sites for wind farms are often far removed from the existing grid, the required extension of the grid is often very expensive for wind energy. Therefore the selection of multiple sites for wind farms is less attractive from a transmission point of view. A consequence of this is that for wind energy, the planning of generation capacity and of transmission capacity should be considered simultaneously. He shows that however very little work has been done on the simultaneous TGEP problem. Mostly generation and transmission expansion are regarded independently, or else a deterministic TGEP model is used. But wind power needs a stochastic model.

Next he presents a two stage stochastic programming formulation of the integrated TGEP problem. The first stage decisions represent the design decisions: where to build generators, and of which type; and where to construct new transmission lines (or upgrade them when there already is one), and of which type. The scenarios reflect all randomness in availability of generator capacities and transmission line capacities and in the demand at the demand nodes. (Reduced availability of generators and transmission lines reflect a certain load scenario of the grid. This load is not explicitly involved in the computations, but it is taken into acount by reducing the availabilities.) The second stage decisions subsequently represent the actual generation at the generation nodes and the power flows over the lines. Second stage variables are constrained by installed generation and transmission capacities, and by flow balance at the nodes. First stage decisions are constrained by renewable portfolio standard requirements (i.e., enough wind energy). System reliability is promoted by putting a penalty on high expected loss-of-load in the objective function. Other terms in the objective are the installation costs of the expansions, and the expected operation costs. Operating costs of a generator for a certain scenario is a linear function of the generation. Summing over all generators and scenarios yields the expected operating cost.

Then the author shows how the quadratic network losses may be linearized by a similar transformation as the one that is applied in this thesis in chapter 5, and he discusses how the L-shaped method by [Van Slyke and Wets 1969] can be applied to the resulting large scale linear program. Because this standard version of Benders' decomposition algorithm performs particularly poorly

[3]Here it must be remarked that although this is true in itself, it only provides commercial opportunities if transportation losses are paid by the parties that cause them. In the current situation in the Netherlands this is not the case: all involved parties share in proportion to their production or consumption. When market design becomes more sophisticated (LMP), it may however become a feasible scenario.

for the wind farm network design problem, the last part of the chapter is devoted to the acceleration of the algorithm. This is achieved by adding quasi-network structure to the restricted master problem. First stage flow variables are added, which yield a rough estimate of the second stage (actual) flow variables. They serve to enforce network connectivity and to make sufficient generation and transmission capacity likely. Moreover knapsack-like constraints are added to ensure faster convergence in subsequent iterations. Computational results show that these improvements have significant impact on the performance of the algorithm. The chapter is ended by summary and conclusions. One of the avenues for future research mentioned, is the extension of the model to a broader class of facility location problems under uncertainty.

2.5 Stochastic programming

Currently there exists a broad interest in the application of Stochastic Programming (SP) techniques. This is displayed in the scientific research of a wide variety of areas. Of course in logistics, goods have to be transported when demand is still uncertain, but also market value of electricity for producers is uncertain, future developments of total energy demand in a country may grow slower or faster, but the extension of generation capacity has to be planned ahead. A few interesting publications in this context:

[Hasche et al. 2006], for example, apply stochastic programming on trading strategies based on forecasts of generated wind power. They try to evaluate the economic merit when wind power forecasts improve.

[Santoso et al. 2005], in a completely different field, describe the application of stochastic programming in the design of supply chain networks. (The link between transportation networks and power networks is discussed in chapter 3 and 4.)

[Wallace and Fleten 2003] give a very useful and illuminating overview of SP models that are used in the entire energy sector. Most applications are in electricity, but also some in gas and oil.

[Shapiro et al. 2009] provide a readable introduction as well as the state of the art in Stochastic Programming. They treat recent developments such as the statistical properties of the Sample Average Approximation (SAA). SAA is the optimization over a random sample of the scenarios. When scenario sets are too large this technique may be required.

3 Power grid modeling

In this chapter a short introduction into power grids is given. Then the power flow equations are formulated, energy loss is described, and by several steps a simplified model for power grid behaviour is derived.

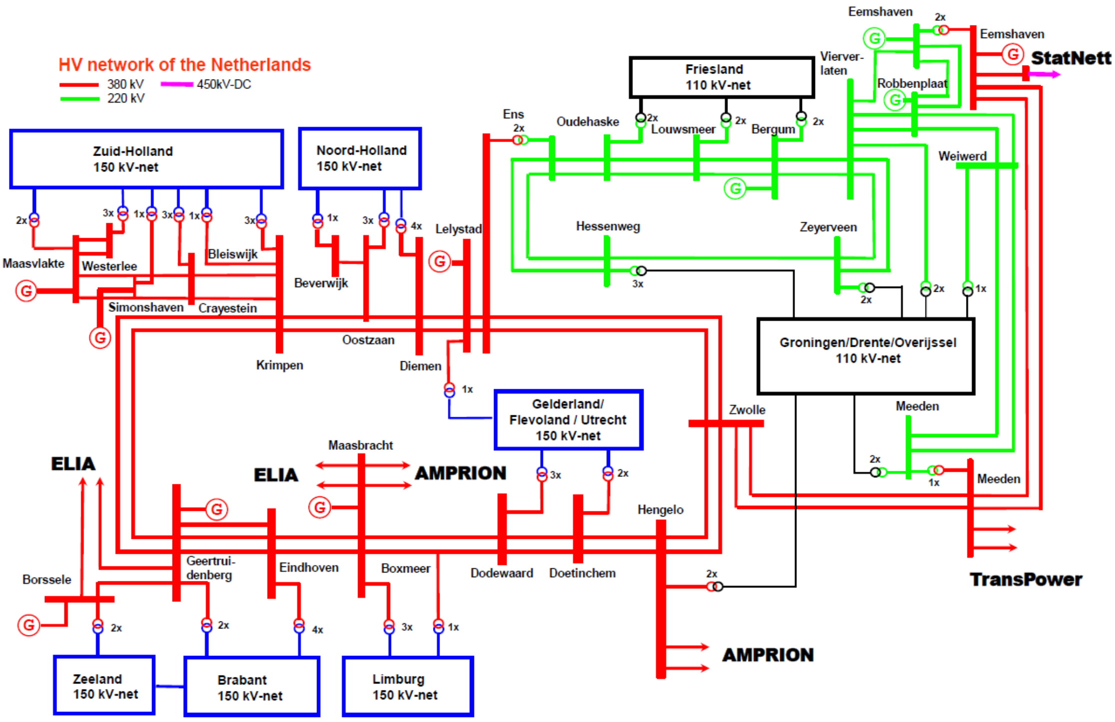


Figure 2: Dutch HV Network

3.1 Power grids

A *power grid* is a set of electrical components called *buses*, which are connected with each other by transmission lines, called *branches*. Some of the buses represent electricity consumers, which have a demand for electrical power. These buses are called *load buses*. Some other buses represent power plants, or other sources of electrical power. These buses are called *generation buses*.

In order to keep the voltages through the entire grid at their demanded level, the system is operated in such a way that supply equals demand at every point in time. As demand varies in time depending on the wishes of the consumers, power generation has to be adjusted accordingly.

Due to resistance of transmission lines, electrical energy is lost during transportation from generation to load buses. These losses are reduced by applying higher voltages, but electricity has to be delivered to consumers at only $230V$. As a result, most power grids are designed in the following way:

- **High Voltage grid**, also called *transmission grid*, because it was designed for transmission over long distances. It is kept at voltages of $110kV$ to $450kV$, it contains all large power plants, and some connections abroad. It was implemented as a number of interconnected double rings. The HV grid of the Netherlands is operated by TenneT (`www.tennet.org`).
- **Medium Voltage grid**, being the link between transmission and distribution. It is kept at voltages between $1kV$ and $110kV$, it is connected to the HV grid by so called *transmis-*

sion stations, and it contains large consumers like process industry and the larger kind of distributed generation like wind turbines and windfarms. MV grids contain partly overhead cables, partly underground cables.

- **Low Voltage grid**, also called *distribution grid*, because its sole purpose is to distribute electricity to the end users. Voltage is below $1kV$, and its network structure is called *radial* because there are no rings and cycles in it. It is connected to the MV grid by so called *distribution stations.* The MV and LV grids in the Netherlands are operated by local (non-commercial) utilities, of which Alliander, Enexis, StedIn and Delta are the biggest.

In figure 2 a schematic overview is shown of the HV parts of the Dutch power grid, which is operated by TenneT. It clearly points out the hierarchical and regional character of the grid.

3.2 Electricity transportation

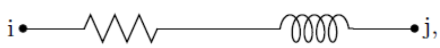
Power systems use alternating current (AC) to transport electrical energy. The main advantage of AC over direct current (DC) is that it can easily be transformed to other voltages, which is clearly very essential to a large scale power grid. Although the same physical laws of electricity apply, analysis of power systems is completely different from electrical circuit analysis.

Because in a system with alternating current all relevant quantities show (ideally) sinusoidal behaviour in time (oscillating around zero), and because calculations with sinuses quickly become very cumbersome, it is common to represent all physical quantities by complex numbers. The calculations become a lot easier, and one can take the real part afterwards.

Note that by writing complex numbers in polar coordinates, it is clear that this way of writing is capable of expressing *magnitude* of a quantity (representing its value at the points in time when it assumes its maximum), and *phase* of a quantity (representing the stage of its oscillation it is in at some reference point in time).

3.3 Power flow equations and power loss

The usual physical representation of a transmission line is:



that is, an ideal line (ij) containing a *resistance* r_{ij}, and a *reactance* x_{ij}. The *impedance* of this line then, is $Z_{ij} = r_{ij} + ix_{ij}$[4]. This means that the relation between current and voltage in the line can be described by Ohm's law $I_{ij} = \frac{V_i - V_j}{Z_{ij}}$. Assuming a complex voltage of $V_i = v_i e^{i\theta_i}$ at bus i, and of $V_j = v_j e^{i\theta_j}$ at bus j, yields exercised powers of

$$\begin{aligned} S_{ij} &= V_i I^* = V_i \frac{V_i^* - V_j^*}{Z^*} = \frac{v_i^2 - v_i v_j e^{i(\theta_i - \theta_j)}}{r - ix} \\ S_{ji} &= \frac{v_j^2 - v_i v_j e^{i(\theta_j - \theta_i)}}{r - ix}. \end{aligned}$$

Note that for ease of notation from here on r and x will be written in stead of r_{ij} and x_{ij}, because the context makes clear which resistances and reactances are meant.

Adding the complex powers exercised by voltages V_i and V_j at both ends of the line yields a total complex power conversion of

$$S_{tot} = S_{ij} + S_{ji} = \frac{v_i^2 + v_j^2 - v_i v_j (e^{i(\theta_i - \theta_j)} - e^{i(\theta_j - \theta_i)})}{r + ix}.$$

[4]capitals are used to represent complex quantities.

The magnitude $|S_{ij}|$ of a complex power is called the *apparent power*. Because of Joule's law $Q = I^2 \Omega t$, it is the current which determines the capacity of a transmission line (before it is damaged by heat production). Moreover for fixed voltage magnitude, the apparent power $|S_{ij}| = |V_i I^*| = |V_i||I|$ is proportional to the current on a line. Therefore in a power grid line capacities can be expressed in maximum apparent power.

But the apparent power says nothing about the power conversion in the line. That can be seen by taking the real parts of the above expressions, representing the *real powers* exercised by voltages V_i and V_j:

$$
\begin{aligned}
p_{ij} &= \frac{1}{r^2 + x^2}[rv_i^2 - rv_i v_j \cos(\theta_i - \theta_j) + xv_i v_j \sin(\theta_i - \theta_j)] \\
p_{ji} &= \frac{1}{r^2 + x^2}[rv_j^2 - rv_i v_j \cos(\theta_i - \theta_j) - xv_i v_j \sin(\theta_i - \theta_j)],
\end{aligned}
$$

and adding them yields a total real power conversion of

$$
p_{loss} = p_{ij} + p_{ji} = \frac{1}{r^2 + x^2}\left[r(v_i^2 + v_j^2) - 2rv_i v_j \cos(\theta_i - \theta_j)\right].
$$

This formula represents energy loss per second, depending on complex voltages on both ends of the line.

If we take the imaginary parts of the complex powers, we get the so called *reactive powers*

$$
\begin{aligned}
q_{ij} &= \frac{1}{r^2 + x^2}[xv_i^2 - rv_i v_j \sin(\theta_i - \theta_j) - xv_i v_j \cos(\theta_i - \theta_j)] \\
q_{ji} &= \frac{1}{r^2 + x^2}[xv_j^2 + rv_i v_j \sin(\theta_i - \theta_j) - xv_i v_j \cos(\theta_i - \theta_j)].
\end{aligned}
$$

And these four equations describe the total (steady state) power grid behaviour, because Kirchhoff's Current Laws demand power flow conservation at every bus i for real power (that is, $\sum_j p_{ij} - \sum_j p_{ji} = 0$), and also for reactive power (that is, $\sum_j q_{ij} - \sum_j q_{ji} = 0$). By specifying certain boundary values at the buses (for example, p_{ij} and v_i at a generation bus i), and one reference node with phase $\theta = 0$, one can then solve for all voltage magnitudes and phases in the system. This is called AC power flow calculation.

3.4 DC load flow model

For technical purposes it is often required to be able to calculate the behaviour of bus voltage magnitudes and angles. It is important to maintain voltage stability, even when power demand and generation is subject to change. Unexpected reactive power behaviour may result in local voltage drops (which is undesirable for customers), or voltage rises (which may damage the electrical components of the system). Therefore the full solution of the above non-linear equations is critical for the technical operation of the system. However for general design questions and performance estimations, it is expedient to simplify a little further. [Purchala et al. 2005] justify the following assumptions for the context of this thesis:

1. The bus voltage magnitudes v_i and v_j are almost equal, so that we can say $v :\approx v_i \approx v_j$. That is, the effect of voltage drop on power system behaviour is neglected. Certainly for high voltage systems this is justified.

2. The difference in voltage angles $(\theta_i - \theta_j)$ is small. In this case one can make the first order approximations $\sin(\theta_i - \theta_j) \approx (\theta_i - \theta_j)$ and $\cos(\theta_i - \theta_j) \approx 1$.

3. An unnecessary assumption for our modeling purposes, but nevertheless true in practice for HV transmission lines, is $r << x$ (resistance ismuch smaller than reactance).

Making these approximations, and putting $v = 1$ (which is called expressing the voltage in per-unit) yields for our real power equations

$$\begin{aligned} p_{ij} &= \frac{1}{r^2+x^2}[rv_i^2 - rv_iv_j\cos(\theta_i-\theta_j) + xv_iv_j\sin(\theta_i-\theta_j)] \\ &\approx \frac{1}{r^2+x^2}[rv^2 - rvv + xv^2(\theta_i-\theta_j)] \\ &= \frac{x}{r^2+x^2}(\theta_i-\theta_j) \approx \frac{1}{x}(\theta_i-\theta_j) \\ p_{ji} &\approx \frac{x}{r^2+x^2}(\theta_j-\theta_i) \approx \frac{1}{x}(\theta_j-\theta_i). \end{aligned}$$

One can see that due to $v_i = v_j$, the first and second term in square brackets cancel out. The justification of this approximation for HV power grid is even more enhanced by the third assumption that x tends to be (much) larger than r, making the third term dominant already.

Now compare a purely resistive DC electrical circuit where the only active elements are current sources connected to ground. In this case the current flow equations are $I_{ij} = \frac{1}{R_{ij}}(V_i - V_j)$ for all lines (ij).[5]

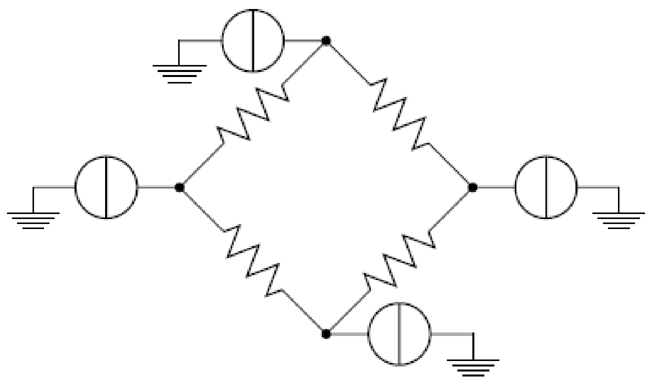


Figure 3: DC electrical circuit

Note the similarity between the simplified power equations of the AC power grid, and the current equations of the DC electrical circuit. Their structure is exactly the same, only quantities differ. Thus it can be seen that after these simplifications a remarkable analogy exists between a large AC power system, and a DC electrical circuit equivalent. The role of currents in the DC load flow is taken over by powers in the AC network; the role of DC voltage drops over resistors correspond to differences between bus voltage angles, and DC resistance values correspond to reactance x of a transmission line.

It is according to this analogy that we like to think about the power grid as if there are power sources, power sinks, a power flow through the network, (even 'power conservation laws'), *en passant* producing line losses.

[5]here capitals represent normal real DC quantities.

3.5 Reactive power and line loss in the DC model

The same approximations 1) and 2) applied to the reactive power equations, yield:

$$\begin{aligned} q_{ij} &= \frac{1}{r^2+x^2}[xv_i^2 - xv_iv_j\cos(\theta_i-\theta_j) - rv_iv_j\sin(\theta_i-\theta_j)] \\ &\approx \frac{1}{r^2+x^2}[xv^2 - xvv + xv^2(\theta_i-\theta_j)] = \frac{-r}{r^2+x^2}(\theta_i-\theta_j) \\ q_{ji} &\approx \frac{-r}{r^2+x^2}(\theta_j-\theta_i) = \frac{r}{x}p_{ij}, \end{aligned}$$

that is, if also $r << x$, reactive power 'flows' vanish, or at least are neglectable in comparison to real power flows.

Similarly, the first order approximation for real power loss yields $p_{loss} = p_{ij} + p_{ji} = 0$. The DC model assumptions formulated at the begining of the previous paragraph have as a direct consequence that zero power loss is assumed. This can easily be verified by the consideration that for all i and j, $v_i = v_j$, which means that there is zero voltage drop in the system, and consequently there can be no net power conversion. But taking the second order approximation $\cos(\theta) \approx 1 - \frac{\theta^2}{2}$ produces

$$\begin{aligned} p_{loss} &= \frac{1}{r^2+x^2}\left[r(v_i^2+v_j^2) - 2rv_iv_j\cos(\theta_i-\theta_j)\right] \\ &\approx \frac{1}{x^2+r^2}\left[2r - 2r(1-\frac{(\theta_i-\theta_j)^2}{2})\right] \\ &\approx rp_{ij}^2. \end{aligned}$$

These powers do not actually disappear from the power flows in the model (which are, like the current in an electrical circuit, preserved), and therefore do not influence the power flow solution.

3.6 Network flow model

Above it was established by simplifying the exact AC power flow equations, that large scale AC power flow behaves approximately like a simple small scale DC electrical circuit. In former times this similarity was used by the old "DC network analyzer", in which each network branch was represented by a resistance proportional to its series reactance and each DC current was proportional to a real power flow. The DC model derived its name from this analogue computing table.

These days of course computer algorithms can be exploited. By basic physics of electricity, Kirchhoff's Current Laws (KCL) and Kirchhoff's Voltage Laws (KVL) yield the equations by which the solution for the nodal voltages and currents over the lines can be computed. KCL demand that the total current into a node must equal the total current out of that node. KVL demand that the directed sum of the voltage drops over every closed loop in the network must equal zero. By simple linear algebra techniques the resulting system of linear equations can be solved. In power flow analysis, the matrix of coefficients of the linear equations is called the *admittance matrix*. The solution procedure comes down to inverting this matrix.

In the context of this thesis (stochastic and constrained optimization) another approach is more useful and also more intuitive. [Crevier 1972] remarks that the currents found by solving both Kirchhoff's Current and Voltage Laws, happen to be the same currents that minimize the total heat dissipation described by $\sum(I_{ij})^2r_{ij}$, while satisfying only Kirchoff's Current Laws.

Because in the terminology of network flow optimization these current laws are just the flow balances at a node; and because the objective function that describes total heat dissipation is clearly separable and convex (quadratic), we know from theory on network flows like [Ahuja et al. 1993] that the power flow problem can be solved efficiently by network flow algorithms. The advantage of this method over the traditional former one is that network algorithms allow for straightforward sensitivity analysis, can cope with additional contraints that may come in, and offer a known framework for a stochastic extension.

Concretely, this means that to find the solution for the currents through a DC network, and therefore an approximate solution for the power flow in an AC power system, the following mathematical program should be solved:

$$\begin{aligned} \text{minimize} \quad & \sum_{(ij)\in A} p_{ij}^2 c_{ij} \\ \text{subject to} \quad & \sum_{j:(ij)\in A} p_{ij} - \sum_{j:(ji)\in A} p_{ji} = b_i \quad i \in N, \\ & p_{ij} \geq 0 \qquad\qquad\qquad\qquad (ij) \in A, \end{aligned}$$

where p_{ij} is the power flow through line (ij), c_{ij} is the resistance of the line, and b_i is the power supply/demand (the *balance*) of each node i.

From this it can be seen that the problem of computing the power flow solution of a power grid, is equivalent to solving a so called minimum cost flow problem with quadratic arc costs. Recall (e.g. from [Ahuja et al. 1993]) that the flow balance constraints may be written as $B\vec{p} = b$, where B is the so-called *node-arc incidence matrix* of the network, $\vec{p}$ is the vector of power flows through the arcs, and b is the vector of supplies at the nodes. Because of notational conventions, in the next chapters we will use y_{ij} in stead of p_{ij} as symbol for the power flows.

3.7 Additional remarks

For the test case in this thesis equal voltages are assumed throughout the grid. But when this is not the case, the above network flow program can easily be adjusted to account for the voltage differences. Define all arc cost coefficients

$$c_{ij} = \frac{x_{ij}}{v_{ij}},$$

where v_{ij} is the voltage at which the line is kept. The intuitive interpretation of this is, that the high voltage lines are 'cheaper' to travel over, and therefore are more likely to attract flow than lower voltage lines. This is nice since higher voltage lines typically cause less power loss, and have higher capacities.

The application of network algorithms to solve the DC power flow problem seems to be little advocated in literature. Yet modern day algorithms such as network simplex and ϵ-relaxation can solve quadratic cost flows extremely efficiently. It has been suggested that these methods are faster than the admittance matrix inversion of classical DC algorithms (e.g., [Ahuja et al. 1993], chapter 1). Moreover, strongly polynomial algorithms have been developed for separable quadratic cost *generalized* flows. In such network problems, a so-called *arc multiplier* is added to all arcs. The outgoing flow of the arc is then defined as the incoming flow times the arc multiplier. The QRELAXG algorithm, together with mathematical proof of its polynomiality, is presented in [Tseng and Bertsekas 1996].

This opens up the possibility to a DC algorithm where power losses are incorporated by a linear (first order) approximation of the losses over a branch. Compare [Stott et al. 2009] where a zero'th order approximation is suggested for the losses by subtracting fixed loss estimates from the node balances. There it is proposed (section VII-B) that incorporation of the losses could be achieved by estimating all powerflows p_{ij} beforehand by $\hat{p}_{ij}$. Then $r_{ij}\hat{p}_{ij}^2$ estimates the losses over this line. These losses can be subtracted from the node balances of the adjacent nodes before computing the power flows. According to [Stott et al. 2009], the losses "usually converge" when this process is iterated (using the resulting power flow as the new estimates $\hat{p}_{ij}$).

The quadratic cost generalized flow however, immediately yields losses that are proportional to the power flow over the branch. To determine the arc multipliers one should still estimate the power flows beforehand. Then, based on the estimate $\hat{p}_{ij}$, put the arc multiplier $\gamma_{ij} = r_{ij}\hat{p}_{ij}$ such that the linear loss $\gamma_{ij}p_{ij} = r_{ij}\hat{p}_{ij}p_{ij}$ approximates the actual loss $r_{ij}p_{ij}^2$. The network programming formulation looks like:

$$\begin{array}{lll}
\text{minimize} & \sum_{(ij)\in A} p_{ij}^2 c_{ij} & \\
\text{subject to} & \sum_{j:(ij)\in A} p_{ij} - \sum_{j:(ji)\in A} \gamma_{ji} p_{ji} = b_i & i \in N, \\
& p_{ij} \geq 0 & (ij) \in A,
\end{array}$$

and take arc multipliers $\gamma_{ij} = r_{ij}\hat{p}_{ij}$. Iteration by using the resulting p_{ij} as the new flow estimates $\hat{p}_{ij}$, can be expected to converge both faster and more likely than the method of [Stott et al. 2009] because linear approximation of the quadratic loss function is significantly better than approximation by a constant function.

The model can be improved yet one step further. As will be shown in chapter 5, every quadratic network flow can be approximated to any desired degree of accuracy by a linear network flow in which parallel arcs are added. The flows over these parallel arcs sum up to the flow over the original quadratic cost arc. When decreasing arc multipliers are associated with these parallel arcs, a piecewise linear approximation of the quadratic loss can be realized. This is shown in figure 4. Thus it can be seen that the network flow approximation to the DC power flow problem is flexible enough to realize the incorporation of quadratic losses into the DC power flow equations. I have seen this remark nowhere else in literature. It would be interesting to assess if this improvement makes the DC power flow model more suitable for MV grids, or very large scale HV systems, where power losses are more substantial.

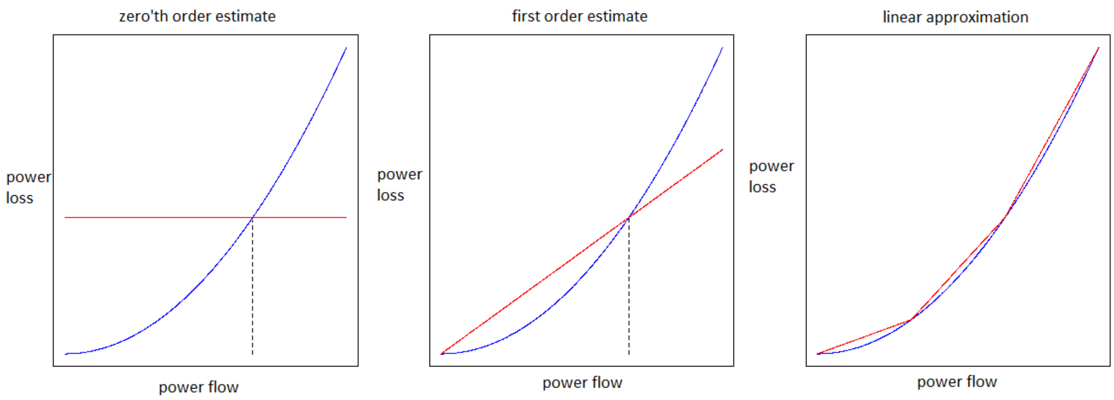


Figure 4: Relations of actual and estimated power loss on a line for the three methods. On the left: the method of [Stott et al. 2009]. In the middle: quadratic generalized network flow. On the right: linear generalized flow on the network with parallel arcs.

4 Mathematical model for the placement problem

In the previous chapter a simple network model was derived from power engineering principles. In this chapter the model is extended so that it includes the wind mills that should be placed on the grid. The goal is to arrive at a mathematical model of which the solution will directly answer the research question itself.

4.1 General outline

Because we expect HV and MV power grids to satisfy our DC load flow assumptions, and because the supply and demand of power have to be equal at every point in time, network behavior of the HV/MV network can be modeled by solving a minimum cost flow. The node balances required by such a model can be obtained from the load data of the power network in the Netherlands. These data can be taken from a moment of peak load or of average load. The connection of a wind mill to the grid corresponds to the increase of the supply at that particular node. In practice however, wind power production is highly uncertain. Although power demand by consumers also involves uncertainty, it is far better to predict than wind power production, and for that reason the assumption of deterministic demand seems justified. This assumption is also made by [Carpinelli et al. 2001].

In the figure below a rough sketch is provided, which may help to visualize the model described. In the original situation, power supply by the power plants (which are mainly connected to the upper parts of the grid) equals total demand by consumers (which are located mainly in the lower parts of the grid). The wind mills to be placed provide an extra (stochastic) supply.

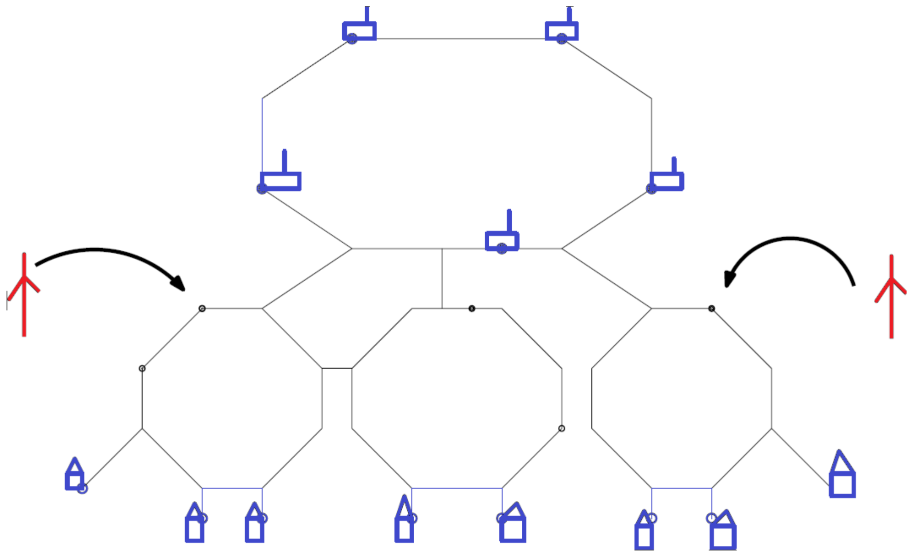


Figure 5: Schematic overview of placement problem

Of course, wind speed is geographically correlated: a strong wind in Den Haag makes a strong wind in Middelburg more likely. In order to model wind speed distribution over the Netherlands, we shall use wind speed measurements throughout the country to partition the Netherlands into several zones. Within these zones we will assume wind speed to be 100% correlated, and we shall use the measurement data to estimate the correlations between the zones. Justification of these assumptions is discussed in chapter 6.

A few things can be observed: first, because power supply and demand always have to be equal, a choice has to be made what to do with the excess power in the system, resulting from the wind turbines we place. A few possibilities are: sell excess power abroad; increase demand of some demand node(s); decrease generation of some supply nodes. In this last case a choice has to be made which power plants to adjust. Because in practice market balancing happens non-locally, we propose to adjust for every MW of wind power, certain other power plants distributed over the country.

Second: in general wind mills are placed 'lower' on the grid than normal power plants. So as long as wind power production does not exceed local demand, wind turbines will have a positive effect on the transportation losses.

Third: there is no intrinsic mechanism in the model that promotes placing the wind turbines on a spot with more wind than elsewhere (except when it can fulfill local demand). This might seem counterintuitive, but it is the consequence of our model. Models that minimize operating costs, do promote windy spots, because wind mills have high building costs but low marginal costs. Therefore windy spots will make more profitable investments.

Fourth: The model does promote spreading the wind mills over different zones, but only if peak wind power production is higher than local demand. In that case the peak spreading results in less power transport, and therefore less energy losses.

> In this network model terminology, the research question translates to finding the optimal location of some wind mills that have uncertain but correlated supply (depending on the geographical zone), such that the expected transportation losses are minimized. In order to maintain global balance, we adjust for the supply of these wind mills, by decreasing the supply of the original supply nodes.

In the following two sections, a step-by-step procedure is described to capture the model and the research question in a mathematical formulation. In section 4.2 the model is described for computing the power flow and line losses, once all power productions are given. This may be called the second stage of the problem, because wind mills have to be placed first. In section 4.3 the model is described for the placement of the wind mills, and for computing the produced wind power and node balance corrections. This may be called the first stage of the problem. Note that the second stage will be described prior to the first stage. The two stages and the stochasticity involved are represented intuitively in figure 6.

4.2 Power flow problem

Define an (undirected) graph $G = (N, A)$, taking a node $n \in N$ for every bus in the power grid, and an arc $a \in A$ for every transmission line connecting two buses.

Define a cost function $c : A \to \mathbb{R}^{\geq 0}$, which associates with every arc $a_{ij} \in A$ a cost c_{ij}, and take a cost c_{ij} equal to the reactance X_{ij} of its corresponding transmission line.

Define a balance function $b : N \to \mathbb{R}$, associating with every node i its supply of power b_i: a negative number for load buses, a positive number for generation buses.

Define a nonnegative power flow $y : A \to \mathbb{N}$ on the arcs of G, satisfying the balance equation $\sum_{j:(ij)\in A} y_{ij} - \sum_{j:(ji)\in A} y_{ji} = b_i$ for each node i, and $y_{ij} \geq 0$ for all $(ij) \in A$.

Assuming a power balance $\sum_{i\in N} b_i = 0$, the so called *power flow solution* is found by minimizing $\sum_{(ij)\in A} c_{ij} y_{ij}^2$ over all feasible power flows.[6]

Once the power flows have been found, $\sum_{(ij)\in A} y_{ij}^2 r_{ij}$ is the expression for the total power loss.

[6]Note that it is necessary that the balances sum up to zero, or, total power production equals total power consumption. This is because the DC power flow model assumes that no power losses occur. Power losses are neglected while computing the power flow solution. Afterwards, when the power flow solution has been computed, an estimate is made of the losses that have occurred.

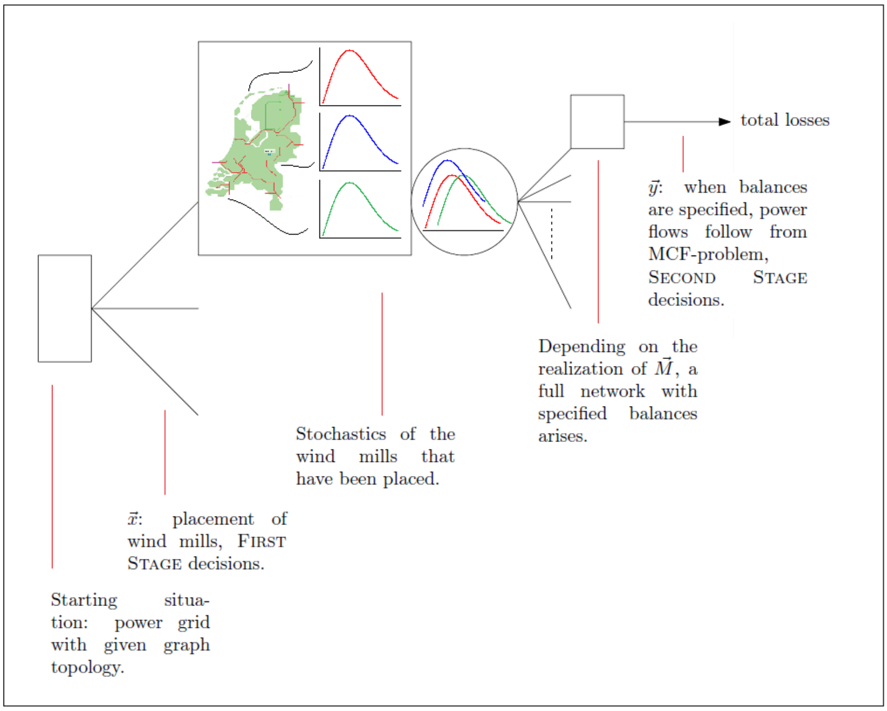


Figure 6: Illustration of the modeling steps

4.3 Wind turbine placement problem

Partition the node set N into k zones such that $N = N_1 \cup N_2 \cup \cdots \cup N_k$.

Introduce the random vector $\widetilde{M} \in \mathbb{R}^k_{\geq 0}$, representing the uncertain wind speeds in the zones at some arbitrary time instance; its k-th element $\widetilde{M}_k$ being the wind speed in zone k.

Transform $\widetilde{M} \in \mathbb{R}^k_{\geq 0}$ into $M \in \mathbb{R}^n_{\geq 0}$ by firstly, computing for each zone $i = 1, \ldots, k$ what wind power is produced by one wind turbine from the wind speed in zone i, and secondly, copying this number for each node in that zone. That is, if two nodes i, j are in the same zone, then $M_i = M_j$. We call M_i the *potential wind speed* at node i.

Introduce decision variables $x_i : i \in N$, representing the choice of placement for the wind turbine. Therefore for all regular $i : x_i \in \{0, 1, \ldots\}$ and $\sum_i x_i = l$ should be satisfied, where l is the number of wind mills to be placed.[7]

Now $\sum_{i \in N} M_i x_i = (\vec{M} \cdot \vec{x})$ represents the total wind power produced, so if $t \in \mathbb{R}^n$ represents the fractions of production adjustments to compensate for the wind power (with $\sum_{i \in N} t_i = 1$), then $b_i + M_i x_i - (\vec{M} \cdot \vec{x}) t_i$ is the new balance for any node $i \in N$.

Note that $\sum_{i \in N} \left[b_i + M_i x_i - (\vec{M} \cdot \vec{x}) t_i \right] = \left(\sum_{i \in N} b_i\right) + \left(\sum_{i \in N} M_i x_i\right) - \sum_{i \in N} (\vec{M} \cdot \vec{x}) t_i = 0 + \vec{M} \cdot \vec{x} + (\vec{M} \cdot \vec{x}) \sum_{i \in N} t_i = 0$, so total power balance is indeed always preserved, no matter what the value of $\vec{M}$ may happen to be.

Then given a placement vector $\vec{x}$ and a realisation $\vec{m}$ of $\vec{M}$, the vector of power flows $\vec{y}$ can be computed as follows:

[7] Do not confuse this decision variable x_i with the symbol x_{ij} or x, used in the previous chapter to denote line reactance. Notational conventions in power engineering demand the use of x_{ij} for reactance, and in mathematical programming it is customary to denote first stage decision variables by x_i.

$$
\begin{aligned}
\vec{y} = \text{argmin} \quad & \sum_{(ij)\in A} c_{ij} y_{ij}^2 \\
\text{subject to} \quad & y_{ij} \geq 0 & (ij) \in A \\
& \sum_{j:(ij)\in A} y_{ij} - \sum_{j:(ji)\in A} y_{ij} = b_i + m_i x_i - (\vec{m}\cdot\vec{x}) t_i & i \in N.
\end{aligned}
$$

It is clear that the power flows y_{ij} depend on the placement decisions x_i and on the random vector M, which determine the balances of the minimum cost flow. Hence, different choices of x_i will mount to different values of y_{ij}, and therefore different transportation losses. Moreover, the expected value of $y_{ij}(\vec{M}, \vec{x})$ over $\vec{M}$ is well defined. The aim then to minimize expected transportation losses by optimally placing the wind mills may be written in the following way:

$$
\begin{aligned}
\text{minimize} \quad & \mathbb{E}_M \left[\sum_{(ij)\in A} y_{ij}^2 r_{ij} \right] \\
\text{subject to} \quad & x_i \in \{0, 1, \dots\} & i \in N \\
& \textstyle\sum_i x_i = l.
\end{aligned}
$$

4.4 The complete model

Combining the two previous sections, a complete formulation is provided by

$$
\begin{aligned}
\text{minimize} \quad & \mathbb{E}_M \left[\textstyle\sum_{(ij)\in A} y_{ij}^2 r_{ij} \right] \\
\text{subject to} \quad & x_i \in \{0, 1, \dots\} & i \in N \\
& \sum_{i\in N} x_i = l \\
\text{and where} \quad & \\
\vec{y} = \text{argmin} \quad & \sum_{(ij)\in A} c_{ij} y_{ij}^2 \\
\text{subject to} \quad & y_{ij} \geq 0 & (ij) \in A \\
& (By)_i = b_i + m_i x_i - (\vec{m}\cdot\vec{x}) t_i & i \in N
\end{aligned}
$$

In the last equation (which represents the flow balance constraints), B is the node-arc incidence matrix for G. The right-hand side $b_i + m_i x_i - (\vec{m}\cdot\vec{x}) t_i$ are the balances, adjusted for the placement of the wind turbines.

List of symbols

Below the symbols used in chapters 4 and 5 are presented in a table:

N	set of nodes
A	set of arcs
S	set of Scenarios
$\{1, \ldots, H\}$	set of linearization intervals
i, j	indices over N
s	index over S
h	index over $\{1, \ldots, H\}$
H	number of linearization intervals
l	number of wind mills to be placed
b_i	node balance at node i
t_i	correction parameter for node i
M_i	random variable representing *potential wind power* at node i
m_i	any realization of M_i
m_i^s	realization of M_i in scenario s
p^s	probability of scenario s
c_{ij}	cost parameter of arc ij, representing its reactance
r_{ij}	resistance of arc ij
$C_{ij,h}$	cost parameter of parallel arc ij, h in linearized graph
$R_{ij,h}$	resistance parameter of parallel arc ij, h in linearized graph
$U_{ij,h}$	capacity of parallel arc ij, h in linearized graph
π_i^s	node potential of node i in scenario s
x_i	number of wind mills placed at node i
y_{ij}	power flow over arc ij
y_{ij}^s	power flow over arc ij in scenario s
$y_{ij,h}^s$	power flow over arc ij, h in linearized graph, in scenario s
B	node-arc incidence matrix for the original graph
B^*	node-arc incidence matrix for the linearized graph
$\tilde{b}_i^s$	node balance at node i in scenario s, after wind correction

Table 1: List of symbols used in chapters 4 and 5

In addition, by placing the parameters m_i^s, the first stage decision variables x_i, or the second stage decision variables y_{ij}^s in a vector, the notations $\vec{m^s}$, $\vec{x}$ and $\vec{y^s}$ become available. Similarly $\vec{M}$ represents the random vector of potential wind speeds.

5 Reformulation into deterministic optimization problem

In chapter 4 a complete mathematical formulation of the research question was given. In this chapter mathematical programming techniques are applied in order to reformulate the problem into a standard solvable minimization problem. In section 5.1 the stochastic program is rewritten into a large deterministic one by introducing scenario parameters. These parameters will be obtained by simulation in chapter 6. In section 5.2 a linearization of the network flow that arises in the second stage is described. In section 5.3 it is shown how duality theory of linear network flows may be used to combine the two stages into one large optimization problem. Section 5.4 deals with the computation of the solution.

5.1 Stochastic programming

Models like the one above are called two stage stochastic optimization problems: in the first stage the decisions $\vec{x}$ have to be made while the future behaviour of M is still uncertain; then a realization of M is observed; then in the second stage the laws of power electronics determine the decision variables $\vec{y}$, which results in a value of the objective function, that is, the transportation losses.

The first difficulty in carrying out the minimization above, has to do with the form of the objective function. Deterministic optimization algorithms cannot trivially handle an objective function like

$$\mathbb{E}_M \left[\sum_{(ij)\in A} \left(y_{ij}(\vec{M}, \vec{x}) \right)^2 r_{ij} \right],$$

which contains an expectation. (The notation $y_{ij}(\vec{M}, \vec{x})$ makes the dependence clear of the y_{ij} on $\vec{M}$ and $\vec{x}$.) The random variable $\vec{M}$ has a continuous (multidimensional) distribution, and it is difficult to see how a (multidimensional) integral in the objective can be dealt with.

Every continuous distribution can be approximated by a discrete one, and moreover, one which assumes finitely many values. Therefore also our wind speed distribution may be assumed discrete. And when it is assumed that $\vec{M}$ takes only finitely many different values, the expectation can be computed as a finite sum. To this end, define a set S containing finitely many so-called *scenarios* $s \in S$. Next assume that $\vec{M}$ will take the value $\vec{m^s}$ with probability p^s, where $\sum_{s\in S} p^s = 1$. Then the expectation of a function $f(\vec{y})$ of $\vec{y}$ can be written $\sum_{s\in S} p^s \cdot f(\vec{y^s})$.[8]

In this way the stochastic program of the previous section can be written as an equivalent (large-scale) deterministic extension by summing over the scenario space:

$$\begin{aligned}
\text{minimize} \quad & \sum_{s\in S} p^s \cdot \left(\sum_{(ij)\in A} (y^s_{ij})^2 r_{ij} \right) & \\
\text{subject to} \quad & x_i \in \{0, 1, \dots\} & i \in N \\
& \sum_{i\in N} x_i = l, & \\
\text{and where} \quad & & \\
\vec{y^s} = \text{argmin} \quad & \sum_{(ij)\in A} c_{ij} (y^s_{ij})^2 & \\
\text{subject to} \quad & y^s_{ij} \geq 0 & (ij) \in A, s \in S \\
& \left(B\vec{y^s} \right)_i = b_i + m^s_i x_i - (\vec{m^s} \cdot \vec{x}) t_i & i \in N, s \in S.
\end{aligned}$$

[8]The scenario set S is a kind of finite discretization of the sample space Ω of M: in fact, the expectation in the objective function could have been written $\int_{\omega\in\Omega_M} \cdot \left(\sum_{(ij)\in A} (y^\omega_{ij})^2 r_{ij} \right) d\omega$, whereas the approximation by introducing the scenario set S looks like $\sum_{s\in S} p^s \cdot \left(\sum_{(ij)\in A} (y^s_{ij})^2 r_{ij} \right)$.

5.2 Linearization

Still this formulation does not fit a standard mathematical programming model. The first stage decisions $\vec{x}$ aim to minimize the expectation of some function of the second stage decisions $\vec{y^s}$, whereas these themselves must be chosen in such a way that they minimize some other function. At first sight it is not clear how to formulate the entire problem as a single optimization model. The constraint that the $\vec{y^s}$ are the arguments which solve another minimization problem, is quite an unusual one. But by a few steps this constraint may be written in a standard form which can be handled by optimization software. The first of these steps is the linearization of the objective functions.

Take the second stage minimization problem

$$\begin{aligned} \vec{y^s} = \text{argmin} \quad & \sum_{(ij)\in A} c_{ij}(y^s_{ij})^2 & \\ \text{subject to} \quad & y^s_{ij} \geq 0 & (ij) \in A, s \in S \\ & \left(B\vec{y^s}\right)_i = b_i + m^s_i x_i - (\vec{m^s} \cdot \vec{x})t_i & i \in N, s \in S. \end{aligned}$$

As already observed, these second stage decisions solve $\#S$ independent but very similar convex network flows. (The only difference between scenarios is a smaller or larger change in supply of the wind mills with their corresponding corrections.) The similarity between the scenarios may be used to obtain the solution of all $\#S-1$ network flows efficiently by starting out from the solution to the first one. This practical remark will be returned to in chapters 7 and 8. Of convex network flows it is known that they can be approximated by linear network flows to any desired degree of accuracy by adding multiple arcs between the connected nodes, and defining appropriate arc cost and capacity parameters (as, for example, [Ahuja et al. 1993] do). This can be understood as follows.

Suppose some arc $(ij) \in A$ has capacity u_{ij} and cost function $c_{ij}(y_{ij})^2$, and suppose the cost function is approximated by a piecewise linear function consisting of H line segments as in figure 7:

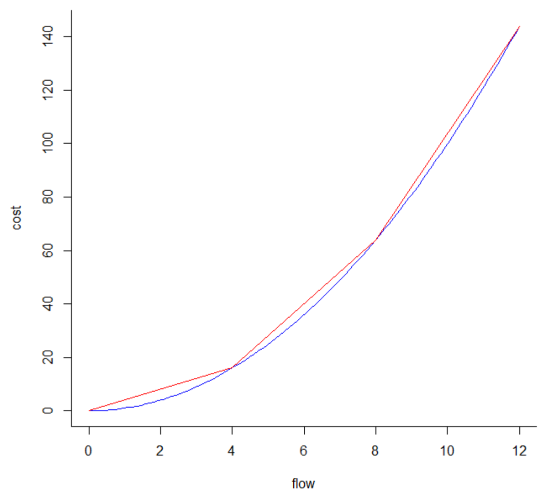


Figure 7: Linearization of the cost function

Now this arc (ij) and the piecewise linear cost function on it behave in the same way as H parallel arcs (ij,h), $h = 1,\ldots,H$ with capacity $U_{ij,h} = \frac{u_{ij}}{H}$ and linear cost functions $C_{ij,h}$ that have a slope equal to the slope of the h'th line segment in the picture. This follows because a flow from i to j that should minimize the cost, will allways first fill the cheapest arcs (ij,h), that is, from $h = 1$ upwards to $h = H$. The old cost of the flow over arc (ij) now equals the sum of the new flows over the arcs (ij, h).

By applying this transformation the number of nodes remains the same, and node balance constraints change only in the sense that there is one more arc parameter to sum over. The objective function becomes $\vec{y^s} = \text{argmin} \sum_{(ij \in A)} \sum_{h \in \{1,\dots,H\}} C_{ij,h} y_{ij,h}$.

The value of the first stage objective function $(y_{ij}^s)^2 r_{ij}$ can easily be computed in terms of the new flow variables $y_{ij,h}$ and in a similar way newly defined $R_{ij,h}$. The total linearized problem then looks like:

$$
\begin{aligned}
\text{minimize} \quad & \sum_{s \in S} p^s \cdot \left(\sum_{(ij) \in A} \sum_{h \in \{1,\dots,H\}} R_{ij,h} y_{ij,h}^s \right) & \\
\text{subject to} \quad & x_i \in \{0, 1, \dots\} & i \in N \\
& \sum_{i \in N} x_i = l, & \\
\text{and where} & & \\
\vec{y^s} = \text{argmin} \quad & \sum_{(ij \in A)} \sum_{h \in \{1,\dots,H\}} C_{ij,h} y_{ij,h}^s & \\
\text{subject to} \quad & 0 \le y_{ij,h}^s \le U_{ij,h} & (ij, h) \in A^*, s \in S \\
& \sum_{j,h:(ij,h) \in A^*} y_{ij,h}^s - \sum_{j,h:(ji,h) \in A^*} y_{ij,h}^s = b_i + m_i^s x_i - (\vec{m^s} \cdot \vec{x}) t_i & i \in N, s \in S,
\end{aligned}
$$

where A^* denotes the new (strongly enlarged) set of arcs with linear cost, indexed by (ij, h), where $(ij) \in A, h \in \{1, \dots, H\}$. Again the last equation represents the node balance constraints. By the introduction of the parallel arcs, each node now has more arcs incident to it than before, but the structure of the graph remains the same. If B^* denotes the node-arc incidence matrix for the new set of arcs, the flow balance constraints could have been written $\left(B^* \vec{y^s}\right)_i = b_i + m_i^s x_i - (\vec{m^s} \cdot \vec{x}) t_i$ for $i \in N, s \in S$.

5.3 Reformulating the second stage constraint

The second and final step of rewriting the problem is performed in this section. When this has been done the problem of minimizing the transportation losses will be formulated in the form of a standard mathematical program.

The point of linearizing the network flow is that for linear network flows a well-known duality framework exists, in which *strong duality* holds: solving the minimum cost flow under flow non-negativity and balance constraints is equivalent to (and yields the same optimal objective value as) solving another, related, linear program, called its *dual*. [Bertsimas and Tsitsiklis 1997] give a very nice introduction into general duality in chapter 4, and apply it to network flows in chapter 7. Alternatively [Ahuja et al. 1993] chapter 9 derive the dual network problem from first principles.

The general results may be summarized by the following. Let $z(y^*)$ denote the cost value of a minimum cost flow with flow variables $\vec{y}$ and balance vector $\vec{\tilde{b}}$. Define $\pi : N \to \mathbb{R}$, called the *node potential* which associates with every node i a real number π_i. Then maximizing $w(\pi) = \sum_{i \in N} \tilde{b}_i \pi_i - \sum_{(ij) \in A} \max(0, \pi_i - \pi_j - c_{ij}) u_{ij}$ yields an optimal node potential π^*, and $w(\pi^*) = z(y^*)$.Therefore, when faced with a linear network flow problem, it can be checked if it has been solved to optimality by a simple condition: does a node potential π exist for which $w(\pi)$ equals the current flow cost?

This theory can be applied in the following way. In the two stage model above, one of the constraints is that the $\vec{y^s}$ have to be chosen such that they minimize some flow cost objective $\sum_{ij,h \in A^*} C_{ij,h} y_{ij,h}^s$ called the *primal value*. For the optimal $\vec{y^s}$, there exist node potentials $\vec{\pi^s}$

for wich the *dual value* $w(\pi^s)$ equals the primal. Hence in stead of the 'argmin'-constraint, it is equivalent to add free node potential variables for every scenario, and add the constraint that for every scenario primal value should equal dual value.

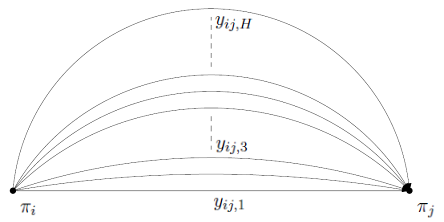


Figure 8: Node potentials and flow variables for the parallel arcs

Let $\tilde{b_i^s} = b_i + m_i^s x_i - (\vec{m^s} \cdot \vec{x})t_i$ be the balance at node i, corrected for scenario s and the placement decisions $\vec{x}$. Then the condition that $\vec{y^s} = \text{argmin} \sum_{ij,h \in A^*} C_{ij,h} y_{ij,h}$, where $\vec{y^s}$ is a feasible flow, is equivalent to the condition that $\sum_{i \in N} \tilde{b}_i^s \pi_i^s - \sum_{(ij,h) \in A^*} \max(0, \pi_i^s - \pi_j^s - C_{ij,h}) U_{ij,h} = \sum_{ij,h \in A^*} C_{ij,h} y_{ij,h}$, where $\vec{y^s}$ is a feasible flow and $\vec{\pi^s}$ a node potential vector.

After these new variables are introduced and the constraint changed as described above, the model can finally be formulated as a standard optimization problem:

$$
\begin{array}{lll}
\text{minimize} & \displaystyle\sum_{s \in S} p^s \cdot \left(\sum_{ij,h \in A^*} R_{ij,h} y_{ij,h}^s \right) & \\
\text{subject to} & x_i \in \{0, 1, \dots\} & i \in N \\
& \sum_i x_i = l & \\
& 0 \leq y_{ij,h}^s \leq U_{ij,h} & (ij,h) \in A^*, s \in S \\
& \pi_i^s \in \mathbb{R} & i \in N, s \in S \\
& \left(B^* \vec{y^s}\right)_i = b_i + m_i^s x_i - (\vec{m^s} \cdot \vec{x}) t_i & i \in N, s \in S \\
& \displaystyle\sum_{i \in N} \tilde{b}_i^s \pi_i^s - \sum_{(ij,h) \in A^*} \max(0, \pi_i^s - \pi_j^s - C_{ij,h}) U_{ij,h} & \\
& \displaystyle\qquad\qquad = \sum_{ij,h \in A^*} C_{ij,h} y_{ij,h}^s & s \in S,
\end{array}
$$

where $\tilde{b}_i^s$ denotes the corrected node balances, and B^* the node-arc incidence matrix adjusted for the parallel arcs.

5.4 Solution strategies

With respect to the solution of the mathematical program formulated in the last section, there are a few things that should be kept in mind. By the introduction of the new indices s for stochastic programming and h for linearization of the convex cost arcs, and the introduction of the new variables π^s, the whole program quickly becomes giant. For example, in the next chapter, 729 scenarios are defined, that is, $\#S = 729$, and if a transmission line of heat capacity 1000MVA is to be modeled with precision $1MW$, then $H = 1000$. The result is, that the number of flow variables is multiplied by almost a million, and that in addition $\#S \cdot n$ new (node potential) variables enter the problem.

But there is one more (somewhat hidden) drawback of the method described in this chapter.

The constraint that the dual must equal the primal value of the second stage was written

$$\sum_{i\in N}\tilde{b}_i^s\pi_i^s - \sum_{(ij,h)\in A^*}\max(0,\pi_i^s-\pi_j^s-C_{ij,h})U_{ij,h} = \sum_{ij,h\in A^*}C_{ij,h}y_{ij,h}^s, s\in S;$$

but the balance $\tilde{b}_i^s$ used in this equation crucially depends on the first stage decisions $\vec{x}$. It was defined $\tilde{b}_i^s = b_i + m_i^s x_i - (\vec{m^s}\cdot\vec{x})t_i$, and therefore the product $\tilde{b}_i^s\pi_i^s$ contains a term $x_i\pi_i^s$; which makes the model nonlinear. Naive rigorous solution when the model is defined on any realistic scale (which may include hundreds of nodes and arcs, as well as hundreds of scenarios and parallel arcs) will probably be infeasible, as shown in section 7.5. Specialized algorithms may be designed by exploiting the primal-dual relationship between variables $y_{ij,h}^s$ and π_i^s, but are beyond the scope of this thesis.

It seems impossible to rewrite the current model into a form which maintains the linearity of all constraints. But if it could be done, the problem would be a large scale linear program. Standard stochastic programming methods exist to decompose such programs into several smaller ones in order to reduce solving time. The most common one is Benders' algorithm, but when there is special structure (such as network structure) in the program, more specialized methods have been proposed. [Nielsen and Zenios 1990], for example, describe the Row Action algorithm and related methods. These methods work by replicating first stage decision variables x_i, $i\in N$ into $\#S$ copies $\vec{x}^s$, $s\in S$, and adding the constraint that all $\vec{x}^s$ should be equal. By relaxing this last constraint, one obtains $\#S$ independent subproblems.

Therefore in this thesis two solution strategies are presented: In chapter 7 steps are described needed for the rigorous solution, and the problem is solved for $\#S$ and H small enough. In chapter 8 some heuristics are developed in order to obtain suboptimal solutions for larger $\#S$ and H. In chapter 9 the results of a heuristic method are used. Our ability to solve small scale problems rigorously could help to establish a quality estimate of the heuristic method. This is one of the suggestions for further research mentioned in chapter 11.

6 Wind power simulation

In this chapter a simulation of wind speeds in the Netherlands is performed, based on empirical data and on wind speed models from literature. The goal is to compute reasonable values for the model parameters p^s and m^s that were used in the optimization model in chapter 5. Recall that the interpretation of p^s is the probability of a scenario $s \in S$, whereas m^s is a vector containing for scenario s the joint wind power output if on every node there would be exactly one wind turbine.

The Dutch meteorological intstitute (`www.knmi.nl`) publishes wind speed measurements of over 50 measurement stations. Figure 9 shows how they are distributed over the country.

Figure 9: Wind stations in the Netherlands

For this thesis data were taken from the hourly wind speed measurements of all these stations during the years 2001 until 2010. The first aim is to reduce the dimension of the data, i.e., come up with some partition of the Netherlands into a few different wind zones, within which we will assume equal wind speeds. For reasons suggested in section 9.2, 6 zones seems appropriate. In this way one realisation of all wind speeds reduces from a stochastic vector of dimension 53 to dimension 6. This reduction maintains the most important national differences, while it neglects unimportant small local differences.

6.1 Data input and inspection

The statistical software tool R [The R project 2012] was used to perform the computations. This program is quite efficient in handling large datasets (for example, the hourly data for only these ten years contains over 80,000 measurements), and contains a lot of standard statistical procedures. The data themselves are available as text files on the site `http://www.knmi.nl/klimatologie/onderzoeksgegevens/potentiele_wind/`. They can be downloaded and read into R as follows:

```
> setwd("C:/Users/leenmants/winddata")
> a=read.csv("potwind_210_2001", skip=21)
> row.names(a)=paste(a$DATE,a$TIME, sep="")
> a$DATE=a$TIME=a$DD=a$QDD=a$QUP=NULL
> names(a)[1]="210"
> station=c(225,229,235,240,242,248,249,251,252,253,254,258,260,
265,267,269,270,271,273,275,277,278,279,280,283,285,286,290,308,
310,311,312,313,315,316,319,320,321,323,324,330,331,340,343,344,
348,350,356,370,375,377,380,391)
> station=as.character(station)
> for(k in station){
>   b=read.csv(paste("potwind", k, "2001", sep="_"), skip=21)
>  row.names(b)=paste(b$DATE,b$TIME, sep="")
>  a[rownames(b),k]=b$UP
> }
> rm(b, station)
```

In this way `a` has become a *data frame* containing 53 (column) vectors of `length(a$"225")` rows with row names indicating date and time of the measurement. Note how the command line `a[rownames(b),k]=b$UP` makes sure that all measurements corresponding to the same time are put in the same row. This is evidently important, as can be seen by typing `sum(is.na(a))` which returns `239616`, that is, many data points are missing.

The first thing to do is scanning the correlation structure. This can be done by typing `c=cor(a, use="pairwise")`", which returns the correlation matrix. `sum(is.na(c))`, `min(c, na.rm=TRUE)` and `mean(c, na.rm=T)` reveal some properties of the correlation matrix, which can be inspected by `fix(c)`.

6.2 Zone partition

Dividing the whole area into different wind zones now comes down to determining groups of variables which have high inner correlation, and low correlation between the groups. The multivariate statistical techniques to do this, are called *clustering* techniques. Hierarchical clustering starts out with as many clusters as there are variables, and then subsequently merges clusters which lie 'nearest' to each other (have highest correlation, or lowest distance between them). Exact description can be found in [Johnson et al. 2007]. The standard procedure to perform this analysis in R is:

```
> c[is.na(c)]=.5
> dd=as.dist(1-c/2)
> dd=round(1000*dd)
> plot(hclust(dd), xlab="Wind Stations", main="Cluster Dendro
gram for Wind Zones"),
```

yielding the following plot.

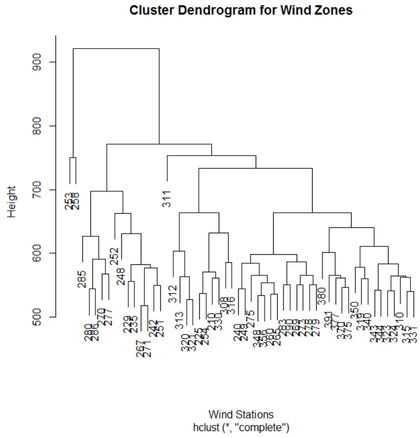


Figure 10: Cluster Dendrogram

When choosing a height at the y-axis, and drawing a horizontal line at that height, every intersection with the dendrogram corresponds to a cluster (which contains all stations that are the descendants of that intersection point). By visual inspection, an appropriate height can be chosen to make the cut.

To this end, several things were noticed: The stations 253 and 258 on the left of the dendrogram lie too far away to be on the map, and therefore should be left out. Station 252 is the K13 measurement station far on the north sea, which was left out as well because building wind turbines so far from the coast would require total different investments from anywhere on shore. Furthermore station 313 (Hoofdplaat) was ignored, because the smallness of its area would not justify treating this single station as a separate zone.

In this way the following clustering and corresponding wind zones of the Netherlands were obtained:

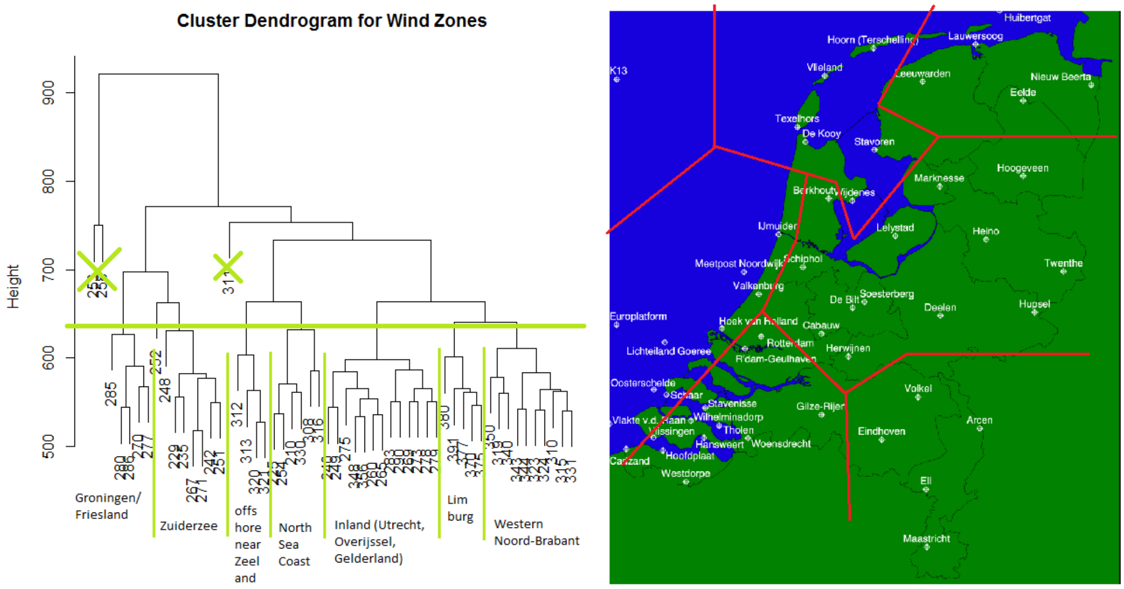


Figure 11: Clustering and Zone Partition

6.3 Marginal distributions

As to the marginal distribution of the wind speed at a particular site. In literature, the Weibull or even Rayleigh (which is a special case of the Weibull) distribution is advocated. See, for example, [Feijoo et al. 1999], or [Carpinelli et al. 2001]. At first glance, both the histogram and the empirical distribution function fit reasonably well with the Rayleigh distribution. The scale parameter is derived from the sample mean of the measurements. By typing

```
> m=mean(a$"210", na.rm=TRUE)
> wa=2
> wb=m/(1+1/wa)
> par(mfcol=c(1,2))
> hist(a$"210", main="Histogram for Valkenburg", xlab="wind
speeds", mfcol=c(1,1))
> q=seq(0,220, length.out=length(a$"210"))
> lines(q,length(a$"210")*10*dweibull(q,wa,wb), col='red')
> plot(ecdf(a$"210"),xlim=range(a$"210"), main="ECDF for Valk
enburg", xlab="Wind speed")
> lines(q,pweibull(q,wa,wb), col='red')
```

the graphs below can be observed:

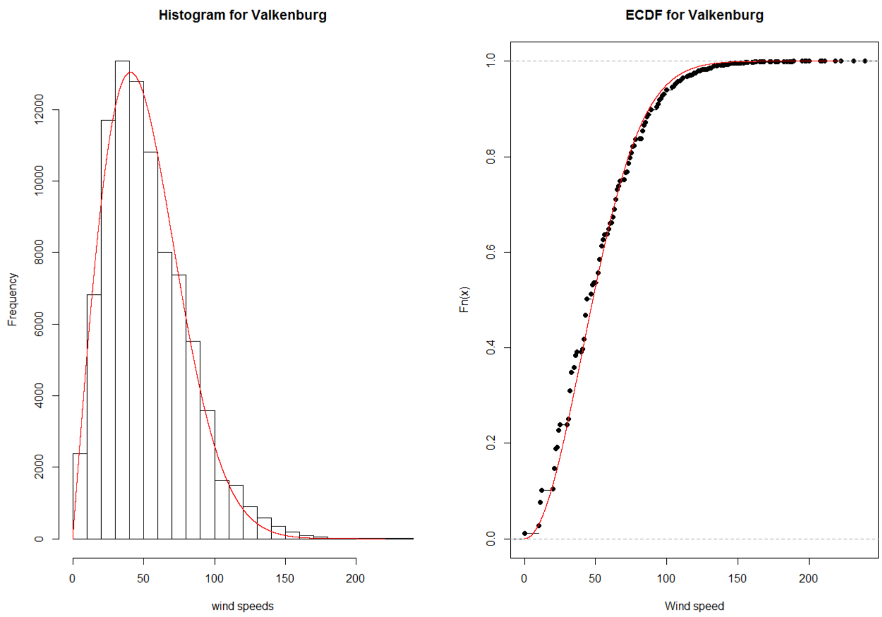


Figure 12: Histogram and ECDF for the marginal

At first sight there seems to be a good match, but this method of visual inspection of the histogram and empirical distribution function is inadequate for judging distribution fits. A little more revealing is looking at the sample qqplots below. From these it becomes clear that although the Rayleigh fits reasonably well, a yet more ideal fit is obtained if we adjust the shape parameter of the Weibull from 2 to for example 1.8. The way to get these fits is

```
> p=seq(from=0, to=1,length.out=length(a$"210"))
> qqplot(a$"210",qweibull(p,wa,wb))
> abline(0,1, col="red")
> qqplot(a$"210",qweibull(p,1.8,m/gamma(1+1/1.8)))
> abline(0,1, col="red").
```

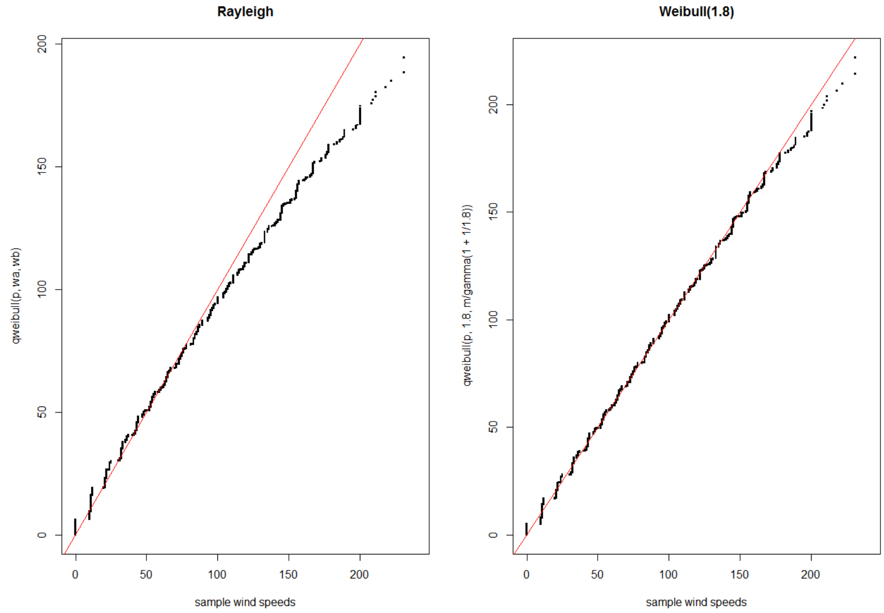


Figure 13: qqplots for Rayleigh (left) and Weibull (right)

The formal way to test hypotheses about the distribution of a Random Variable is the Kolmogorov Smirnov test. The test statistic D of this test is the integral of the absolute difference between the empirical distribution function and the theoretical distribution function. (The test is called *one-sided* if the ecdf is compared to a theoretical distribution, and *two-sided* if two ecdfs are compared.) Details about the derivation of the distribution of the test statistic and computation of the p-value can be found in [Kvam et al. 2007], pp $81-89$.

Again the procedure in R is very simple.

```
> ks.test(a$"210","pweibull",2,m/gamma(1.5))
```

returns a test statistic $D = 0.0693$, and a p-value $< 2.2e^{-16}$, which seems a bit disappointing. But the reason for this rude rejection of the hypothesis, is the enormous amount of data. By the large size of the sample, the variance of the test statistic becomes so small, that every deviation from the hypothesized distribution is detected. And because empirical data never follow perfectly a theoretical distribution, this rejection was to be expected.

By typing the command `ks.test(a$"210"[seq(1,80000, by=1000)], "pweibull",2,m/gamma(1.5))`, which performs the test on an excerpt of 80 approximately independent time instances, a p-value of 0.2319 was obtained. In other words, 80 observations are not enough to refute the assumption of Rayleigh distributed steady state wind speeds. (Needless to say, by the time-correlation of wind speeds clearly `ks.test(a$"210"[1:80],"pweibull",2, m/gamma(1.5))` results in a p-value of $< 2.2e^{-16}$ again.) Similar results hold for all measurement stations. Conclusion is that the assumption of Rayleigh distributed marginals is sufficiently justified.

6.4 Joint distribution and simulation

But what about the joint distribution of the wind speeds in the several regions? [Feijoo et al. 1999] suggest a multivariate distribution which is easy to simulate, and has marginals with approximate Rayleigh properties. They cite [Lee 1979] and [Jensen 1970] as mathematically more precise, but of a (for this purpose) unjustified complexity.

Their simulation method is based on a few observations. Given an n-dimensional random vector z with mean vector μ_z and covariance matrix Σ_z, a lower triangular matrix L and some vector

μ_y; then the transformed vector $y = Lz + \mu_y$ has expectation $E[y] = L\mu_z + \mu_y$ and covariance structure $\Sigma_y = L\Sigma_z L^T$. Moreover, when z is multivariate normally distributed, y is as well (and its distribution is completely determined by mean and covariance matrix).

As applied to simulation, when z is taken to be an independent vector of standardized normals (so $\mu_z = 0$ and Σ_z is the identity matrix), any desired multivariate normally distributed y can be obtained by decomposing $\Sigma_y = LL^T$, and putting $y = Lz + \mu_y$. This decomposition of Σ_y is always possible because any covariance matrix is positive definite, and can be done for example by Cholesky's decomposition.

All this is well known from theory. The idea now is to perform this same transformation on independent standardized Rayleigh random variables. [Feijoo et al. 1999] go on to show that the resulting random vector not only has the desired covariance structure, but also respects the Rayleigh properties of the marginals, and most features of the joint distribution.

This procedure was performed in R by extracting one representing station for every wind zone from the original data, and computing their marginal Rayleigh parameters:

```
> aa=a[,c("330","267","260", "370","340","286")]
> sum(is.na(aa))   #no missing data
> exp=mean(aa),
> mvb=exp/gamma(1+.5)
```

then computing the sample covariance and its decomposition

```
> bb=var(aa)
> L=t(chol(bb))
> sum(abs(L%*%t(L)-bb)>1e-10) #check: correct decomposition.
```

In order to simulate the Rayleigh distributed variables, one should simulate uniformly distributed variables, and transform them by the inverse of the Rayleigh distribution function. The proof that this yields a correct simulation can be found for example in [Madras 2002]. Now the Weibull distribution function looks like

$$F(x) = 1 - e^{-(\frac{x}{b})^a} \text{ on domain } x \in [0, \rightarrow),$$

with a the shape parameter and b the scale parameter. Furthermore it holds for the expectation and variance

$$\begin{aligned} E[X] &= b \cdot \Gamma(1 + \frac{1}{a}) \\ Var(X) &= b^2 \left[\Gamma(1 + \frac{2}{a}) - \Gamma(1 + \frac{1}{a})^2 \right]. \end{aligned}$$

So the command lines `u=runif(10000)` and `x=30*sqrt(-log(1-u))` would yield 10000 independent Weibull(2,30) variables. And combining this with the above described procedure to transform to correlated variables, the command lines

```
> sd=sqrt(mvb^2*(gamma(2)-gamma(1.5)^2))
> uu=matrix(runif(6*10000,0,1),ncol=6)
> transformtoWP=function(v, mvb, exp, sd,L){
>   L%*%matrix((mvb*sqrt(-log(1-v))-exp)/sd, ncol=1)+exp
> }
> ww=apply(uu,1,transformtoWP,mvb,exp,sd,L)
> ww=t(ww)
> simbb=var(ww)
```

show very similar sample covariances from the data and from the simulation. (Of course, this is not surprising because the transformation was designed to have the simulated sample covariances converge to the measured ones.) Also, the marginals endured a quick Kolmogorov-Smirnov test by

```
> ks.test(ww[1:80,1],"pweibull",2,exp[1]/gamma(1.5)).
```

Though it is obvious that this transformation can in no way be exact: for instance, `sum(ww<0)` reveals negative simulation outcomes which have no physical meaning. These negative values are simply truncated to zero by the transformation in the next chapter, so they pose no problem in our context.

6.5 Transformation to wind power and computing the scenario parameters

Having simulated 10000 independent instances of joint wind speeds, these should now be transformed into the powers that are produced by a wind turbine within these zones. Four main characteristics are mentioned in literature: a wind speed value v_{ci} (cut-in velocity) below which the turbine has no output, a rated velocity v_r above which the turbine has its maximum output, and a cut-out velocity v_{co} above which the turbine is deactivated in order to prevent damage. The fourth caracteristic is the rated or maximum power P_r itself.

[Li Yang Chen 2010] takes as power output curve the left graph in the figure below, which is doubtless the more accurate model. But we think that the linear approximation used by [Carpinelli et al. 2001] and [Papaefthymiou and Klöckl 2008], and shown on the right in the figure 14, suffices for our purposes.

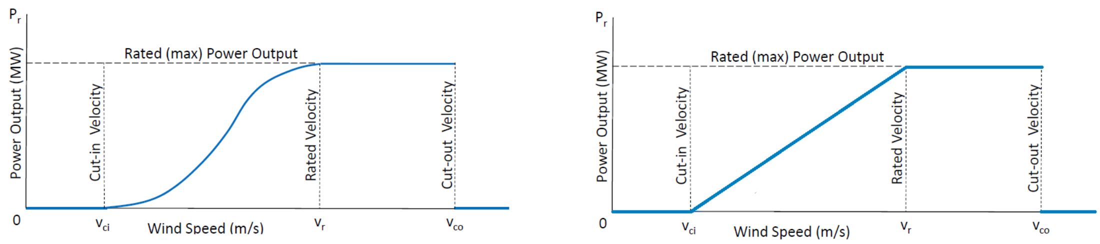


Figure 14: Non-linear and linear power curves

There are many different sizes and types of wind turbines available in industry. For this thesis typical characteristics for a modern large $2500kW$ turbine were taken. They are specified at the webpage of the [Wind Power Program] as follows: $v_{ci} = 3.5m/s$, $v_r = 14m/s$, $v_{co} = 25m/s$ and $P_r = 2500kW$. Note that although the marginal wind *speeds* are Rayleigh distributed, the marginal wind *powers*, are quite differently distributed. This is illustrated in figure 15. The form of the output power distribution is in agreement with the result of [Papaefthymiou and Klöckl 2008]. Note that the height of the bar on the left increases when v_{ci} is increased or v_{co} is decreased; and the bar at the right grows when v_r is decreased or v_{co} is increased.

These graphs were obtained by transforming the wind speed simulation data by our specified power output curve thus:

```
> maxwind=250
> minwind=30
> refwind=140
> maxpow=2500 #in kW
> zz=matrix(0,ncol=dim(ww)[2], nrow=dim(ww)[1])
> zz[ww<minwind | ww>maxwind]=0
> zz[refwind<ww & ww<=maxwind]=maxpow
> zz[minwind<=ww & ww<=refwind]=(ww[minwind<=ww & ww<=refwind]
-minwind)*maxpow/(refwind-minwind)
> par(mfcol=c(1,2))
> hist(aa[,1]/10, main="Histogram of wind speeds", xlab="speed
```

```
(m/s)")
> hist(zz[,1], main="Histogram of power outputs", xlab="power
(kW)")
```

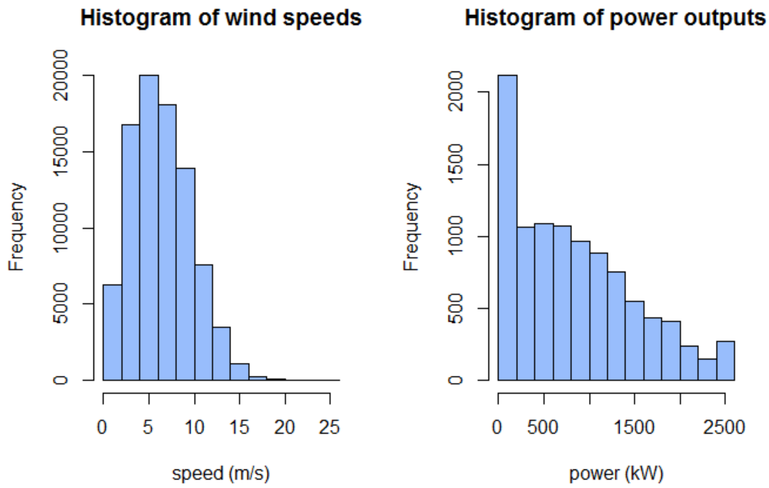


Figure 15: Histograms of wind speed and wind power

By the procedure described until here, it is possible to simulate correctly correlated wind power outputs for single wind plants over the different regions. Total wind power output in a region can then be obtained by multiplying that value by the number of turbines in the region, as described in the previous chapter. This simulation procedure can be used to compute values for the scenario parameters p^s and m^s, defined for the stochastic optimization program.

p^s, the probability of a scenario $s \in S$, could be computed by partitioning the 6-dimensional range of power outputs, and counting which parts of the simulated values would lie in the particular areas. m^s, the value of the scenario, is then obtained by averaging over these values only. But the fact that we started the simulation from independent uniforms, provides us with an opportunity for simplification: by simply partitioning the range of the multivariate uniforms (that is, $[0,1]^6$), then mapping them to output powers and compute the m^s, it becomes easy to define scenarios which have all equal probability p^s.

This procedure was performed by distinguishing 3 areas in every dimension: low, middle or high value; and then doing the computations for every scenario independently. In this way $3^6 = 729$ different scenarios were created. Because the whole mapping from the uniform random variable to the wind power output is monotone, these distinctions between high and low values are preserved. Obviously $p^s = \texttt{ps}$ is a vector of length 729, containing $\frac{1}{729}$ at every entry; and the corresponding $m^s = \texttt{ms}$ were computed as follows:

```
> ps=rep(1/3^6,3^6)
> ms=matrix(rep(0,6*3^6), ncol=6)
> for(i1 in 0:2){
>   for(i2 in 0:2){
>     for(i3 in 0:2){
>       for(i4 in 0:2){
>         for(i5 in 0:2){
>           for(i6 in 0:2){
>             uu=cbind(runif(1e4,i1/3,(i1+1)/3),runif(1e4,i2/3
,(i2+1)/3),runif(1e4,i3/3,(i3+1)/3),runif(1e4,i4/3,(i4+1)/3),r
unif(1e4,i5/3,(i5+1)/3),runif(1e4,i6/3,(i6+1)/3))
```

```
>               ww=apply(uu,1,transformtoWP,mvb,exp,sd,L)
>               ww=t(ww)
>               zz=matrix(0,ncol=dim(ww)[2], nrow=dim(ww)[1])
>               zz[ww<minwind | ww>maxwind]=0
>               zz[refwind<ww & ww<=maxwind]=maxpow
>               zz[minwind<=ww & ww<=refwind]=(ww[minwind<=ww &
ww<=refwind]-minwind)*maxpow/(refwind-minwind)
>               ms[1+i1+3*i2+9*i3+27*i4+81*i5+243*i6,]=as.vector
(colMeans(zz))
>               rm(uu,ww,zz)
>             }
>           }
>         }
>       }
>     }
> }
```

Now that these parameters p^s and m^s, $s \in S$ have been computed, the deterministic optimization problem that was drawn up at the end of the last chapter, can be solved.

7 Solving the mathematical program

In the chapters 4 and 5 the model was formally derived by which the research question is to be answered. In chapter 6 the model parameters p^s and m^s, associated with the stochastic wind energy were computed by simulation. In the current chapter it is shown how this model can be implemented step by step in the optimization software program Aimms. There are two reasons for describing the several intermediate steps. The first reason is that in this way it is easier to understand the implementation. The whole model looks rather complicated, but it is a construction of a few simple consecutive ideas. The practical advantage of implementing these steps successively is that it simplifies error detection for the programmer: when the whole model is implemented at once, and then a runtime error or an unexpected outcome occurs, it can be very difficult to trace back the cause of the error. The second reason for describing the intermediate steps is that the precise model described here will be too complex to solve. However, the heuristics described in the next chapter will rely on the basic structure of partial models that are presented in the first few steps.

7.1 Network data preparation

Topological network data including peak load or standard load data for the European High Voltage grid are available as a text file on the website belonging to the article by [Bialek 2005]. It includes line capacity data, line impedances, bus load and generation data, including the bus voltages required to do an AC power flow computation. Also all 1254 buses are linked to the country they are in (because the original study had to do with the border exchanges of electrical power).

The bus data and the branch data were split; the branch data were filtered such that all branches having one of their endpoints within the Netherlands remained; and then the bus data were filtered by deleting all the buses that did not belong to the Dutch grid or are directly adjacent to it. Furthermore only the branch data columns were retained that correspond to: its bus endpoints (which were called `FromB` and `ToB`), resistances, reactances and capacities. For the bus data only bus numbers, loads and generation magnitudes need to be retained. But three more things were required to complete the bus data: first, total load did not equal total generation anymore. This was corrected by adding the difference to the generation values of the nodes outside the Netherlands which were retained; so the shortage of generation in the Netherlands is imported from Germany and Belgium. Second, for the model described in chapters 4 and 5 a vector t is needed which describes which nodes adjust their generation when a new wind mill is added to the grid. We made the choice to decrease the value in proportion to their current generation. This vector is defined in the column `t`. Lastly all nodes have to be linked to one of the wind zones defined in the previous chapter. Although the bus data lacked coordinates of the buses, the structure of the network combined with a few bus names and HV grid data from [TenneT 2011] made it possible to retrieve the location of all buses, and link them to their appropriate wind zone. The result is shown in figure 16.

The borders of the wind zones have been copied from the picture in chapter 6. Because of the proximity of Ens (Node 614) to the IJsselmeer it was decided to link Ens to the zone `Zuiderzee`.

The bus data, prepared in the way described above, now look like:

```
COMPOSITE TABLE:
i   Loads     Generation  t       BusInWindZone
534 264.61323 0.00000     0.00000 WestNoordBrabant
547 0.00000   2777.18000  0.20840 Limburg
556 0.00000   239.12000   0.01795 Limburg
607 0.00000   710.41000   0.05331 GroningenFriesland
...
```

and the branch data:

```
COMPOSITE TABLE:
```

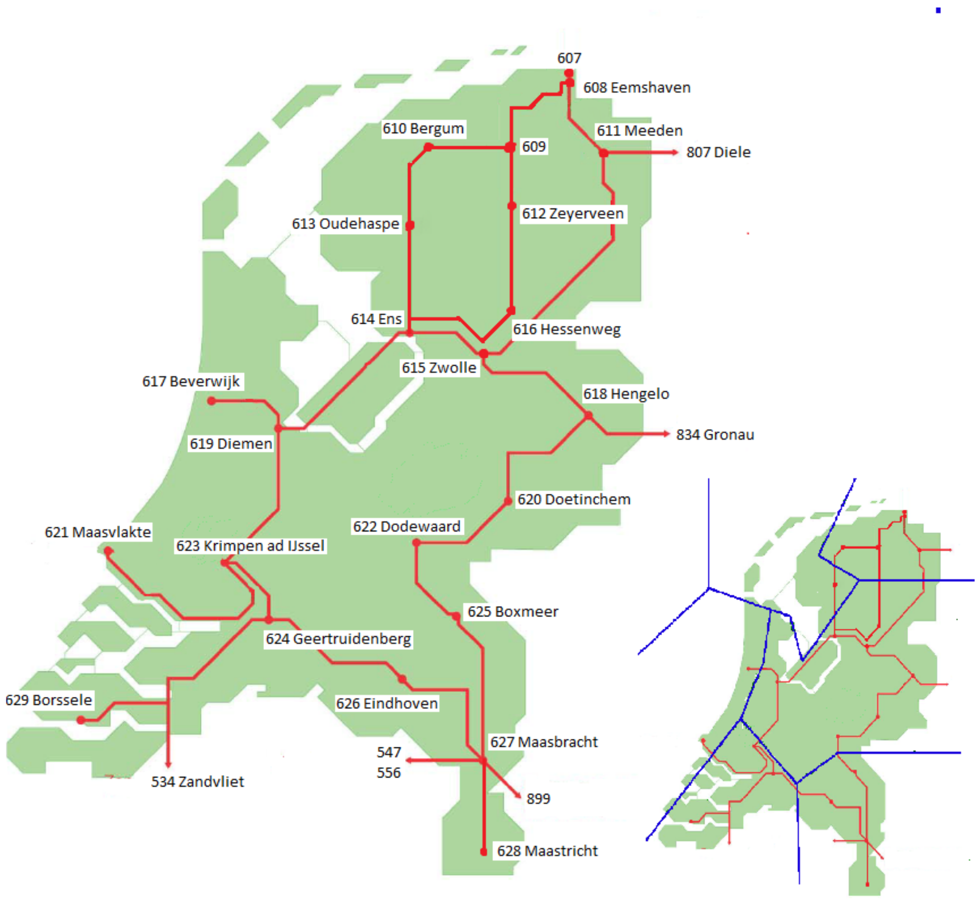


Figure 16: The network used for the model, with allocation to wind zones

ij	FromB	ToB	Resistances	Reactances	Capacities
1	534	624	0.001	0.01593	1650
2	534	629	0.001	0.00993	1476
3	627	547	0.001	0.03674	9900
4	556	627	0.001	0.0152	1320

...

In this format the data can easily be read into Aimms.

For ease of computation, all lines are assumed to be of equal voltage 380kV. Therefore the cost coefficients on the arcs can be taken to equal line reactances. No scaling is required. In section 3.7 it is explained how the cost coefficients should be scaled in case different voltages occur in the network.

7.2 Linearized network flow model

The first and simplest thing that can be done with the data, is to compute the power flow solution without the placement of any wind mills. For this case all wind and stochasticity can be left out. All that has to be done is to define a set `Buses` with index `i`, define the parameters `Loads` and `Generation` over index domain `i`, and do accordingly with the branches and their endpoints and electrical characteristics. The data can then be read by the command lines `read from file 'BusData.dat';` and `read from file 'BranchData.dat';` in a procedure.

```
SET:
   identifier :  Buses
   index      :  i

PARAMETER:
   identifier   :  Loads
   index domain :  (i)

PARAMETER:
   identifier   :  Generation
   index domain :  (i)

PARAMETER:
   identifier : Supplies
   index      : (i)
   definition : Generation(i)-Loads(i)

SET:
   identifier :  Branches
   index      :  ij

ELEMENT PARAMETER:
   identifier   :  FromB
   index domain :  (ij)
   range        :  Buses

ELEMENT PARAMETER:
   identifier   :  ToB
   index domain :  (ij)
   range        :  Buses

PARAMETER:
   identifier   :  Reactances
   index domain :  (ij)

PARAMETER:
   identifier   :  Capacities
   index domain :  (ij)
```

Note that the Supplies are defined as the difference between Loads and Generation. It remains to define the flow variables, the flow constraints, and the objective function. In Aimms this works as follows: any flow variable is called an Arc, and it requires as its parameters a Node as its tail and a Node as its head, and optionally a cost parameter. By defining its range nonnegative (and smaller than the capacity), flow feasibility at the arcs is ensured. And the flow constraints should be enforced by entering in the definition of the Nodes that are referred to in the Arc properties: netoutflow(i)=netinflow(i)+Supplies(i). Because Aimms regards Arcs as having a direction whereas electrical cables do not have a prescribed direction, it is necessary to add arcs in the opposite direction as well. Then an objective function variable can be defined as the sum over all Arcs squared times the Reactances, and Aimms will be able to solve the convex network flow.

But first the linearization from section 5.2 is described. This consists in adding extra parallel arcs of capacity H and a cost parameter which equals the slope of the line segments of figure 7. The following set and parameter declarations serve to compute and store all parallel and inverted arcs from the original data. Precision is the length of the intervals over which the linearization is defined; the set ArcPartition with its index h distinguishes the parallel arcs; and the set ArcDirection labels forward and backward arcs by the labels fw and bw:

```
PARAMETER:
   identifier :  Precision

PARAMETER:
   identifier :  LargestCapacity

SET:
   identifier :  ArcPartition
   subset of  :  Integers
   index      :  h
   definition :  {1 .. ceil(
                 LargestCapacity/Precision)}

SET:
   identifier :  ArcDirection
   index      :  v
   definition :  data {fw,bw}

ELEMENT PARAMETER:
   identifier   :  FromB2
   index domain :  (ij,h,v)
   range        :  Buses
   definition   :  if(v='fw') then
                   FromB(ij) else
                   ToB(ij) endif

ELEMENT PARAMETER:
   identifier   :  ToB2
   index domain :  (ij,h,v)
   range        :  Buses
   definition   :  if(v='fw') then
                   ToB(ij) else
                   FromB(ij) endif

PARAMETER:
   identifier   :  c_react
   index domain :  (ij,h)

PARAMETER:
   identifier   :  u
   index domain :  (ij,h)
```

Then all new arc parameters can be computed by the procedure below. It first sets H, then determines how many parallel arcs are needed, and then computes the parameters. Note that costs and capacities are the same for backward and forward arcs, so the algorithm needs to make no distinction for index v.

```
PROCEDURE
  identifier :  ComputeLinearization
  body       :
    Precision:=10;
    LargestCapacity:=0;
    for(ij in Branches) do
    if(Capacities(ij)> LargestCapacity) then
    LargestCapacity:=Capacities(ij) endif;
    endfor;

    for(ij in Branches, h in ArcPartition) do
    c_react(ij,h):=Reactances(ij)*Precision*(2*h-1);
    if(h*Precision<Capacities(ij)) then u(ij,h):= Precision
    else    if((h-1)*Precision<Capacities(ij))
            then u(ij,h):= Capacities(ij)-(h-1)*Precision;
            else u(ij,h):= 0;
            endif
    endif
    endfor
ENDPROCEDURE  ;
```

Finally the decision variables (that is the `Arcs`), the constraints (that is the `Nodes`), objective function (the variable `TotalFlowCost`) and the mathematical program can be entered:

```
ARC:
   identifier   :  Flow
   index domain :  (ij,h,v)
   range        :  [0, u(ij, h)]
   from         :  NetworkNode(
                   FromB2(ij,h,v))
   to           :  NetworkNode(
                   ToB2(ij,h,v))
   cost         :  c(ij,h)

NODE:
   identifier   :  NetworkNode
   index domain :  (i)
   definition   :  netoutflow=
                   netinflow+Supply(i)
```

```
VARIABLE:
   identifier :  TotalFlowCost
   range      :  free
   definition :  sum[(ij,h,v),
                 c(ij,h)
                 *Flow(ij,h,v)]

MATHEMATICAL PROGRAM:
   identifier   :  MinCostFlow
   objective    :  FlowCost
   direction    :  minimize
   constraints  :  AllConstraints
   variables    :  AllVariables
   type         :  Automatic
```

Now the command line `solve MinCostFlow` will solve the problem. Note that the variable `TotalFlowCost` is superfluous because Aimms recognizes the standard network flow objective function itself when 'FlowCost' in stead of 'TotalFlowCost' is entered as objective in the Mathematical Program.

7.3 Expected transportation losses for a given placement

In this section a given placement of wind turbines is assumed, and the goal is to extend the model of the former section in order to compute the expected line losses for this particular placement.

When wind turbines are introduced to the model, the wind zone partition of the Netherlands and wind stochasticity become relevant. For that, new sets `WindZones` and `Scenarios` have to be defined with their indices `k` and `s`. The simulated parameter m^s which links to every wind zone the potential wind power generated by one wind turbine, can be read into model parameter `PotentialWindPower(k,s)` by the command line `read from file 'WindData.dat'` when the data file looks like:

```
COMPOSITE TABLE:
k                  s    PotentialWindPower
NorthSeaCoast      1    137.9987923
```

```
ZuiderZee           1    2.480694625
Inlands             1    0
Limburg             1    0
WestNoordBrabant    1    0
GroningenFriesland  1    0
NorthSeaCoast       2    735.6025377
ZuiderZee           2    174.6701529
...
```

An element parameter `BusInWindZone(i)` links every node to a wind zone. Moreover the vector x is introduced as the model parameter `Placement(i)`, containing for every i the number of wind turbines placed at that node. The realized wind power at every node `ActualWindPower(i)`, the total wind power at all nodes $\vec{m^s} \cdot \vec{x}$, called `TotalWindPower`, the vector of adjustment fractions `t` and the new vector of corrected supplies may then be defined and computed as follows:

```
PARAMETER:
   identifier   :  t
   index domain :  (i)

ELEMENT PARAMETER:
   identifier   :  BusInWindZone
   index domain :  (i)
   range        :  WindZones

PARAMETER:
   identifier   :  Supply
   index domain :  (i,s)
   definition   :  Generation(i)+
                   ActualWindPowe
                   r(i,s)-Loads(i)
                   -t(i)*TotalWi
                   ndPower(s)

SET:
   identifier :  WindZones
   index      :  k

SET:
   identifier :  Scenarios
   index      :  s

PARAMETER:
   identifier   :  PotentialWindPower
   index domain :  (k,s)

PARAMETER:
   identifier   :  ActualWindPower
   index domain :  (i,s)
   definition   :  Placement(i)*Pot
                   entialWindPower(
                   BusInWindZone(i),s)

PARAMETER:
   identifier   :  TotalWindPower
   index domain :  s
   definition   :  sum[i,ActualWind
                   Power(i,s)]

PARAMETER:
   identifier   :  Placement
   index domain :  (i)
   definition   :  3.
```

Note that network characteristics and placement of the wind turbines are scenario independent, but all potential and actual wind power and therefore the supplies differ for all scenarios. Hence, also (in accordance with the model from chapter 5) the flow variables and constraints must run over the extra index `s`. Their definitions however remain the same.

The quadratic function which describes the line losses as a function of the flow over the line may be linearized by the same intervals as the function `Reactances(ij)`·`Flow(ij)`^{2}. Similarly, the parameters `c_resist(ij,h)` can be determined. These computations can be inserted next to the computation of `c_react(ij,h)` in procedure `ComputeLinearization`. The objective `TotalFlowCost` now should sum over all scenarios as well. They all have the same probability, so the devision by 729 can be done at once. The procedure `DetermineLosses` then solves the mathematical program (already defined in the last section), finding values for all flow variables. Afterwards it computes the line losses resulting from those power flows, summing over the scenarios in order to take the expectation.

As to the correct unities, recall from chapter 3 the expression for the powerflows $p_{ij} = \frac{v^2}{x_{ij}}(\theta_i - \theta_j)$ and for the power loss $p_{loss} = \frac{r_{ij}v^2}{x_{ij}^2}(\theta_i - \theta_j)^2$. Therefore,

$$p_{loss} = \frac{r_{ij}p_{ij}^2}{v^2}.$$

When power flows p_{ij} are expressed in MW and voltages in kV, the losses are also expressed in MW. Because this factor v^2 is equal for all terms, the devision by $(380kV)^2$ can be done at once afterwards. Note that the power flow solution itself is not influenced by the factor v^2, because the voltage is assumed equal throughout the grid.

```
PARAMETER:
   identifier   :  c_resist
   index domain :  (ij,h)

VARIABLE:
   identifier :  TotalFlowCost
   range      :  free
   definition :  sum[(ij,h,v,s),
                 c_react(ij,h)*
                 Flow(ij,h,v,s)]/729
```

```
PARAMETER:
   identifier :  TotalExpectedLosses

PROCEDURE
  identifier :  DetermineLosses
  body       :
    solve DetermineFlow;
    TotalExpectedLosses:=sum[(ij,
    h,v,s), c_resist(ij,h)*Flow(
    ij,h,v,s)]/729;

ENDPROCEDURE  ;
```

Above a `Placement` parameter was defined where every node had 3 wind turbines attached to it. But these values can easily be read from a file, and by varying the values one can compare the expected line losses resulting from various placement decisions.

For every scenario $s \in S$ a set of node balances is given and flow variables should be determined. But the flows of two different scenarios do not influence eachother. So the above formulation needs not be regarded as one giant network flow, but may be seen as $\#S$ separate network flows. Solving them subsequently will have a positive influence on the solution time needed by Aimms. Observe that for two subsequent scenarios, the balances on the node sets look very similar: only wind power is scenario-dependent, and even the wind speeds of two subsequent scenarios do not differ that much. Therefore solution time for a scenario $s \geq 2$ can be strongly improved by taking as a starting solution the flow variable values of the former scenario. In this way the whole problem can be solved by 1 network flow on n nodes and $n \cdot H$ arcs, and $\#S - 1$ sensitivity analyses.

7.4 Placement by direct minimization of the objective

Although the objective function of the mathematical program used in the two previous sections does not represent the expected energy losses, it is of the same structure as the formula for `TotalExpectedLosses`: it is a separable quadratic function of the flows between all nodes. This suggests one more intermediate step before defining the full model.

By making the placement vector an integer variable, and adding the constraint that the sum over its elements should equal a particular number `NumberOfWindmills`, the model becomes a placement problem. Of course `ActualWindPower`, `TotalWindPower` and `Supplies` depend on `Placement`, so they should be turned into variables as well. But nothing changes in their definition.

In this way Aimms will determine an optimal placement, not by the criterion of (in the terms of chapter 3) minimizing $\sum_{ij} r_{ij} p_{ij}^2$, that is, the total real power loss, but by minimizing the similar, but certainly different function $\sum_{ij} x_{ij} p_{ij}^2$.

Note that in the extremely special case that the proportions of the line reactances and resistances $\frac{x_{ij}}{r_{ij}}$ are equal for the whole network, the same arguments will minimize both functions; and therefore the direct minimization of the objective function would immediately yield a placement that also minimizes transportation losses.

7.5 Full model

The moment has come to define the full model. This consists of another five steps. First define `PrimalValue(s)`: for each scenario the cost of the flow based on the `c_react(ij)` parameters. Second, add for each scenario a set of node potentials `NodePotentials(i,s)`. Third, define `DualValue(s)`, for each scenario the dual of the cost of the flow based on the `c_react(ij)` parameters. Fourth: add `DualConstraint(s)`, requiring for each scenario `s` that the `PrimalValue` equals the DualValue. The consequence is that in every scenario the power flow minimizes $\sum_{ij} x_{ij} p_{ij}^2$. Fifth: take the TotalExpectedLosses (as defined above) as the objective function.

In Aimms this looks like:

```
VARIABLE:
   identifier   :  PrimalValue
   index domain :  s
   range        :  free
   definition   :  sum[(ij,h,v),
                   c_react(ij,h)*
                   Flow(ij,h,v,s)]

VARIABLE:
   identifier   :  NodePotentials
   index domain :  (i,s)
   range        :  free

VARIABLE:
   identifier   :  DualValue
   index domain :  s
   range        :  free
   definition   :  sum[i,Supply(i,s)*NodePotentials(i,s)] -
                   sum[(ij,h,v), u(ij,h)*max(0,-c_react(ij,h)+NodePotent
                   ials(FromB2(ij,h,v),s)-NodePotentials(ToB2(ij,h,v),s))]
```

```
CONSTRAINT:
   identifier   :  DualConstraint
   index domain :  (s)
   definition   :  DualValue(s)=
                   PrimalValue(s)

VARIABLE:
   identifier :  TotalExpectedLosses
   range      :  free
   definition :  sum[(ij,h,v,s),
                 c_resist(ij,h)*
                 Flow(ij,h,v,s)]
                 /729,
```

and the procedure

```
PROCEDURE
  identifier :  DeterminePlacement
  body       :
    ReadNetworkData;
    ComputeLinearization;
    ReadWindData;
    OptimalPlacement.CallbackAOA := 'OuterApprox::BasicAlgorithm';
    OuterApprox::IterationMax := 5;   ! Optional
    solve OptimalPlacement;

ENDPROCEDURE  ;
```

will read the data, and then try to solve the mixed integer nonlinear optimization problem (MINLP) by application of the so-called Outer Approximation (AOA) Algorithm. The AOA solver in Aimms iterates between a Non-linear program (NLP) relaxation of the problem and a small scale linear Mixed Integer Program (MIP), in order to solve the MINLP. However this algorithm in Aimms seems unable to solve the problem even for reduced precision in the linearization and with a smaller scenario space. It keeps iterating until a memory error occurs. There was no need to do further research into the cause of this because of the availability of the two heuristics discussed in the next chapter.

8 Heuristic methods

Two heuristics based on the partial model implementations from chapter 7 are explained and discussed. Relevant observations are presented concisely point by point in order to capture the main idea, but to leaving the details of implementation to the reader.

8.1 Greedy heuristic

The idea is based on section 7.3; by subsequently placing single wind mills a good approximation of the optimal configuration of (for example) 30 wind mills can be achieved.

There are only 29 nodes in our network. Therefore in order to find the optimal location (least expected losses) for a single wind mill, only 29 placement possibilities need to be computed and compared. At the best node one wind mill is placed. Next for the second wind mill again 29 possibilities need to be checked. The complexity of this heuristic equals the complexity of computing expected losses for a given placement (section 7.3) multiplied by the product of the number of nodes n in the network and number of wind mills l that should be placed.

Particularly when network size increases, this heuristic may become problematic. In that case a selection of the nodes to be checked should be made beforehand.

For large number of wind mills this heuristic will produce near-optimal solutions, as may be conjectured from the placement results in chapter 9. Placing additional units does not seem to affect the optimal configuration of the previously placed ones. Poorest performance is to be expected when a small number of units with large individual supplies should be placed. In that case the heuristic below or rigorous solution is advised.

8.2 Adjusted impedance parameter heuristic

This heuristic is based on section 7.4; the idea arises from the observation that the functions $\sum_{ij} r_{ij} p_{ij}^2$ and $\sum_{ij} x_{ij} p_{ij}^2$ have the same structure.

It was already observed in chapter 7 that if for all branches ij, the fraction $\frac{x_{ij}}{r_{ij}}$ is the same, then minimization of $\sum_{ij} x_{ij} p_{ij}^2$ would immediately yield a placement that minimizes losses.

Now minimize objective function $\sum_{ij} c_{ij} p_{ij}^2$, where the cost parameters c_{ij} are a mixture of reactance x_{ij} and resistance r_{ij}. Minimization of $\sum_{ij} x_{ij} p_{ij}^2$ is needed for correct power flows; minimization of $\sum_{ij} r_{ij} p_{ij}^2$ yields loss minimizing objective. So a certain tradeoff has to be made which aspect is more important for an arc.

One simple suggestion could be to take cost parameters $c_{ij} = \frac{1}{2}(r_{ij} + x_{ij})$; or another

$$c_{ij'} = r_{ij'} \cdot \left(\frac{1}{m} \sum_{ij} \frac{r_{ij}}{x_{ij}} \right)^{-1} .$$

Two extreme parameter choices can be identified. First: disregard resistances completely (that is, take $c_{ij} = x_{ij}$), which respects power flows, but disregards losses[9]. And second: disregard reactances completely (that is, take $c_{ij} = r_{ij}$), which does minimize losses, but does not respect power flows. [Li Yang Chen 2010] makes this second choice (he does not respect power flows; only flow balances).

It is difficult to say something intuitive about the quality performance of this heuristic; in general it will perform poorer if $\frac{x_{ij}}{r_{ij}}$ differs much for different arcs. Implementation is easy, and speed of the heuristic is high enough to achieve the computation times mentioned in table 5.

[9]This is the choice made to produce the results in chapter 9.

9 Results

In this chapter the results of the performed computations are presented. In order to visualize the tradeoffs mentioned in section 1.1, first some preliminary results are shown. The tradeoff between high expected value and small variance of the total wind power, is illustrated in section 9.1. The tradeoff between local and central placement for stochastic supply, is illustrated in section 9.2.

In section 9.3 the results of the full model are shown. Next it is discussed how the results can be interpreted by combining the results of sections 9.1 (where only wind zones are regarded) and the results of 9.2 (where only network structure of the power grid is regarded).

9.1 Significance of wind zones

In figure 17, normalized output power histograms are shown for 6 wind turbines. The left picture applies to the situation where all outputs are 100% correlated, i.e., all turbines are in a single zone. The region of Groningen and Friesland is taken as an example, because it has medium expected wind speed. The right picture applies when the turbines are equally distributed over 6 different wind zones. The results are obtained from the data used and discussed in chapter 6.

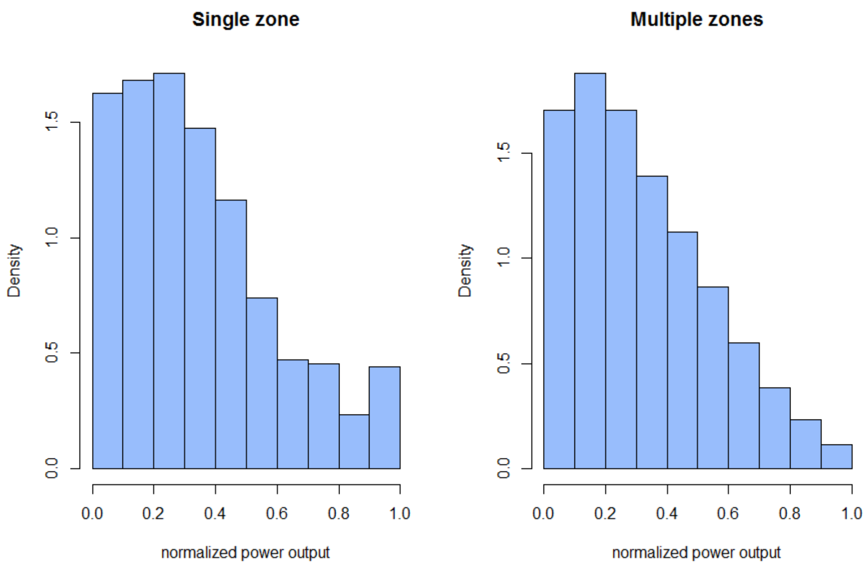


Figure 17: Histograms for power output in single and multiple zones.

Although the effect is not very strong, one can see that the distribution for the single zone is significantly flatter than the distribution for multiple zones. In fact, the variance of the single zone output is 1.5 times higher than for multiple zones. This suggests that there could be benefits from spreading wind turbines over the different wind zones. Of course the benefits in larger countries would be greater (for example, [Li Yang Chen 2010] presents similar histograms for various sites in the United States, indicating a stronger effect in variance reduction). Figure 17 was obtained in R as follows:

```
> par(mfcol=c(1,2))
> gh[aa<minwind | aa>maxwind]=0
> gh[refwind<aa & aa<=maxwind]=maxpow
> gh[minwind<=aa & aa<=refwind]=(aa[minwind<=aa & aa<=refwind]-
minwind)*maxpow/(refwind-minwind)
> hist(6*gh[,6]/15000, col="grey", breaks=11, main="Single zone", xlab="norm
alized power output", freq=FALSE)
```

```
> hist((gh[,1]+gh[,2]+gh[,3]+gh[,4]+gh[,5]+gh[,6])/15000, col="grey", breaks
=11, main="Multiple zones", xlab="normalized power output", freq=FALSE)
```

Now suppose that we distribute 20 wind turbines over the various zones, yielding a placement vector p which contains for every zone $i = 1, \ldots, 6$ the number of wind turbines p_i that are allocated to that zone. Define the output of a single wind turbine in zone i as X_i, $i = 1, \ldots, 6$ and define the total power output of the 20 turbines $X_p = \sum_{i=1}^{6} p_i X_i$.

For various reasons a high variance of the output could be undesirable. In our context high variance of power output would mean that it is more difficult to match it to local demand, resulting in higher power losses by transportation. In finance and trade, high variance of production means less predictable volumes to be traded, resulting in lower selling prices. In order to compare expected value and variance, it has been proposed (for example [Steinbach 2001]) to maximize the function $f(p) = E[X_p] - \lambda Var(X_p)$, where λ represents the costs of volatility. This so-called *mean-variance model* is common in financial optimization and other disciplines. The variance of X_p can be expressed as

$$var(X_p) = var(\sum_{i=1}^{6} p_i X_i) = \sum_{i=1}^{6}\sum_{j=1}^{6} cov(p_i X_i, p_j X_j) = \sum_{i=1}^{6}\sum_{j=1}^{6} p_i p_j cov(X_i, X_j),$$

where the covariance matrix Σ of the X_i is known from the wind data discussed in chapter 6. Let $\mu_i = E[X_i]$, and let $\sigma_{i,j} = cov(X_i, X_j)$ be fixed parameters. Then the function $f(p)$ can be written in the nonlinear form

$$f(p) = \sum_{i=1}^{6} p_i \mu_i - \lambda \sum_{i=1}^{6}\sum_{j=1}^{6} p_i p_j \cdot \sigma_{i,j},$$

and its variables p_i, $i = 1, \ldots, 6$ are subject to the constraints $p_i \in \{0, \ldots, 20\}$, and $\sum_{i=1}^{6} p_i = 20$. The parameters are given by

$$\mu = \begin{pmatrix} 65.77 \\ 56.95 \\ 37.00 \\ 41.74 \\ 41.18 \\ 50.47 \end{pmatrix}, \text{ and } \Sigma = \begin{pmatrix} 1008.68 & 699.088 & 480.541 & 542.839 & 585.349 & 538.844 \\ 699.089 & 781.241 & 415.072 & 438.205 & 455.226 & 569.415 \\ 480.541 & 415.072 & 389.708 & 381.418 & 394.284 & 377.386 \\ 542.839 & 438.205 & 381.418 & 526.130 & 479.518 & 409.670 \\ 585.349 & 455.226 & 394.284 & 479.518 & 591.539 & 412.965 \\ 538.844 & 569.415 & 377.386 & 409.670 & 412.965 & 678.272 \end{pmatrix}.$$

Again Aimms can solve this for various λ. The problem is a Mixed Integer Quadratic Programming (MIQP) problem with 6 variables. Table 2 shows results for various values of λ. Because of the high values in matrix Σ compared to μ, the range of λ for which interesting changes occur, is quite near to zero ($\lambda \in [0, 0.01]$).

100λ	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	1
North Sea Coast	20	15	9	4	2	0	0	0	0
Zuiderzee	0	4	4	3	3	2	2	1	0
Inlands	0	0	0	6	9	11	13	14	15
Limburg	0	0	0	2	2	3	2	2	2
WN Brabant	0	0	0	0	0	0	0	0	0
Groningen Friesl	0	1	7	5	4	4	3	3	3

Table 2: Distribution of turbines for various λ in the mean variance model

For $\lambda = 0$, naturally all turbines are placed in the zone with highest average wind speed. For λ too high, the term $-\lambda Var(X_p)$ becomes dominant, and most turbines are allocated to the zone ‘Inlands’. This zone has lowest variance, but it has also low expected value.

The value of λ that should be taken for practical optimization, is determined by external factors (market volatility, share of wind power in the generation portfolio of the power company). One can clearly observe that the optimal wind mill distribution for a power company strongly depends on the value of its λ.

9.2 Network structure

In the former section only wind speed behaviour was taken into account, while network structure of the power grid was neglected. An optimal dispersion over the various wind zones was achieved, depending on the relative importance of expectation and variance of the output. In this section, by contrast, the structure of the power grid is analysed. This is done by placing Distributed Generation (DG) units in the network, and assessing their effect on the total line losses on the grid. Optimal locations for both deterministic and stochastic DG units are presented. The problem is similar to the placement of wind turbines, but simpler because wind speed correlation structure is disregarded.

As a first orientation into the structure and the load of the HV grid, the simple power flow problem can be solved. It yields the following observations:

-The global structure of the grid resembles two connected rings, as depicted in the left part of figure 18. The province of Groningen and the loose ends in the west are exceptions to this, but the major part of the transportation is taking place at one cycle (mainly from West to East).

-For a achieving minimum transportation losses, DG units are very likely placed on nodes with negative supply. Therefore if node 619 (Diemen), which has positive supply, is moved to the wind zone 'North Sea Coast', this is likely to have no consequences for any solutions.

-Almost all nodes in 'North Sea Coast' and in 'West Noord-Brabant' are supply nodes.[10] It is unattractive to place DG units there. When these two zones are united, and the rest of the nodes are sorted according to their wind zones, a picture can be made of the power exchanges between the several zones. This is shown in the right part of figure 18. Net supplies are cumputed per zone, and the power exchange between the various zones is the sum over one or two transmission lines. Both quantities are in MW.

-From these pictures it is intuitively clear that DG units will be placed most likely in the 'Inlands' or 'Limburg'.

-Total load on the grid is 13325MW. About half of it is exchanged between the zones. The network losses for the standard power flow solution amount to 7.5MW.

In order to investigate network preferences for the placement of DG, tables 3 and 4 are presented. They contain the results of a sequence of placement problems for varying DG penetration, and for varying size of the generation units. By penetration is meant: total added generation capacity divided by total original generation capacity (13325MW).

In the limit of 100% penetration and small enough unit size, it is clear that transportation losses can be made to vanish. When distributed generation equals demand at every node, no power transportation takes place at all.

The extreme case of placing only one DG unit can be used to obtain the notion of a kind of centrality of a node. When placing one unit of 1MW, the location which minimizes the loss is the point for which the network losses are most sensitive to a change in demand. One DG unit with 100% penetration yields the best location if the whole network would be powered from one source node.[11] The resulting source node turns out to be Doetinchem for small size, Dodewaard when a size of $6000MW$ is reached, and Boxmeer when 100% penetration is approached.

[10] The large supply at Beverwijk is explained by offshore wind farms at the height of Noord Holland. The large supplies at Maasvlakte and Geertruidenberg are explained by the large number of Combined Heat and Power (CHP) systems used by greenhouses and industry. Borssele is home to a nuclear power plant.

[11] Note that this is purely theoretical because of the following. In our model extra DG power is compensated by decreasing supply of the original supply nodes. This choice was made in order to approximate the situation in which all commercial power plants in the country take equal share in the compensation. When the share of DG becomes too large, this approximation fails.

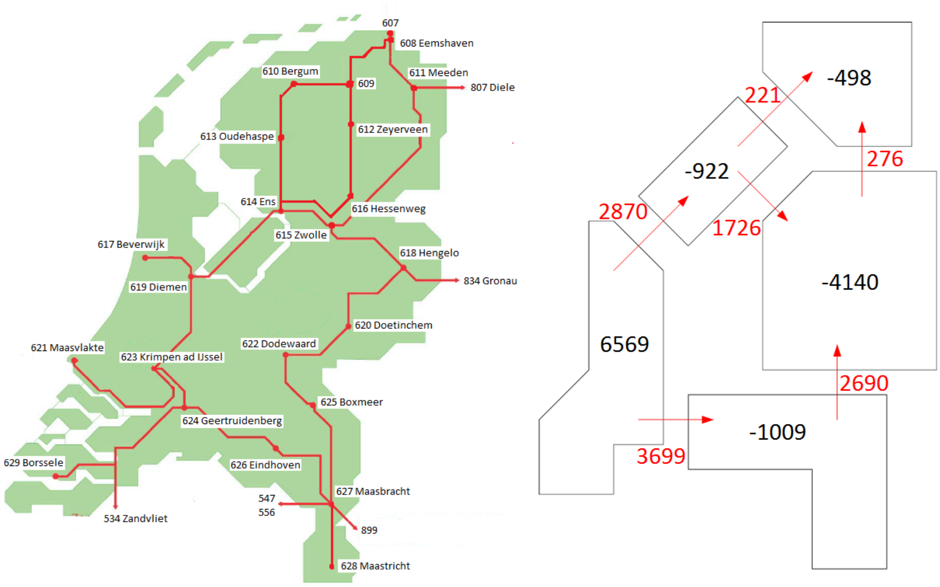


Figure 18: Power exchanges in MW between the wind zones for the standard power flow solution

The computations were done using the heuristic described in section 8.2. Table 3 confirms the intuition that both smaller unit size and increasing DG penetration have positive effect on transportation losses.

#DG units	30	60	120	6	12	24	3	6
unit size (MW)	100	100	100	500	500	500	2000	2000
total DG capacity (MW)	3000	6000	12000	3000	6000	12000	6000	12000
Coast and Brabant	0	0	1	0	0	0	0	0
Zuiderzee	0	0	6	0	0	1	0	0
Inlands	29	41	44	5	8	9	3	2
Limburg	0	9	48	0	2	10	0	3
Groningen Friesland	1	10	18	1	2	4	0	1
Network losses (MW)	3.71	1.60	0.04	3.72	1.59	0.06	1.82	0.46

Table 3: Dispersion of deterministic DG over wind zones

The most remarkable feature of these results is the enormous sensitivity of the total losses to DG. Even for small penetration the losses drop to half the original value. Certainly for lower penetration, the size of the DG units makes little difference. Next it is notable that the first DG units are mostly placed in the inlands, but at a certain point Limburg takes over the central role.

The same computations were done for stochastic DG units: a placement is found for which expected transportation losses are minimized. For each case, the expected output of the DG units equals the size of its deterministic equivalent in the table above. Three scenarios were used: one with zero DG output, one with its expected output, and one with output of twice the expectation. The results are presented in table 4.

It immediately stands out that the strong positive effect of deterministic DG on the transportation losses, is partly neutralized when DG units are stochastic. Interestingly, all penetrations yield comparable resulting net losses. Furthermore, stochastic DG units seem to be more dispersed than their deterministic equivalents. The strongest dispersion effect is seen for high penetration. However, it should be observed that for high expected penetration, there is excess of DG power in the

#DG units	30	60	120	6	12	24	3	6
mean unit output (MW)	100	100	100	500	500	500	2000	2000
mean DG capacity (MW)	3000	6000	12000	3000	6000	12000	6000	12000
Coast and Brabant	0	0	22	0	0	5	0	1
Zuiderzee	0	1	6	0	0	1	0	0
Inlands	24	27	29	5	5	6	2	1
Limburg	2	22	47	0	5	9	1	3
Groningen Friesland	4	10	16	1	2	3	0	1
Network losses (MW)	4.287	3.17	3.0	4.3	3.209	3.0	3.651	4.055

Table 4: Dispersion of stochastic DG over wind zones

third scenario (where all units produce twice their mean output). This results in the situation that the original supply nodes become demand nodes, making the North Sea Coast an attractive zone to place the DG units. As mentioned before, the results for this situation should be distrusted.

9.3 Wind turbine placement

In this section results of the solutions to the wind turbine placement problem are presented and discussed. The computations were done using the heuristic discussed in section 8.2. Recall that the maximum output power of the turbine that was regarded in chapter 6 equals 2.5MW. The network results presented in the former section make clear that this quantity is almost neglegible compared to the total load. Therefore, the output power was multiplied by successively 40, 200 and 800 in order to obtain interesting results. That is, the 'DG units' of maximum unit output of $100MW$, $500MW$ and $2000MW$ correspond to groups of 40, 200 and 800 large wind mills.

The 729 scenarios that were simulated in chapter 6 were used. These wind power scenarios reflect mutual correlation between the zones, and the Rayleigh distributed wind speeds for each zone apart. The resulting mathematical program is a mixed integer quadratic program (MIQP) with 64155 constraints and 88240 variables, of which 29 integer. Aimms' CPLEX 12.4 solver is able to solve this within a few minutes on a standard PC. It can be observed that both unit size and number of units have a negative influence on the computation time. (A larger number of units obviously increases the number of basic feasible solutions strongly. It is less evident why larger unit size requires longer computation time.)

Consider the wind turbine placement problem. Placement of 30 wind turbines with maximum output of 100MW yields the result that all turbines are placed in Groningen, the average total output being less than 600MW. This shows that the solution of the wind turbine placement problem may differ from what one might expect when just wind or network data are considered. Highest wind speeds are to be found at the North Sea Coast, but there are no demand nodes there. Second highest wind speeds occur in the 'Zuiderzee' region which does have a shortage of 922MW. But apparently lower production in Groningen can cause a greater decrease in power transportation than a somewhat higher production in the 'Zuiderzee' region would cause. However for the Inlands and Limburg (which have an even more attractive position on the network for DG placement) wind speed is just too low.

Table 5 below shows some results for higher penetrations.

It is interesting to see how the Inlands, which seemed so attractive both in the mean variance model for wind power, and in the network structure analysis, are less attractive in the wind turbine placement problem. Units are placed in the Inlands only after Groningen has reached a kind of saturation, and there are already many wind turbines in Limburg. This phenomenon might occur because the expected wind generation is simply too low in the Inlands; but seeing that the difference in expected outputs between Inlands and Limburg is smaller than between Groningen and Zuiderzee, some other reason seems likely.

When there are more than 30 units of 100MW (that is, 3000MW of installed wind capacity) in

#wind turbines	60	120	240	12	24	48	3	6
max unit output (MW)	100	100	100	500	500	500	2000	2000
mean DG capacity (MW)	1120	2060	3710	1140	2072	3750	1140	2140
Coast and Brabant	0	0	0	0	0	0	0	0
Zuiderzee	0	0	7	0	0	2	0	0
Inlands	0	1	57	0	0	11	0	0
Limburg	23	70	120	4	14	24	1	3
Groningen Friesland	37	49	56	8	10	11	2	3
Network losses (MW)	6.309	5.438	4.282	6.829	5.43	4.272	6.829	5.448
Solution time	157s	-	341s	-	-	-	374s	524s

Table 5: Dispersion of wind turbines over wind zones. In the lowest row some computation times are displayed, which it took Aimms 3.12 on a computer with 4 GB RAM, Intel Duo core 3.0GHz processor, running under 64-bit Windows 7.

Groningen, there is excess of wind power in that region during the strong wind scenarios. Some of that power will flow to the neighbouring Inlands. Consequently, Limburg is more attractive than the Inlands because of its greater distance to Groningen.

Also neigbouring zones do not have full correlations: highest wind generation scenarios of one zone do not necessarily coincide with highest generation in the other zone. The probability of local power excess may therefore be decreased by placing units in the neigbouring zone. This may explain why the results for wind turbine placement show a higher spreading of generation units than the stochastic DG placement regarded in the former section (where correlations were 100%).

10 Relevance of the research question

In this chapter the topics and results of this thesis are discussed from a practical point of view. The sections below contain five different aspects of DG, stochastic generation and loss minimization in practice.

10.1 Realistic scale of DG and losses

According to [Energy in the Netherlands 2011], page 9, $126.6 \cdot 10^9 kWh$ of electrical energy was transported over the Dutch power grid during the year 2011. $4.45 \cdot 10^9 kWh$ of that energy was lost during transportation, which is around 3.5%. This exceptional efficiency is due to the high quality of the Dutch grid infrastructure and operation. Other factors which contribute to this result is the large proportion of underground cables (which would be more difficult with rocky soil), and high population density.

From the tables on page 9 and page 65 in [Energy in the Netherlands 2011], the average price of a kWh excluding taxes, can be deduced to be about 0.10 euros. Multiplying this price by the suffered losses yields an economic value of 445 million euros which was lost on the Dutch grid during 2011 by technical inefficiency. From this economic point of view, any significant improvement of grid efficiency seems worth to pursue.

For a good estimate of the economic value of the results in chapter 9, more data would be needed. There it was concluded that losses on the HV grid could be reduced significantly by relatively small penetration of deterministic DG, or by somewhat higher penetration of wind generation. To compute the economic value of this reduction, it should be known what percentage of the losses occurs in the HV grid. (These data are not available in [Energy in the Netherlands 2011], though doubtless they are known to TenneT.) In general it can be said that the majority of losses occur in the lower voltage grids by their enormous lengths (LV $0.4kV$: $200,000km$, MV $0.4 - 25kV$: $100,000km$, HV $\geq 50kV$: $11,000km$), even with much smaller power flow.

The proportion of DG in the Netherlands is very high (around 25% of produced electricity). This is mainly due to the Combined Heat and Power (CHP) systems in chemical industry and horticulture. (Because these businesses tend to cluster in the same area, this DG does not contribute to grid efficiency, but rather is a frequent source for congestion.) The penetration of installed wind generation capacity in the Netherlands was $2200MW$, that is about 9% in 2011. But there is a difference between installed capacity and actual production as can be seen from figure 19.

Although installed wind capacity is about 9% of total installed capacity, the share of electricity produced from wind is about 2.5%. Most of the time wind turbines do not operate at full power. This agrees with the results in section 9.3 where many units had to be placed in order to get significant mean output. Typically nuclear power plants have higher share in produced electricity than in installed capacity. This is because from all conventional power sources, nuclear power has lowest marginal costs, and production is not very flexible. Gas plants have highest marginal cost and are flexible in operation, and are therefore typically used to compensate for fluctuations in demand or other production. Coal plants are in both respects intermediate. Furthermore it can be observed that hydropower by far is the most successful renewable energy source, but it is not abundantly available in every country. Denmark is the only country with significant share of wind energy in total production (17%).

[Reza 2006] has written a PhD thesis in which he analyses maximum penetration of DG for renewable power. He concludes that a penetration of 30% is possible without the occurrence of stability problems. It is clear that this scenario is still far ahead, but wind energy is (except for hydropower) the most promising renewable DG. This distant future scenario of 30% is the highest penetration scenario that was analyzed in section 9.3. It can be seen from table 5 that for this penetration a high spreading should be advised.

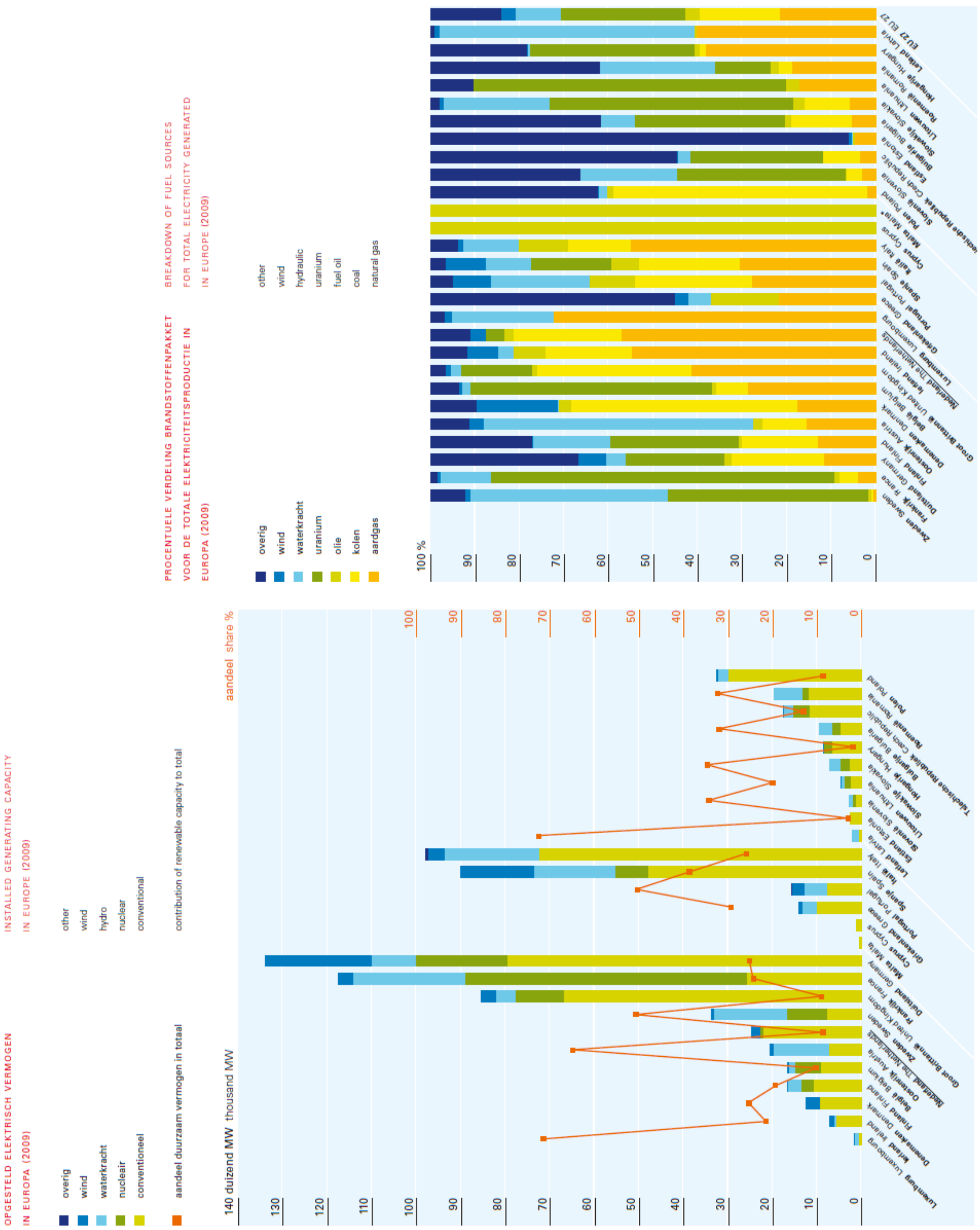


Figure 19: Share of different power sources

10.2 DG for reducing transportation costs

The costs for transportation of electricity consist of building and maintaining the infrastructure itself (the grid), and the losses occuring during the actual transportation. In certain contexts, DG can be useful for the reduction of both.

Note that for traditional power plants (coal, gas, nuclear), large scale generation is most efficient. Only in more recent times it became possible to build small scale power plants (particularly gas) of reasonable efficiency. Therefore research into these uses of DG emerged only some 30 years ago. It slightly predated the appearance of CHP-systems and of renewable power generation like wind turbines and photovoltaic cells (PV). Because these 'green' generation units also had low capacity, and connect to lower voltage parts of the grid, they were also called DG. Renewable, stochastic DG can be used for the same purposes (but in a less ideal way) as the small scale gas plants.

In what manner DG is used for reduction of grid reinforcement costs, can be understood as follows. For example if a certain group of consumers sometimes overloads the transformers or transmission lines by which it is connected to the higher voltage grid, two things can be done to solve the problem: first, the utility could build larger transformers or thicker transmission lines. Second, a small scale DG unit could be built in the neighbourhood of the consumers. This second solution simply relieves the load on the overloaded components. Sometimes it is cheaper or faster to build the DG unit, especially when a long distance of underground cable has to be replaced.

DG can also be used for reduction of power losses. It is clear that by placing DG in the neighbourhood of consumers, less power is transported, and therefore losses decrease. The effects of this are largest, when great distances have to be covered, or when power is transported at a low voltage. It is usually too expensive to build a high voltage connection only to a small number of consumers. So in regions where small villages are far remote from other towns, this use of DG is most relevant. It can also be relevant in mountainous areas where low voltage distribution lines do not travel underground, because it is too expensive to dig the ditches. Overhead lines suffer significantly higher losses than underground lines (because of changing weather conditions there is more leakage due to humidity and higher resistance due to temperature changes), which makes power loss reduction more profitable in this context than if underground lines can be used.

10.3 Power companies

In the Netherlands, liberalization of the energy market was established in 2004. Since then, production of power has become a commercial activity, performed by commercial companies. The responsibility for expansion, maintenance, and operation of the grid still lies with the utilities. Before 2004 both these activities were performed by the same utility, and consequently generation expansion and grid expansion could be planned simultaneously. But now construction of new power plants is commercial, and in principle independent of the grid situation. Commercial companies maximize profit, and therefore they look for cheap building sites and low operational costs (for heat production this means good logistic environments and availability of cooling water, for wind turbines this means high wind speed).

Location selection for wind turbines in liberalized electricity markets is quite independent of the power grid. However, there are two remarks to be made, as it can lead to negative consequences when not managed properly.

First: in power grids where congestion is a frequent phenomenon, this might cause a loss of generated power or in a prescribed restriction of the maximum allowed generation. In either case, revenues are missed. In this case, the grid situation has a direct effect on the effective operational costs of the plant. [Li Yang Chen 2010] for example, calls this 'loss of load', and he incorporates this factor into his optimization model for selecting locations for wind turbines on the basis of profit maximization. In congestion sensitive areas, economic value of a wind turbine can be directly related to its location on the grid. This seems more relevant for the American context of [Li Yang Chen 2010] than for the Dutch situation. See however the remarks in section

10.4.

Second: in the current situation in the Netherlands, the costs for the power lost during transportation over the grid are divided proportionally to all participants in production and consumption. The costs for a producer are independent of the locations of his production units. Therefore, no economic incentive for producers exists which promotes the reduction of transportation losses. There is no commercial reason to select locations for their production units, that have an advantageous effect on grid efficiency. One solution to improve on this situation, is to make the price of electricity dependent on the location on the grid where it is produced or consumed. In that way it could be tried to redesign the market such that producers would make higher profits from power produced at locations where transportation losses are low. We return to this in section 10.5.

10.4 Grid operators

It was stated above that the utilities that carry responsibility for the power grid (the *grid operators*) have nothing to say in the location selection of new production units. The ideal of market liberalization of 2004 is that maintenance of infrastructure and commercial activity on the infrastructure are completely independent. Therefore by law TenneT is bound to accept every request to connect a new power plant or other generation unit to the grid. However, there are some areas in which this new generator would result in congestion of the transmission lines in the neighbourhood. So in practice TenneT uses a waiting list for these new requests, until capacity expansion of the grid allows for the new generation unit. Figure 20[12] shows that it is predominantly those areas of the Netherlands which are apt for wind mills, that also suffer from these waiting lists. The black arrow indicates a waiting list, the green bars indicate currently installed wind power in the area.

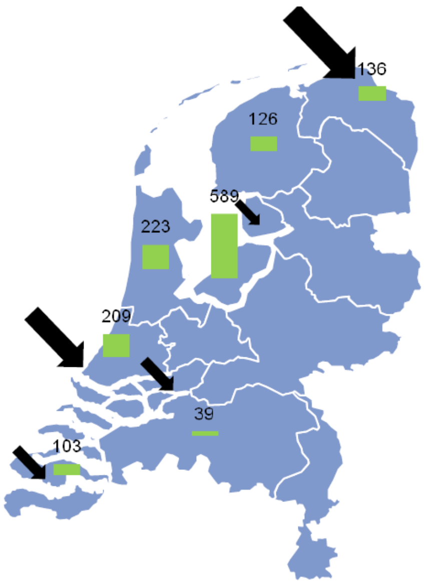


Figure 20: Waiting lists for new generation units compared to amount of installed wind power.

It can be concluded that in practice for wind generation, generation expansion and grid expansion are not independent at all. Grid operators influence the location selection of wind turbines by denying permission if grid capacity is too low.

Moreover, in [Toveren met Vermogen 2011], page 54, Prof. van der Sluis asserts that grid operators at the moment look at the maximum scenario case. That is, in order to analyse if the connection of

[12]source: [Deloitte 2007]. The picture shows the situation in 2006, but we expect it to reflect the current situation as well.

a certain wind turbine would cause congestion, they assume the wind turbine to produce maximum power. He suggests that if grid operators change their criterion, and demand that only 95% of the time grid capacity suffices, many more wind turbines could be installed at the same grid. Van der Sluis even predicts that in the future consumers will pay less when they accept higher risk of power shortage (which results in voltage drop) due to congestion.

In this context, it might interest grid operators which have to deal with a new request for the building of some wind turbines in an already saturated network region, to find a place for the turbines which minimizes congestion probability. This placement problem is related to the problem discussed in this thesis.

10.5 Locational Marginal Pricing

Locational Marginal Pricing (LMP) originally sprung up from the idea to influence power demand and supply by price changes in order to avoid congestion on the grid. For example if a transmission line to some consumers is in danger of becoming congested, LMP methods increase the price of electricity for those consumers. With high enough price elasticity, economic theory states that demand will decrease, and in that way the danger of congestion will be averted.

This is an attempt to achieve technical goals (congestion management) by economic means (market design). There exist a variety of methods based on locational or regional pricing. LMP methods seem most appropriate for networks with weakly connected parts, where these parts themselves have high inner connectivity. Then it seems reasonable to let the price depend on the conditions of the local part of the network. Research into European power market integration is an example where such questions are currently encountered. Because of historical development, the distinct countries have high inner connectivity, but there is lower transmission capacity between countries (although there is certainly progress, cf. the recent construction of both NorNed and BritNed cables[13]). With one central European market without price differentiation per country, border capacities would easily be exceeded. Countries like US, Australia and New Zealand already have some kind of implementation of LMP. TenneT is currently looking into possible advantages for the Netherlands.

The problem with electricity markets is that there is very low elasticity, certainly for private consumers. Moreover, it can be questioned if it is desirable that families in regions with higher congestion risk should pay more for their electricity than families in regions with low congestion risk. Maybe prices should be differentiated for producers while they remain fixed for consumers. There is an infinite number of different implementations possible. This location dependent pricing might also serve as a financial incentive for power producers to place production units at sites where they contribute to higher grid efficiency.

This topic is large enough for a research of its own. Before implementation of such methods, it should also be checked if the price differences have actual effect on the decisions of power companies.

[13]High Voltage cables connecting the Netherlands with respectively Norway and England.

11 Summary and conclusions

In this chapter the content of this thesis is summarized; conclusions are drawn from the computational results; and suggestions are made concerning topics for potential further research.

11.1 Summary

In this thesis:

- an introduction was provided into power grid structures and into modeling of power flows with network losses;
- an overview was given of relevant literature concerning DC power flow modeling, power grid design, generation expansion planning, power loss minimization, and the use of stochastic programming methods;
- suggestions were made to improve on the currently available DC power flow models, by incorporating quadratic network losses;
- a mathematical model was formulated in order to capture the wind power Generation Expansion Planning problem for loss minimization in a so-called two stage stochastic program;
- network flow duality was applied such as to arrive at a (possibly novel) standard LP formulation for the stochastic Generation Expansion Planning problem for loss minimization;
- a procedure was presented to simulate appropriate scenarios for steady state wind power studies in multiple areas;
- implementation of the model and possible heuristics were explained;
- computational results for the Dutch HV network were presented and discussed;
- the relation of the research question of this thesis to current situation and developments in the electricity sector were discussed. It was conjectured how DG placement problems might be of value to practical power grid design and planning studies in congestion sensitive areas. Special mention was made of the field of Locational Marginal Pricing.

11.2 Conclusions

From the computational results it was possible to conclude:

- that the effect of reducing transmission losses by placing DG units is significantly weaker for stochastic than for deterministic units;
- that in order to achieve high transmission efficiency, stochastic DG units should be placed further apart than deterministic DG units;
- that there exists a non-trivial relation between the locations of wind turbines and transportation losses, but that interesting effects occur only for higher penetration;
- that the optimal distribution of wind energy over the Netherlands shows a higher spreading than can be explained merely from its stochasticity;
- that, consequently, the location dependent nature of wind energy is relevant for its effect on power grid performance;
- that because of the load situation of the Dutch HV grid Groningen and Friesland are the only regions with both attractive wind speeds, and where placement of wind turbines would result in reduction of transmission losses in the HV grid;

- that at the current scale of wind energy, a spreading strategy for wind turbines does not seem to be relevant, neither from a commercial point of view (variance reduction of total output for the Dutch market) nor from a technical point of view (reducing transmission losses)[14];

- however, that if in a future scenario wind energy were to acquire a significant share in HV power flows, location selection may become important for grid performance.

Following from the research into the context of power losses and wind energy generation, also some other observations were made.

- The loss minimizing uses of DG seem more relevant in large countries like US or Australia, than in a smaller and densely populated country like the Netherlands.

- Location selection of new wind turbines is done by individual commercial power companies, whereas costs for transportation are paid by all parties equally. Therefore in practice there is no financial incentive which stimulates locations that improve grid efficiency.

- LMP methods may open the possibility to let the causer of transportation losses pay for the loss. In such a scenario using DG for reducing losses could also have economic value.

11.3 Suggestions

There is a number of topics which were briefly touched upon in this thesis, but which it was not possible to elaborate. Several suggestions for further research are described below.

- The first extension this thesis calls for is an assessment of the quality of the two heuristics presented in chapter 8. This could be combined with computational testing of the solution methods. Most efficient implementation requires the sensitivity analysis described in section 7.3. In this way one of the advantages of the network flow approach to power flow analysis could be illustrated.

- A second extension would consider the scalability of the results presented in chapter 9. For smaller scale MV or LV networks, power flows are lower and therefore significant wind penetration is more realistic. MV grids cover smaller areas, so location dependence of wind power seems to be less relevant. However, in coastal or urban regions (which have strong difference in so called *roughness* of the surface), it is imaginable that it is still an important factor. Mean wind speeds and covariances for the different locations in combination with MV network data would be needed for this research.

- The idea to regard the DC power flow as a separable quadratic generalized cost flow (as discussed in section 3.7) could be elaborated. Comparison of computational performance between the customary DC methods, and the network flow approach could be carried out; assessment of the improvements that are achieved by incorporating linear or quadratic line losses could be done; and it could be investigated whether the improved method would improve on the guidelines of [Purchala et al. 2005] on the usefulness of the method.

- Research could be done into the question to which extend LMP-methods are already (or can be) equipped with transmission losses. The idea would be to let all parties pay for the losses they actually cause. This may help to let commercial interests coincide with grid efficiency. An overview could be made of the different methods to do this with their respective advantages and disadvantages.

- In stead of addressing transportation losses, it could be asked how to place (stochastic) DG in order to minimize the probability of congestion in the network. As discussed, congestion seems a more pressing worry for grid operators than transportation losses. The two themes

[14] This concerns the HV grid: whether spreading strategies on a MV network scale would be desirable from a technical point of view, should be investigated in a different case study (cf. the second suggestion in section 11.3).

certainly seem related, because both congestion and high losses occur for a high load of a line.

- Inspired by a project that currently runs at TNO (Flexiquest): one might wonder what the consequences would be if on a large scale consumers would change their consumption behaviour, based on electricity prices on the market. This kind of stochasticity would be location independent, because of the location independence of the deals on the APX market. Sampling distributions could be deduced from analysis of APX price fluctuations.

- The mean variance model from section 9.1 gives rise to the question at which scale of a country, or for which market behaviour, or at which penetration of wind energy, spreading of wind turbines would become interesting from a purely commercial point of view (to maximize expected market value).

References

[Ababei and Kavasseri 2011] Ababei, C. and Kavasseri, R., *Efficient network reconfiguration using minimum cost maximum flow-based branch exchanges and random walks-based loss estimations*, IEEE Transactions on Power Systems, vol. 26 no. 1 (2011).

[Ahuja et al. 1993] Ahuja, R.K., Magnanti, T.L., Orlin, J.B., *Network Flows*, Prentice-Hall, Englewood Cliffs (1993).

[Andersson 2008] Andersson, G., *Modeling and Analysis of Electric Power Systems*, Lecture notes for ETH Zürich (2008).

[Baran and Wu 1989] Baran, M.E., Wu, F.F., *Network reconfiguration in distribution systems for loss reduction and load balancing*, IEEE Transactions on Power Delivery, vol. 4 no. 2 (1989).

[Bertsimas and Tsitsiklis 1997] Bertsimas, D., Tsitsiklis, J.N., *Introduction to Linear Optimization*, Athena Scientific, Belmont, Massachusetts (1997).

[Bialek 2005] Bialek, J., *Approximate Model of European Interconnected System as a Benchmark System to Study Effects of Cross-Border Trades* IEEE Transactions on Power Systems, vol. 20 no. 2 (2005). http://www.see.ed.ac.uk/~jbialek/Europe_load_flow/

[Boyles and Waller] Boyles, S, Waller, S.T., *A Mean-Variance Model for the Minimum Cost Flow Problem with Stochastic Arc Costs* Paper submitted for review (2010).

[Carpinelli et al. 2001] Carpinelli, G., Celli, G., Pilo, F., Russo, A., *Distributed generation sizing and siting under uncertainty*, conference paper for IEEE Porto Power Tech Conference (2001).

[Crevier 1972] Crevier, D., *Approximate Transmission Network Models for Use in Analysis and Design*, MIT Energy Laboratory Report (1972).

[Croes 2011] Croes, N.N., *Impact of distributed energy generation on energy loss: finding the optimal mix*. MSc Thesis, University of Groningen (2011).

[Deloitte 2007] Deloitte Consulting B.V., *Voorrang duurzaam. Onderzoek naar effecten op investeringsklimaat.*, Final report of consulting project for TenneT and Ministerie van EZ (2007).

[Energy in the Netherlands 2011] *Energie in Nederland*, Energy report by Energie Nederland and Netbeheer Nederland (2011).

[Enslin et al. 2003] Enslin, J.H.R., Knijp, J., Jansen, C.P.J., Bauer, P., *Integrated approach to network stability and wind energy technology for On-shore and Offshore Applications*, 24th International Conference for Power Electronics, Neuremberg (2003).

[Feijoo et al. 1999] Feijoo, A.E., Cidras, J., Dornelas, J.L.G., *Wind speed simulation in wind farms for steady-state security assessment of electrical power systems*, IEEE Transactions on Energy Conversions, vol. 14 no. 4 pp. 1582-1588 (1999).

[Grijalva et al. 2003] Grijalva, S., Sauer, P.W., Weber, J.W., *Enhancement of Linear ATC Calculations by the Incorporation of Reactive Power Flows*, IEEE Transactions on Power Systems, vol. 18 no. 2 (2003).

[Hasche et al. 2006] Hasche, B., Barth, R., Swider, D.J., *Effects of improved wind forecasts on operational costs in the german electricity system*, Research paper for the Institut für Energiewirtschaft und rationelle Energieanwendung, Universität Stuttgart (2006).

[Hydra Project] KNMI Hydra Project: Wind Climate assessment of the Netherlands. Online available: http://www.knmi.nl/klimatologie/ onderzoeksgegevens/potentiele_wind/.

[Jensen 1970] Jensen, D.R., *A generalization of the multivariate Rayleigh distribution*, The Indian journal of statistics, vol. A no. 32 pp. 193-208 (1970).

[Johnson et al. 2007] Johnson, R.A., Wichern, D.W., *Applied Multivariate Statistical Analysis*, sixth edition, Prentice-Hall, Upper Saddle River (2007).

[Kvam et al. 2007] Kvam, P., Vidakovic, B., *Nonparametric Statistics with applications to Science and Engineering*, John Wiley and Sons (2007).

[Le et al. 2007] Le, A.D.T., Kashem, M.A., Negnevitsky, M., Ledwich, G., *Optimal Distributed Generation parameters for reducing losses with economic consideration* Research paper for the Australian Research Council (2007).

[Lee 1979] Lee, L., *Multivariate distributions having Weibull properties*, Journal of multivariate analysis, no. 9 pp. 267-277 (1979).

[Li Yang Chen 2010] Li Yang Chen, R., *Models and Algorithms for Stochastic Network Design and Flow Problems: Applications in Truckload Procurement Auctions and Renewable Energy*, PhD Thesis, University of Michigan (2010).

[Madras 2002] Madras, N., *Lectures on Monte Carlo Methods*, Fields Institute Monographs, American Mathematical Society (2002).

[Nielsen and Zenios 1990] Nielsen, S.S., Zenios, S.A., *A Massively Parallel Algorithm for Nonlinear Stochastic Network Problems*, Decision Sciences Department, University of Pennsylvania (1990).

[Papaefthymiou and Klöckl 2008] Papaefthymiou, G., Klöckl, B., *MCMC for Wind Power Simulation*, IEEE Transactions on Energy Conversion, vol. 23 no. 1 (2008).

[Purchala et al. 2005] Purchala, K., Meeus, L., Van Dommelen, D., Belmans, R., *Usefulness of DC Power Flow for Active Power Flow Analysis*, Proc. IEEE PES General Meeting (2005).

[Reza 2006] Reza, M., *Stability analysis of transmission systems with high penetration of distributed generation*, PhD Thesis, Delft University of Technology (2006).

[The R project 2012] See for all relevant information `http://www.r-project.org/`.

[Quezada et al. 2006] Mendez Quezada, V.H., Abbad, J.R., San Roman, T.G., *Assessment of energy distribution losses for increasing penetration of Distributed Generation*, IEEE Transactions on Power Systems, vol. 21 no. 2 (2006).

[Santoso et al. 2005] Santoso,T., Ahmed, S., Goetschalckx,M., Shapiro, A., *A stochastic programming approach for supply chain network design under uncertainty*, European Journal of Operational Research 167 (2005).

[Schavemaker and Van der Sluis 2008] Schavemaker, P., Van der Sluis, L., *Electrical Power System Essentials*. John Wiley & Sons (2008)

[Shapiro et al. 2009] Shapiro, A., Dentcheva, D., Ruzczynsky, A., *Lectures on stochastic programming: modeling and theory*, SIAM and MPS (2009).

[Srinivasa Rao and Narasimham 2008] Srinivasa Rao, R., Narasimham, S.V.L., *Optimal capacitor placement in a radial distribution system using plant growth simulation algorithm*, World Academy of Science, Engineering and Technology, no. 45 (2008).

[Steinbach 2001] Steinbach, M., *Markowitz revisited: Mean-variance models in financial portfolio analysis*, SIAM Review vol. 43 pp. 31-85 (2001).

[Stott et al. 2009] Stott, B., Jardim, J., Alsac, O., *DC Power Flow Revisited*, IEEE Transactions on Power Systems, vol. 24 no. 3 (2009).

[Toveren met Vermogen 2011] *Toveren met vermogen, Terugblik op tien jaar IOP EMVT*, Report published by Agentschap NL, ministerie van EL&I (2011).

[TenneT 2011] TenneT Holding B.V., *Kwaliteits- en Capaciteitsdocument 2011*, Public report on the operation of the power grid (2011).

[Tseng and Bertsekas 1996] Tseng, P., Bertsekas, D., *An ϵ-relaxation method for separable convex cost generalized network flow problems*, Proceedings of the 5th International IPCO conference, Vancouver (1996).

[Van Slyke and Wets 1969] Van Slyke, R., Wets, R.J.B., *L-shaped linear programs with applications to optimal control and stochastic programming*, SIAM Journal of applied mathematics, no. 17 pp. 638-663 (1969).

[Wallace and Fleten 2003] Wallace, S.W., Fleten, S.E., *Stochastic Programming Models in Energy*, Chapter 10 of A. Ruszczynski and A. Shapiro, Eds., Handbooks in OR & MS, vol. 10 Science B.V. (2003).

[Wang and Nehrir 2004] Wang, C., Hashem Nehrir, M., *Analytical approaches for optimal placement of Distributed Generation sources in Power Systems*, IEEE Transactions on Power Systems, vol. 19 no. 4 (2004).

[Wind Power Program] `http://www.wind-power-program.com/turbine_characteristics.htm` (2012).