MASTER'S THESIS

Optimizing wind farm layout – more bang for the buck using mixed integer linear programming

Patrik Fagerfjäll

Department of Mathematical Sciences
CHALMERS UNIVERSITY OF TECHNOLOGY
GOTHENBURG UNIVERSITY
Göteborg, Sweden 2010

Thesis for the Degree of Master of Science

Optimizing wind farm layout – more bang for the buck using mixed integer linear programming

Patrik Fagerfjäll

CHALMERS

UNIVERSITY OF GOTHENBURG

Department of Mathematical Sciences
Chalmers University of Technology and Gothenburg University
SE-412 96 Göteborg, Sweden

Göteborg, June 2010

Mathematical Sciences
Göteborg 2010

Abstract

Designing a wind farm layout is as of today mainly performed manually. Many different factors affecting the revenue of a wind farm and a large amount of information must be considered, making a close to optimum layout a hard problem to solve manually. In this thesis, mixed integer linear programming models capable of optimizing many common layout problems are developed.

We have formulated two optimization models, the Wind Farm *Production* Optimization model and the Wind Farm *Infrastructure* Optimization model. In the *Production model*, wind turbines are positioned with respect to a minimum separation distance. Furthermore, production losses due to wake effects between wind turbines are accounted for. The model also superpositions sound pressures in nearby areas for different wind directions and is also capable of maximizing profit instead of maximizing the production. The *Infrastructure model* connects the wind turbines by roads and cables, where the latter can include choosing cable dimensions and calculating cable power losses relating to the current. The model can also optimize the positioning of the transformer station.

The two models are capable of optimizing the problems stated at a reasonable computation time. The *Production model* was compared to the commercial *WindPRO 2.6 Optimize* module in verification examples housing 20 to 30 wind turbines. For the problem of maximizing the total production in the given geographic areas, the *Production model* of this thesis managed to find locations for many more wind turbines than *Optimize*, yielding about 40% higher total production. When restricted to allow only as many wind turbines as the software *Optimize* was able to place, the *Production model* still performed equal or better.

Our tests show that the field of mixed integer linear programming and the mathematical models of this thesis possess a great potential to aid in the process of increasing the return of wind farms. These tools are versatile and can be adopted to many different scenarios. They need further development to be able to handle really large wind farm projects but are still, as of today, capable of delivering more bang for the buck in wind farm layout design.

Sammanfattning

The following is an extensive Swedish abstract.

Målet med detta examensarbete är att, givet beräkningsmetoder för produktion, vakeffekter, ljudutbredning och kostnader, skapa matematiska modeller för att optimera total produktion, kostnader eller vinst vid designen av en vindkraftparks layout. Arbetet utvärderar inte olika fysikaliska modeller eller andra optimeringsmetoder, men jämför en av de skapade modellerna med motsvarigheten i den kommersiella programvaran *WindPRO 2.6 Optimize* genom exempelproblem, dock ej genom någon kvantitativ studie.

Layouten hos en vindkraftpark designas idag ofta för hand. Som underlag har man ett projektområde och utifrån en karta över terräng eller beräknad energitäthet skall man placera ut vindturbiner för att nå en hög total produktion. En liten ändring av en vindturbins medelvind ger stora skillnader i effektuttag, och en effektiv inbördes placering är därmed viktigt för att få bra avkastning från parken. Lokala variationer i vindstyrka och vindfrekvens i olika riktningar för olika delar av ett parkområde medför mycket information att hålla reda på. Lägg till detta hur vakeffekterna mellan två vindturbiner påverkar dessas produktion negativt samt hur ljudnivåer från flera vindturbiner samverkar till att öka minsta tillåtna avståndet till bebyggelse, och processen blir väldigt detaljrik. Utöver dessa produktionsaspekter innefattar designen också problem som att välja rätt dimension av kablar samt att dra dessa och interna vägar på billigast möjliga sätt.

Två modeller skapas, av typen linjära heltalsprogram, där den ena representerar problemet att maximera produktion och den andra problemet att minimera kostnader för kablar, vägnät och placering av transformatorstation. Båda modellerna kan beskrivas med hjälp av grafer, vilka består av en mängd geografiskt utspridda noder samt bågar som sammanbinder dessa. Noderna representerar möjliga placeringar av vindturbiner, där en nod i vilken en vindturbin placerats kallas aktiv, och bågarna symboliserar vakeffekter, kablar eller vägar mellan noder. En aktiv båge i *Infrastruktursmodellen* kan exempelvis motsvara en kabeldragning mellan två vindturbiner.

I *Produktionsmodellen* implementeras ett minsta tillåtna avstånd mellan två vindturbiner som ett tätpackningsproblem av cirklar vars radier är halva detta minsta avstånd. Två placerade vindturbiner medför produktionsförluster i form av vakeffekter vilket representeras av en aktiv båge mellan motsvarande noder i grafen. Den kombinerade effekten av vakeffekter från flera källor beräknas genom linjär överlagring. Då andra kombinationsmetoder tidigare har visat sig stämma bättre överens med verkligheten, exempelvis att summera kvadraterna av hastighetsförlusterna, skulle detta kunna vålla problem för Produktionsmodellen, och en metodik för att i dessa modeller approximera andra samband än linjär överlagring presenteras. Vidare kompletteras modellen med bivillkor som begränsar den kombinerade ljudnivån i en viss punkt från alla vindturbiner och för olika vindriktningar. Detta innebär att det minsta avståndet från en vindturbin till en närliggande bostad beror av antalet vindturbiner, i vilka riktningar från bostaden de är placerade samt hur vindkarakteristiken för dessa olika riktningar ser ut. Denna modellering medför större möjligheter än om man endast anger ett minsta avstånd till bebyggelse, vilket görs i *Optimize*.

I *Infrastrukturmodellen* består bågarna i grafen av möjliga vägval för kabel- och vägdragning. Här skapas en trädstruktur av effektflöden från vindturbinerna till den gemensamma transformatorstationen. Bågarna tillåts även att ansluta till icke aktiva noder för att möjliggöra förgreningar utanför vindturbinernas positioner. Slutna slingor i trädet, vilket är ett problem som måste hanteras vid trädoptimering, undviks då detta effektflöde endast kan skapas i vindturbinerna och allt flöde måste nå transformatorn. Från transformatorn skickas effektflödet i den dyrare exportkabeln till valfri anslutningspunkt. Modellen kan behandla kabeltyper med olika kapacitet samt modellera dubbla kablar och de effektförluster ström genom en kabel ger upphov till. De olika typerna modelleras genom att varje båge i grafen kan representeras av flera matematiska variabler med olika övre gräns för mängden effektflöde och olika kostnader. Utöver kostnaderna för kablar och vägar ingår i *Infrastrukturmodellen* terräng- och platsberoende kostnader för fundament till transformatorstation samt för kabelförgreningar i de fall dessa sker i kopplingsstationer istället för vid befintlig vindturbin.

De båda modellerna kan genom att prissätta produktionen kombineras till en gemensam modell med målet att maximera avkastningen, vilket typiskt leder till att enskilda vindturbiner inte byggs för långt ifrån de resterande eller på en plats dit vägbygge eller kabeldragning blir för dyrt.

Utöver de implementerade modellerna föreslås tillägg för att öka hastigheten med vilken *Produktionsmodellen* kan lösa problem med hög upplösning. Denna höga upplösning har mindre betydelse för storleken på vakeffekterna än för möjligheten till precis och tät placering av vindturbinerna. För att antalet bågvariabler i grafen därför inte skall växa som kvadraten på antalet noder, föreslås en uppdelning av ytan i sektorer. Vindturbiner kan placeras noggrant över det högupplösta rutnätet, medan vakeffekter samt eventuellt kablar och vägar approximeras med bågar mellan de sektorer som innehåller vindturbiner. Antalet bågar motsvarar då antalet sektorer i kvadrat oberoende av upplösningen. Storleken på dessa sektorer väljs så att det minsta tillåtna avståndet mellan vindturbiner förbjuder fler än en vindturbin per sektor samt så att vakeffekterna mellan sektorer förblir en god approximation av vakeffekterna mellan de riktiga positionerna för vindturbinerna.

De utvecklade modellerna visar sig fungera väl för sina respektive ändamål och kan användas i processen att skapa en layout för en vindkraftpark. *Produktionsmodellen* kan optimera vindturbinernas placeringar med avseende på antingen produktion eller vinst, och kan därmed bidra till en högre total produktion och större intäkter. *Infrastrukturmodellen* hittar prisvärda lösningar på kabeldragningar med val av kabeltyp samt vägnät, vilket kan bidra till att minska kostnaderna för dessa. Då båda modellerna kan hantera mycket av den ingående problematiken kan också en större del av layoutprocessen automatiseras och därmed tiden för det manuella arbetet minskas. Kombinationen av de båda modellerna till en enda optimeringsmodell är intressant ifall kostnaderna för kabel- och vägnät är av betydande storlek relativt vinsten från vindturbinernas produktion, då förbättringar hos dessa annars endast leder till obetydliga förbättringar i totala vinsten och vindturbinernas position sannolikt inte påverkas. Relationen mellan produktionsvinsten och kostnaden för kabel- och vägnät beror på elpris samt vilken önskad återbetalningstid för parken som används. Den kombinerade modellen blir dock stor och tar lång tid att lösa och vissa föreslagna genvägar för att minska antalet noder i infrastrukturmodellen inte kan användas.

För verifiering av *Produktionsmodellen* i ett riktigt projektområde jämfördes lösningarna funna efter en timme för två upplösningar med lösningarna funna genom den kommersiella optimeringsprogramvaran *WindPRO 2.6 Optimize*. Som facit på de placerade vindturbinernas produktion genomfördes produktionsberäkningar för de framtagna layouterna med hjälp av verktyget *WindPRO 2.6 PARK*, vilket är ett verktyg för att beräkna total produktion för en viss layout. Båda metoderna fick samma indata, använde samma modeller för produktion- och vakberäkningar och optimerade med målet att maximera total produktion. Kombinationen av multipla vakeffekter hanteras dock på olika sätt: *Produktionsmodellen* som överlagrar enskilda vakeffekter överskattar den kombinerade förlusten från två vakar jämfört med *Optimize* och *PARK* som båda kombinerar vakar som summan av kvadraterna av hastighetsförlusterna. Då *Optimize* inte kan hantera överlagring av ljudintensitet från flera vindturbiner, utan endast använder ett minsta tillåtna avstånd från var och en av dessa, skalades *Produktionsmodellen* ner till att inte inkludera denna del. Jämförelsen visade att *Produktionsmodellen* finner en högre total produktion än *Optimize*. Detta då den kunde placera ut många fler vindturbiner än vad *Optimize* lyckades med. I fallet med en upplösning om 100m lyckades *Optimize* placera ut endast 18 vindturbiner medan motsvarande antal för *Produktionsmodellen* var 28 stycken, vilket innebar en 43% högre total produktion. För upplösningen 50m var motsvarande antal vindturbiner 20 respektive 30 stycken, vilket gav en 38% högre total produktion. De av *Produktionsmodellen* utplacerade antal vindturbiner motsvarar också de maximala antalen för respektive upplösning. Om antalet vindturbiner begränsades i *Produktionsmodellen* till de 18 respektive 20 stycken som *Optimize* lyckades placera ut gav denna en total produktion som var ungefär en procent högre än *Optimize* i problemet med upplösningen 100m och ungefär lika i problemet med upplösning 50m. I det senare problemet begränsades *Produktionsmodellen* av programvarans tilldelade minne till att inte kunna inkludera vakeffekter på längre avstånd än 900m. Trots att *Optimize* och facit *PARK* kombinerar multipla vakeffekter annorlunda än *Produktionsmodellen* gav denna alltså lösningar med lika bra eller bättre total produktion.

De utvecklade modellerna kan hitta värdefulla förbättringar i många av de delmoment som ingår i en designprocess, och kan dessutom spara in på den tid det skulle tagit att skapa dessa för hand. De kan hantera projekt på land och offshore, med fri positionering eller bunden av visuella aspekter eller specifika möjliga positioner. Produktionsmodellen kan hantera avancerade ljudutbredningsmodeller som beror av vindriktningar. *Infrastrukturmodellen* har också kompletterats med tillägg som möjliggör att optimera sammankoppling av vindkraftparker istället för enskilda vindturbiner, vilket möjliggör användande av denna modell i ett vidare sammanhang. Programmen behöver kompletteras och testas, dels med föreslagna förenklingar så som sektoruppdelning av vakvariabler, dels med förfinade optimeringsparametrar, såsom bättre val av branching strategy, för att klara av att lösa riktigt storskaliga problem på rimlig beräkningstid. För parker på ett antal kvadratkilometer med några tiotals vindkraftverk går problemen dock redan nu att optimera till tillräcklig nivå på någon timme på en persondator. Linjär heltalsprogrammering visar sig vara en fungerande teknik för att optimera dessa layouter och ytterligare utveckling bör utföras inom området. Modellerna kan redan idag, och kommer än bättre i framtiden kunna, skapa layouter som medför en bättre avkastning från vindkraftparker.

Contents

Acknowledgements

This master's thesis was performed in collaboration with Vattenfall Power Consultant at their facilities. Firstly I want to thank the two people at this company who initially trusted in me and granted me the possibility to undertake an optimization work of my choice – my supervisor Lasse Johansson and my boss Mattias Törnkvist. I am also grateful for them allowing me to take part in all various types of social and professional events.

I wish to thank all of my co-workers at the Wind unit of Vattenfall Power Consultant, who have supported my work with ideas and with answers to numerous questions, as well as provided me with a very likable working environment. Out of these, I especially wish to extend my gratitude towards my two supervisors, Lasse Johansson and Anna Larsson, for their aid and expertise.

For the mathematical input and discussions, as well for the extensive reading and improvement of this theses, I would like to thank my supervisor and examiner, Docent Ann-Brith Strömberg at the department of Mathematical Sciences at Chalmers University of Technology.

Göteborg, May 2010

Patrik Fagerfjäll

Chapter 1

Introduction

1.1 Structure of the thesis

This thesis describes the work of developing models for addressing the optimization of production and cost of the wind farm layout optimization problem. It relies on existing methods and models for calculating wind resources in a geographic area, production, wakes, sound propagation and so on, and is a method for combining these models to optimize the locations of wind turbines and the road and cable networks. The thesis does not evaluate physical models or other means of optimization. It will through verification examples compare one of the models developed to the corresponding module in the commercial software *WindPRO*, however, no extensive quantitative study is undertaken. The aim of the thesis is to develop models and to conclude *whether* it is possible to increase production and profits, rather than to indicate by *how much* these may be increased.

The thesis starts by listing many important subproblems of the wind farm layout design problem. It then gives an introduction to the modeling tool used in this thesis, *mixed integer linear programming.* Following these initial chapters, the two models - the *Wind Farm Layout Production Optimization model* and the *Wind Farm Layout Infrastructure Optimization model* - capable of solving the problems stated are built, some performance results are listed and the *Production model* is verified and compared with commercial software in a real life example. The application of the models in a real layout processes is the main focus of the *Discussion* section.

1.2 The need for wind farm layout optimization

As of today much of the wind farm layout process is done manually, partly since there does not seem to exist any good optimization tools (one commercial optimization tool will be presented in this thesis). Care should be taken so that the iterative work flow of a design process is not restricted too much. However, due to the enormous amount of information that must be handled when searching for the best wind turbine locations, there is a lot of room for computer aided optimization. Although the uncertainties in the models used for modeling, for example, wind flows and the spread of a wake, need not be small [22], this is the information that financing decisions are currently based upon. Thus, increasing the modeled total profit is of great value. Since this profit is highly dependent on the power production, and thus the positioning of the wind turbines, an effort creating tools that yield a higher production, even if merely a percent higher, would be well worth.

1.3 From a mathematical point of view

The wind farm layout optimization problem consists of several theoretically interesting subproblems. Circles are to be packed tightly in an area and nodes in a graph are to be connected by a spanning tree with the possibility but not the obligation of using additional nodes to shorten the total length of the tree, which transforms the spanning tree problem into a so called Steiner tree problem. Optimization in wind farm layout design is a young field of application and while heuristic or genetic algorithms seem to have been the first solution methods chosen [7, 19, 17, 12, 23] it is important to investigate the performance of an integer linear programming approach.

Chapter 2

Wind farm layout

Designing a wind farm layout is a process involving many branches of expertise, including wind analysis, construction, power grid dimensioning, environmental impact analysis and in some sense politics. This chapter will list a few of the considerations important in the layout process, which will all be incorporated in the model created. They are not exclusive, but represent most of the important subproblems of wind farm layout design.

2.1 Maximizing production in a given area

The first and major part of the wind farm layout optimization problem is to find the optimum wind turbine locations. This section will define important factors affecting permitted locations and wind farm production.

2.1.1 Project site area

Many aspects are taken into account when finding and deciding on project site area. These aspects include finding a windy and accessible piece of land, not covered by, for example, wetlands or any type of nature reserve. Moreover, there could be considerations regarding the site area posing as a habitat to endangered species or being the nesting grounds for birds. Restrictions regarding minimum distance from, for example, existing roads could also apply and should be accounted for when defining the project area. In this thesis it is assumed this area is defined with restricted areas removed.

2.1.2 Wind distribution

The power in a wind is proportional to the cubic wind speed, v^3, where v^2 comes from the kinetic energy of the air mass and v from the air mass flow (Figure 2.1). Thus, a small change in wind speed can have great impact on produced electrical output. Expressed relative to the wind speed, a one percent increase of the wind speed v yields a $(1.01)^3 - 1 \approx 3\%$ gain in v^3 and, thus, production. A good site location and proper wind turbine placement within the wind farm are therefore of importance.

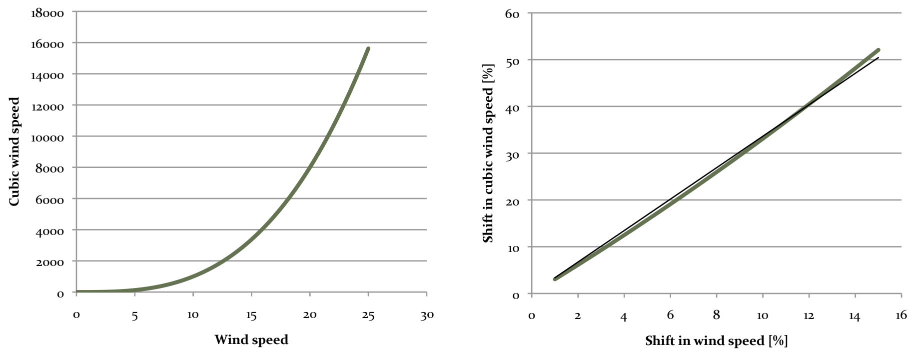


Figure 2.1: **Power in wind proportional the cube of the wind speed.** The right graph describes the percentage increase in cubic wind speed per percentage increase in wind speed, $(\Delta v, (v + \Delta v)^3/v^3)$. The fitted linear line has an incline of about 3.5.

Wind speed distribution and directions are calculated for all of the site area. This calculation is based on wind measurements made over a longer time, preferably from a hub height mast at the project site, which are corrected for deviance from yearly averages. Photos showing two means of measuring wind characteristics, a mast and a sodar, are shown in Figure 2.2. Softwares used in the modeling could be based on computational fluid dynamics (CFD) or some terrain flow model, and incorporate, for example, elevation, roughness of the ground surface for a large area surrounding the site [18, 8]. The output of this software and the input to the thesis models will be Weibull distribution parameters in a number of angular sectors for all points in a grid covering the project area and surroundings. The Weibull distribution has been shown to represent wind speed frequencies well and is commonly used [24]. A Weibull distribution along with an example of a power curve and the resulting production distribution is shown in Figure 2.3. An illustrative representation of the Weibull parameters and wind direction frequency for a number of sectors is the wind rose, seen in Figure 2.4.

2.1.3 Minimum distance between wind turbines

Wind turbines cannot be placed too close to each other. The most obvious distance is two rotor radii apart to prevent the rotor blades from colliding. There are however usually other requirements or demands on minimum

(a) Mast (b) Sodar

Figure 2.2: **Means for measuring wind speed.** Picture (a) shows an 80m high mast fitted with anemometers whereas in picture (b) a sodar, measuring wind speed at different heights by observing mainly the doppler shift of returning aucustic impulses, is seen. Both pictures are taken in Sourva, Sweden.

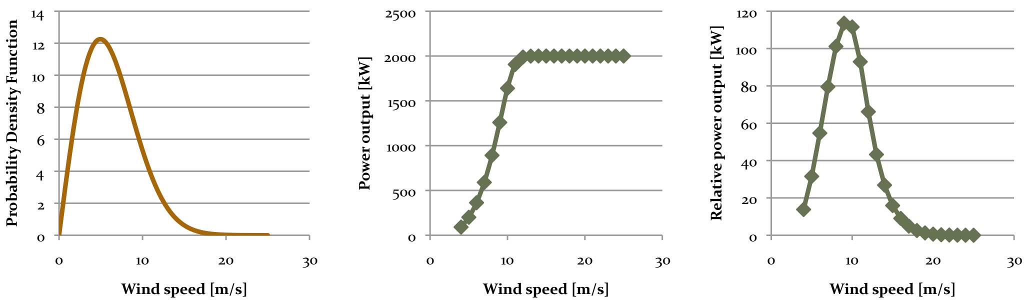


Figure 2.3: **Weibull distribution, a power curve example and the resulting production distribution.** The latter is the product of the Weibull distribution and the power curve at each wind speed.

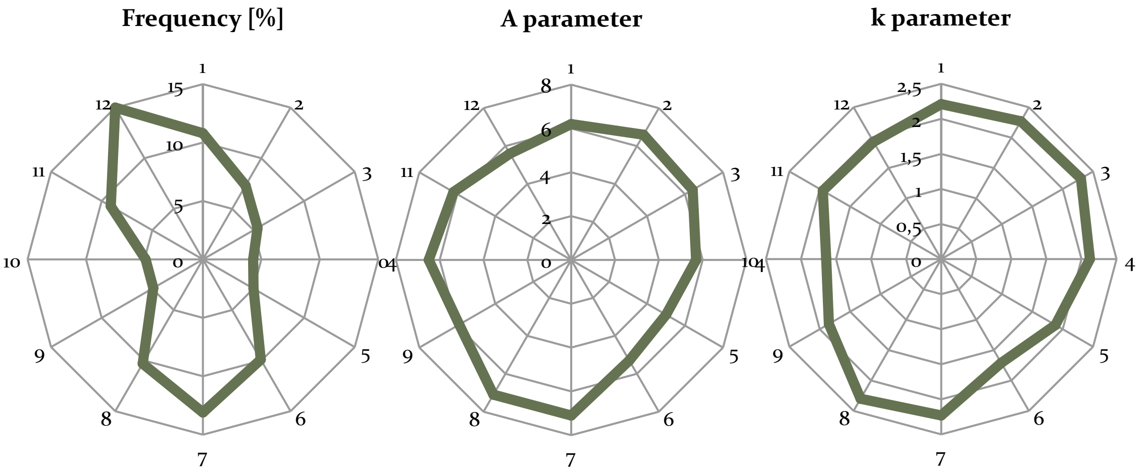


Figure 2.4: **Wind roses,** showing frequency and Weibull A and k parameters for twelve 30 degree sectors, where north is denoted by 1.

separation distance to be met made from manufacturers, partly to prevent excessive tear on the construction due to turbulent winds [15, 24]. A typical minimum distance could be around four or five rotor diameters but would sometimes be set higher when manually placing the wind turbines due to production losses between them, further discussed in the next section.

2.1.4 Production loss between wind turbines

Wind passing through a wind turbine will decrease in speed, not to recover fully for quite some distance. Thus, two wind turbines standing not too far apart will have a negative effect on each others production, referred to as wake effects or array losses. This loss of wind speed could be calculated through a number of models and the optimal positioning of wind turbines will depend on this effect. These wake models are often based on single turbine wakes, and the method for combining several wakes varies. One common way is to sum the kinetic energy deficits, that is the sum of the square of the velocity deficits. This is said to in some cases have a better resemblance with reality than just linear superposition, although not generally accepted why [14, 22, 9]. This thesis will however, due to the linear programming approach, make use of linear superposition and the differences between these methods will be addressed later on. For the purpose of manually positioning wind turbines a distance between two turbines of 5 rotor diameters perpendicular to, and 7 rotor diameters parallel to the prominent direction of wind is often proposed [24].

2.1.5 Sound levels in surrounding areas

There are regulations stating maximum sound levels originating from wind turbines at nearby houses or other sensitive areas. Through the use of models calculating sound levels around a given source, this relates to a minimum possible distance between wind turbines and these areas. The regulations could be formulated in a number of ways, as, for example, "must not exceed an equivalent sound level of 40 dB(A) at nearby houses, and superposition of sound pressure from multiple wind turbines is made as if the sensitive area is downwind of all wind turbines simultaneously", partly taken from the Swedish *Naturvårdsverket* [1]. The minimum distance will however vary depending on the number of wind turbines positioned. Two wind turbines of the same sort in about the same spot will yield double the sound pressure, equivalent to a 3dB sound level rise, in a nearby point. This effectively increases the minimum distance. Sound pressure and sound level relates as $\mathrm{SL} = 10\log_{10}(\mathrm{SP})$.

2.1.6 Choice of wind turbine

The choice of wind turbines will affect the production in a location, the magnitude of the wake effects and the sound levels. In many wind turbines there is also an option to decrease the latter at the cost of also decreasing power production.

2.2 Building infrastructure and reducing costs

It is not only the available amount of wind energy in each point, and thus the possible power output, that will decide the optimal wind turbine locations and return on investment. The cost of constructing a wind farm could vary depending on the amount of work and material needed for foundations as well as roads and cables connecting the wind turbines, transformer station and the external power grid. The problem of optimizing this is in contrast to the above aspects of maximizing production a problem of minimizing costs. Relative to the cost of wind turbines the costs of roads and cables are rather small and it is common these costs will not be considered until after the wind turbine locations have been chosen [24]. If locations are already decided, a model optimizing the costs could be formulated and used straight away. If, however, these costs are to be considered when finding the optimal wind turbine locations, other approaches are needed since the two problems are optimizing different quantities. One could optimize over both production and costs by letting one of them be fixed at a certain level, and do this over and over again using different levels to create a type of trade-off curve, in economics often referred to as a Pareto efficient frontier. Another approach is to bridge the two problems by defining a revenue on electricity production, thus optimizing revenue minus costs over a given time period. This latter approach will be implemented for the purpose of testing problem complexity.

2.2.1 Foundations

Wind turbines should not be built where the terrain, for example, is too steep, too rough or consists of too unstable ground for it to be economically beneficial or even possible to use as a location. To deal with this, the cost of foundation must be priced correctly in different areas. One should also add the cost of each wind turbine, though this cost will generally not vary due to the choice of location. If using multiple types of wind turbines, the costs of these and possibly foundation costs could vary depending on which type is chosen. Figure 2.5 shows pictures from the construction of a small wind farm.

2.2.2 Cables

Wind turbines need to be connected to a transformer station and the existing external grid, and for this interconnecting cables buried underground are generally used. Since the act of transporting and burying a cable requires a road to be constructed along its path, roads and cables are usually constructed taking the same paths.

(a) Nacelle (b) Construction

Figure 2.5: **The construction of a wind turbine,** with hub height 105m. Both pictures are taken in Simmatorp, Sweden.

Interconnecting cables of different dimensions can allow for different current levels and thus different number of connected wind turbines. Several cables can also be placed along side one another requiring only one road to be built and affecting the amount of power losses. The model should thus be able to incorporate different types of cable on each path. The cost of a cable will depend on both the work needed for constructing the road and the burying as well as the type of cable selected. For connecting the transformer station to the external grid, an export cable with high capacity and cost is used. In addition to the construction costs there is the cost of power losses which is proportional to the length of a cable and the square of the current.

Several cables originating from wind turbines could be combined into high capacity cables, as in the branches of a tree, with additional costs for connector stations if this connection is not performed at another wind turbine. If defining a power flow from the wind turbines to the transformer station, flow can be combined but not split.

2.2.3 Roads

Roads connecting the wind turbines and the outside world are needed for transporting the wind turbines to their locations and performing maintenance, where the former is the need defining what roads can be built or not. Wind turbine transports are heavy, with the generator as the heaviest part, or really long, as when transporting the blades (as seen in Figure 2.6). This calls for robust, low-incline roads, which, for example, could result in it not being possible to build on top of a hill even though wind energy levels might be high and foundation costs low.

Figure 2.6: **Wind turbine blade transport.** Each blade is 45m long. The picture is taken in Simmatorp, Sweden.

2.2.4 Transformer station

A large wind farm would usually connect to the external power grid through a transformer station which, just as in placing a wind turbine, will result in foundation as well as road and cable costs along with the cost of the transformer station itself. Its optimal placement is dependent on the cost of the high voltage export cable needed for connecting to the external power grid compared to the cables needed for connecting the wind turbines. The construction cost of the transformer station can also vary depending on its location within the farm. Power losses in the transformer will depend on the current, and thus the number of wind turbines, but not on the location of the wind turbines within the farm. It is possible to allow for different sizes of transformer stations if the number of wind turbines to be placed is not decided prior to optimization.

Chapter 3

About integer linear programming - the mathematical modeling tool

This section gives a short introduction to the field of mixed integer linear programming used in developing the models of the thesis. Examples and basic theory of linear programs, integer linear programs and the definition of graphs and trees are presented, along with a final section defining the notations used in this thesis.

3.1 Linear programming

The word *linear* in linear programming (LP) refers to the variables of a problem being combined in a linear way, that is adding or subtracting variables but not, for example, multiplying them or using them as exponents. Most physical problems, or other types of problems for that matter, cannot be expressed in a linear way, but those who can can be solved in more efficient ways. The concept of linear optimization will here be explained by the use of an example.

Let us say we have discovered that the citizens of New York have great cravings for Swedish *Knäckebröd* and *Messmör*, a form of crispbread and soft whey butter. We see a great business opportunity and decide to bring a cart and a bag full of these goods across the Atlantic to pursue a career as Knäckebröd and Messmör vendor. On the flight to America the cart will take up all of our checked baggage allowance, leaving us with just the carry-on baggage for transporting our goods. If the economic and physical properties of Knäckebröd and Messmör are according to Table 3.1 and the carry-on baggage is restricted to 8 kg and 60 litres respectively, what quantities of the two products should we bring to maximize our profit in the land of opportunities? For the moment, fractions of packages are allowed to be bought and sold and we are able to pack the goods real tight in our baggage.

[per package]	Knäckebröd	Messmör
Cost [SEK]	20	35
Revenue [SEK]	40	60
Profit [SEK]	20	25
Weight [kg]	0.5	0.8
Volume [litre]	7.5	0.6

Table 3.1: **The properties per package of Knäckebröd and Messmör**

The total potential profit, which is what we want to maximize, can be expressed as the linear function of packages of Knäckebröd, K, and Messmör, M, called the *objective function*, to

$$\text{maximize} \quad 20K \; + \; 25M.$$

Since this objective function is ever increasing, the *constraints* of maximum weight and volume of our combined goods have to be added. This is done by introducing two more linear functions, and the *linear program* reads

$$\begin{array}{lrcrcl} \text{maximize} & 20K & + & 25M, & & \\ \text{subject to} & 0.5K & + & 0.8M & \leq & 8, \\ & 7.5K & + & 0.6M & \leq & 60. \end{array}$$

Now, since we should not be able to buy or export a negative number of goods, neither of the variables K and M are allowed to obtain a value less than zero (another effect of allowing negative variable values would be the possibility to sell, for example, Messmör before entering the aircraft, yielding a negative impact on the total weight and volume, thus, making it possible export more than 8 kg and 60 litre of Knäckebröd). The added constraints now yields the linear program to

$$\begin{array}{lrcrcl} \text{maximize} & 20K & + & 25M, & & \\ \text{subject to} & 0.5K & + & 0.8M & \leq & 8, \\ & 7.5K & + & 0.6M & \leq & 60, \\ & & & K,\, M & \geq & 0. \end{array} \tag{3.1}$$

The possible combinations of values, the set of *feasible* solutions, of K and M not violating the constraints can now be illustrated by the bounded area of Figure 3.1. All other combination of variables are referred to as *infeasible.* Both the variables need to be non-negative and are restricted by the two lines representing the weight and volume constraints. This type of area, bounded by straight lines, is a called a *polyhedron* and is in this instance *convex.* Convexity basically means, when referring to an object in an Euclidean space, that there are no holes or dents in the object - the straight line segment joining any pair of points within the object, lies entirely within the object. Mathematically speaking, a set U is convex if for any $u_1, ..., u_n \in U$, and any

$\lambda_1, ..., \lambda_n \geq 0$, such that $\sum_{k=1}^{n} \lambda_k = 1$, the vector $\sum_{k=1}^{n} \lambda_k u_k \in U$. Linear functions on such convex polyhedrons, or in several dimensions convex polyhedrons, have the property of attaining their maximum and minimum values in an extreme point of the polygon or along one of its sides (that is, along the feasible region of a constraint line) [16]. In Figure 3.1 these possible extreme points are marked by circles, and the act of optimizing the problem 3.1 is now a matter of locating the right extreme point. In this two-dimensional problem, this can be seen by just observing the figure while knowing which direction is the gradient of increasing profit. For a problem consisting of many more variables, optimization through such a visual approach is however not possible and mathematical methods are required.

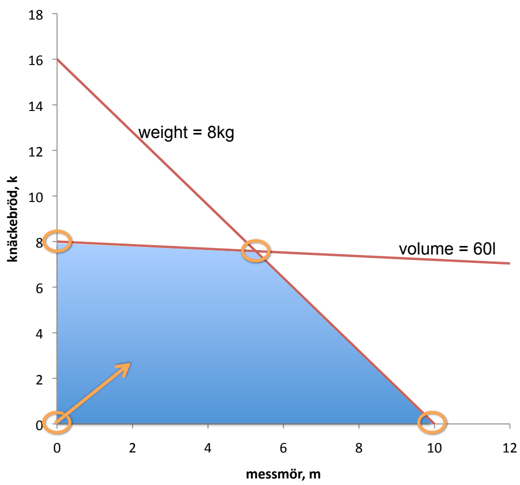


Figure 3.1: **A convex polyhedron of feasible solutions.** Due to the linearity of the objective and constraints, and the set of feasible solutions being a convex polyhedron, the optimum will be in one of the extreme points or along a constraint line. The gradient of increasing profit is illustrated by an arrow.

The optimal solution to the problem of exporting Knäckebröd and Messmör is shown in Figure 3.2. In this particular example, at this intersection of constraint lines, both the weight and the volume of the baggage are utilized to a maximum.

Further readings on linear programming can be found in [16].

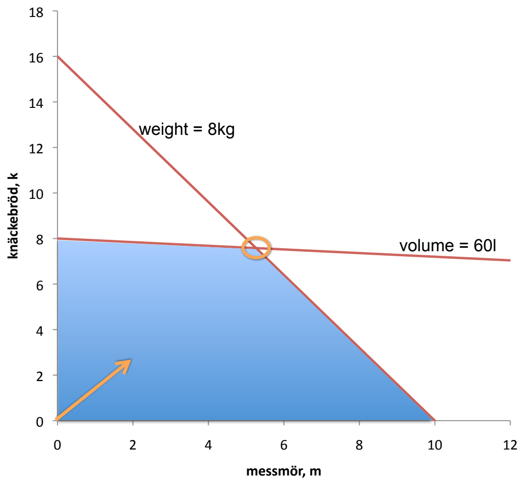


Figure 3.2: **Optimal solution to the problem of selling Knäckebröd and Messmör.** The optimal solution is $K \approx 7.6$ and $M \approx 5.3$.

3.2 Integer linear programming

An *integer linear program (ILP)* is formulated in the same manner as a linear program with the additional restriction of the variables having to be integer valued instead of continuous. If both integer and continuous variables are used, the program is a *mixed integer linear program (MILP)*. Though still expressed in a linear way, solving a integer program is generally a more time consuming task than solving its continuous counterpart. Moreover, changing a problem from continuous to integer restricted means adding restrictions on the variables, and additional restrictions can never increase the value of the optimal solution but will decrease it or perhaps not change it at all.

To get a hint of why integer problems take a long time to solve, one could compare to the complete enumeration of all possible combinations. If having a set of n binary variables, $x_i \in \{0, 1\}$, where $i = 1, ..., n$, the number of possible ways to combine these variable values is 2^n. With a computer capable of computing and checking feasibility of one billion, 10^9, of these combinations per second, the time needed for different number of variables is shown in Table 3.2. If the problem is that of connecting n wind turbines by roads, similar to the formulation derived in Section 4.3.2, the number of integer variables correspond to $n + n^2$, and the complete enumeration time, using the same computer as above, is also seen in the table.

n	time, 2^n
20	0.001 seconds
50	313 hours
80	38 years

n	time, 2^{n+n^2}
5	1 second
6	1.2 hours
7	2.3 years
8	150 thousand years

Table 3.2: **Computational time.** Time required for complete enumeration on a computer capable of 10^9 computations or feasibility checks per second.

Clearly, other means of methods than complete enumeration are required. One commonly used algorithm, which will be used in this thesis, is the branch and bound algorithm. This algorithm *branches* on the fractional variables found as optimal solutions of the linear program (LP) relaxation of an integer problem (say $x^\star_{LP} = 5.4$ being the optimal solution of a LP-relaxed problem where one of the integer variables is denoted by x), one at a time, splitting the problem into two continuous subproblems where the variable is forced to be no larger than the fractional value rounded downwards to the closest integer ($x \leq 5$) and no less than this value plus one ($x \geq 6$), respectively. Each of these problems is a *LP-relaxation*, meaning that the integer requirements are relaxed. Fractional variables in the solutions of the two new problems are branched on, creating two new subproblems per subproblem, and the process is repeated. These iterations continue on a branch until an integer solution is found, in which there are no fractional variables left to split in that branch, or the bounds on the variables renders the problem infeasible. However, since the solution to every subproblem will be an *upper bound* on the objective value of that branch (remembering this value cannot increase by imposing more restrictions), the algorithm will also stop searching a branch if this upper bound is less than the objective of an already found integer solution. This procedure is called *pruning* the solution tree. Throughout the branch and bound process, better integer solutions will be found and the upper LP-relaxation bound will be decreased. The relative difference between these is a measure of the difference between the value of the best integer solution and that of the optimal solution. It is thus a quality measure of the best feasible solution found. It is here referred to as the MIP-gap, defined for a maximization problem as

$$\text{MIP-gap} = \frac{\text{Best LP-relaxation objective value found so far}}{\text{Best integer objective value found so far}} - 1.$$

The principle of the branch and bound method is illustrated in Figure 3.4 for the integer problem example of exporting *Knäckebröd* and *Messmör*, and is described below.

Continuing the example problem 3.1, we may sell fractions of packages, but when buying our goods in Sweden we will have to buy full packages. This translates into requiring the variables representing Knäckebröd and Messmör being integer, that is $K\ M \in \mathbb{N}$, where $\mathbb{N} = 0,\ 1,\ 2\ldots$ which is altered in the mathematical formulation to

$$\begin{array}{lrcrcl} \text{maximize} & 20K & + & 25M, & & \\ \text{subject to} & 0.5K & + & 0.8M & \leq & 8, \\ & 7.5K & + & 0.6M & \leq & 60, \\ & & & K,\ M & \in & \mathbb{N}. \end{array} \tag{3.2}$$

The graphical representation of the set of feasible solutions to this integer problem is shown in Figure 3.3. For example, rounding the obtained continuous optimum to the closest integer feasible point is not a way of optimizing the integer problem.

Now, to solve this problem the branch and bound algorithm is used. The procedure is hereby described and is illustrated in Figure 3.4. We start off with the optimal solution to problem 3.1, $(K,\ M) = (7.6,\ 5.3)$. Now, we choose a fractional variable, in this case K, and branch in to the two continuous subproblems with the additional constraints $K \leq 7$ and $K \geq 8$, respectively. When solving these continuous problems we notice an integer solution, $(8, 0) = 160$, found in the "$K \geq 8$"-branch. This is a solution to the integer program 3.2, but we do not know whether it is optimal. The MIP-gap tells that the value of the real optimum cannot be more than 75% better than this integer solution. This branch can now be pruned and the objective value of 160 is used as a lower bound on the optimal objective value. If another branch yields an upper bound (a LP-relaxation value) lower than this, that branch can be pruned. Continue by branching on the fractional value of M in the leftmost node in level 1, yielding an even better integer solution, $(7, 5) = 265$, and another fractional solution, respectively. From

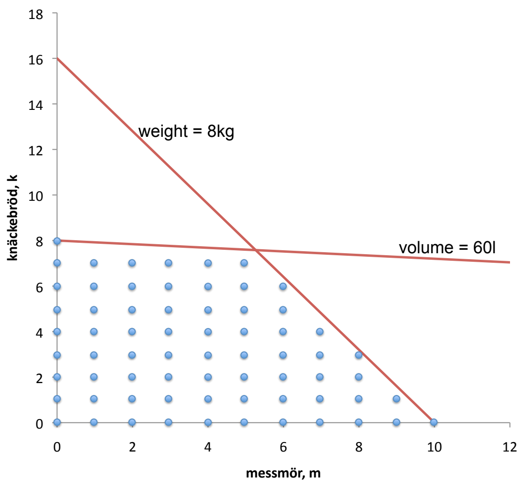


Figure 3.3: **The feasible set of solutions to the integer problem.** The dots all represent possible solutions.

this node we branch into $K \leq 6$ and $K \geq 7$, but due to the earlier constraint of $K \leq 7$ the only option for the rightmost branch is $K = 7$. Now, the configuration of $M \geq 6$ and $K = 7$ is infeasible and searching along that branch is stopped. We carry on, find an even better integer solution, (6,6), with an objective value 270, but still does not know whether it is optimal. The next split however yields one infeasible subproblem and one for which the upper bound is less than the objective of the best integer solution found, and we have proven the solution $(K, M) = (6, 6)$ to be optimal.

The feasible set of solutions to this problem is once again, along with this computed optimum, illustrated in Figure 3.5.

3.3 Graphs and trees

A *graph*, $G = (V, E)$, is a set of *nodes* or *vertices*, V, linked in some way by directed and/or undirected*edges*, E. An example of a graph is a set of cities with air traffic routes linking them together either directly or via another city in the set (but without requiring the existence of paths between any pair of nodes). By representing the nodes and edges with variables (for example, $x_i \in \{0, 1\}$, $i \in V$, and $y_{ij} \in \{0, 1\}$, $(i, j) \in E$, respectively (see

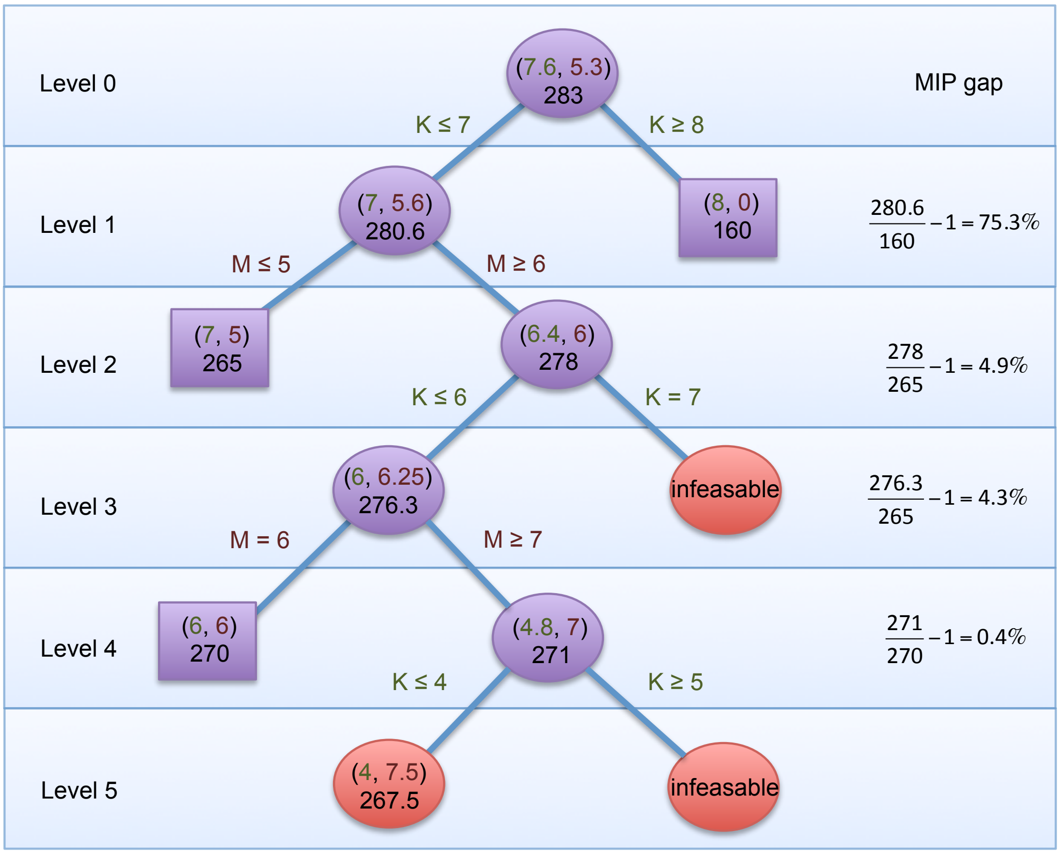


Figure 3.4: **Branch and bound solution tree,** for the problem of exporting *Knäckebröd*, K, and *Messmör*, M.

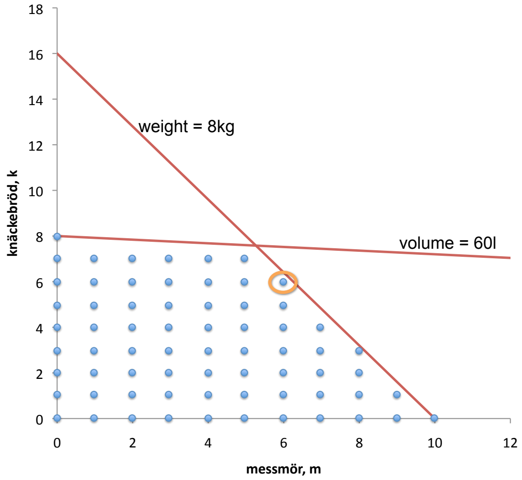


Figure 3.5: **Optimal solution to the problem of selling Knäckebröd and Messmör.** The optimal solution is $K = 6$ and $M = 6$.

illustration in Figure 3.6)), these roads connecting cities can be modeled by the use of constraints in a linear program. It is thus, for example, possible to minimize the total sum of the air route lengths (minimize $\sum_{(i,j)\in E} d_{ij}y_{ij}$, where d_{ij} is the length of the route from node i to j), while requiring each visited city to be connected by both incoming and outgoing routes (subject to $\sum_i y_{ij} \geq x_j,\ j \in V$ and $\sum_j y_{ij} \geq x_i,\ i \in V$).

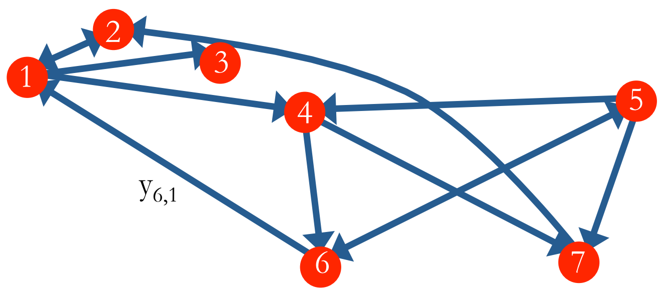


Figure 3.6: **Graph example.** Consisting of geographically distributed nodes with directed edges connecting pairs of nodes.

There are numerous optimization problems formulated using graphs, as, for example, the traveling salesperson problem, where one needs to find the shortest path for a salesperson visiting every city in a set but never visiting the same city twice, or as similar to the problems of this thesis, to find low cost tree structures spanning the nodes. A *minimum spanning tree*, for example, is to find the shortest, or cheapest, combination of edges that connect each node to at least one of the other nodes, creating a structure of edges similar to the branches of a tree. For many of the classic graph theory problems, heuristic solution algorithms have been developed. However, when altering such a problem or creating a somewhat different type of graph problem, as in this thesis, these heuristic algorithms need not be applicable. Finding efficient algorithms producing at least close to optimal solutions using heuristics developed for a new problem could prove to be hard.

Chapter 4

Formulating the optimization model

4.1 What to formulate

As of today, the few wind farm layout optimization softwares available on the market are primarily based on heuristic algorithms trying to find good enough solutions in reasonable time [7, 17, 12, 23]. These implementations, however, cannot provide a guarantee on the solution quality which a MILP formulation can. There are thus incitements for formulating a MILP and analyzing the outcome on realistic project areas.

The problems defined in Chapter 2 are addressed throughout this Chapter. The first problem is that of maximizing the overall production, considering the maximum number of turbines, minimum distances between turbines and negative interference between turbines, so called wake effects. The second problem deals with the construction costs of the park, for example, the foundation cost at each possible turbine position and the costs of cables and roads connecting the turbines.

4.2 Placing wind turbines and maximizing production

4.2.1 Possible production at different locations

The production maximization problem is formulated on a graph G=(V, E) where the set of nodes V defines possible wind turbine locations, and the edges E denote the paths between each pair of nodes. The placement of a wind turbine in a node in V is represented by the binary variable x_i, which attains the value of 1 if a wind turbine is located at node i, and 0 otherwise. The objective is to maximize the sum of the production, p_i MWh, in every active node (i.e., a node i such that $x_i = 1$) with an additional constraint restricting the maximum number of installed wind turbines to n. This constraint could also represent a restriction of the maximum installed capacity. The integer linear program is thus to

$$
\begin{aligned}
\text{maximize} \quad & \sum_{i \in V} p_i x_i, \\
\text{subject to} \quad & \sum_{i \in V} x_i \quad \leq \quad n, \\
& \quad\ x_i \quad \in \quad \{0,1\}, \quad i \in V.
\end{aligned}
$$

Multiple wind turbine models, or the option of choosing a lower production level in a wind turbine, for example, to decrease sound levels, can be formulated by expanding the nodes in V with new sets of x_i^l, where $l \in L$ denotes the different types of wind turbines. Only one type in each location will be chosen due to the constraints formulated in the next section. A maximum installed capacity constraint when using multiple wind turbines would have to be formulated as the slower method restricting the installed capacity, $\sum_{i \in V} \sum_{l \in \text{types}} p_i^l x_i \leq p_{\max}$, where p_i^l is the wind turbine capacity for turbine type l at location i, instead of the number of wind turbines.

4.2.2 Minimum distance between wind turbines

To prevent wind turbines from being placed too close to each other, constraints regarding the minimum separation distance between wind turbines are formulated. This is done by stating that none of the other nodes within a given range from an active node is allowed to be active. This minimum separation distance is defined as a factor α times the rotor radius, R, when modeling a circular minimum separation zone, but it can also be expressed by for instance an ellipse with it's major axis aligned with the primal wind direction.

Setting the sum of all nodes within the minimum separation distance αR of node i less than or equal to one will not do the trick since the constraint has to hold also for a node where no turbine is placed. At this inactive node it is perfectly alright for more than one node within the distance αR to be active without them being less than αR meters apart, as illustrated in Figure 4.2. In a continuous representation of possible locations this would lead to a formulation where a maximum of five nodes strictly less than the minimum separation distance from i would be allowed when node i is not active, and one otherwise. Six active nodes around an active or non active node is possible only when the distances from these to i is larger than or equal to αR, shown in Figure 4.1. The mathematical representation of this constraint will be the sum of five times the value of node x_i plus the values of the surrounding nodes x_j, where $j \in Q_i$, should be less than or equal to five. The set Q_i is defined as all the nodes j with distance from i less than αR, except for the node i itself. $|E_{ij}|$ denotes the distance in meters between the nodes i and j. The constraints read

$$5x_i + \sum_{j \in Q_i} x_j \leq 5, \quad i \in V,$$

where $Q_i = \{j \in V \mid |E_{ij}| < \alpha R\} \setminus \{i\}$, $i \in V$. The solution of this problem is however slow and a different approach is therefore adopted. Here, the sum of active nodes within a certain radius around a node is set to be less than or equal to one. The radius used is of half the minimum separation distance and now the program rely on the points in between each pair of nodes to set the minimum separation distances, as illustrated in Figure 4.4. At most one wind turbine may be placed within each of these circles. The problem becomes one of packing circles as dense as possible, around active and non active nodes, as seen in Figure 4.3. To allow for additional temporary nodes, used only for these constraints but not as possible wind turbine locations, a new set of nodes and edges, $\bar{G} = (\bar{V}, \bar{E})$ where $G \subseteq \bar{G}$, is introduced. The new constraint, which now will be used in the formulating of this *Production model*, is given by

$$\sum_{j \in Q_i} x_j \leq 1, \quad i \in \bar{V}, \quad Q_i = \{j \in \bar{V} \mid |\bar{E}_{ij}| < \alpha R/2\}.$$

The constraint says that the sum of the active nodes in a set Q_i may not be greater than one, where Q_i consists of all nodes j on distance less than $\alpha R/2$ from node i (including i itself). $|\bar{E}_{ij}|$ denotes the distance in meters between the nodes i and j.

In a discrete representation of possible locations situations can occur where there does not exist any useful points in between two nodes, allowing them to be placed within the minimum distance of one another. This is easily corrected by defining the added temporary nodes at half the distance between each pair of nodes. Obviously this is required only for pairs of nodes separated by approximately the minimum distance, or more specifically by more than half but no more than the full minimum separation distance. Creating these

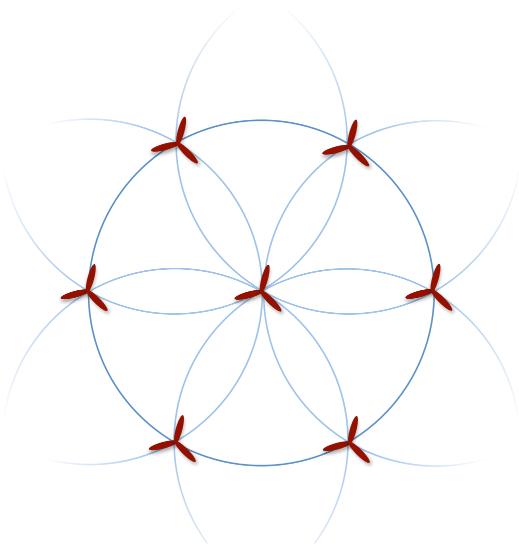


Figure 4.1: **Six wind turbines in an area less than the minimum separation distance of a node is not possible.** The tightest placement of the six are instead on the minimum separation distance.

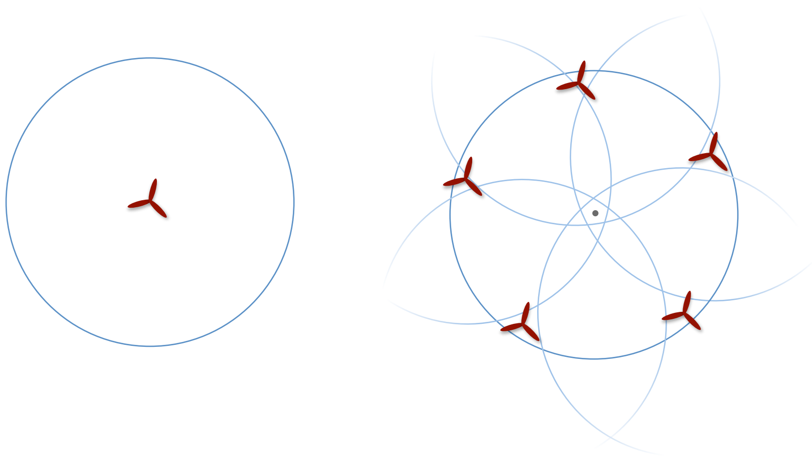


Figure 4.2: **Strictly less than the minimum separation distance of a node,** a maximum of five wind turbines can be placed, depending on the status of the center node.

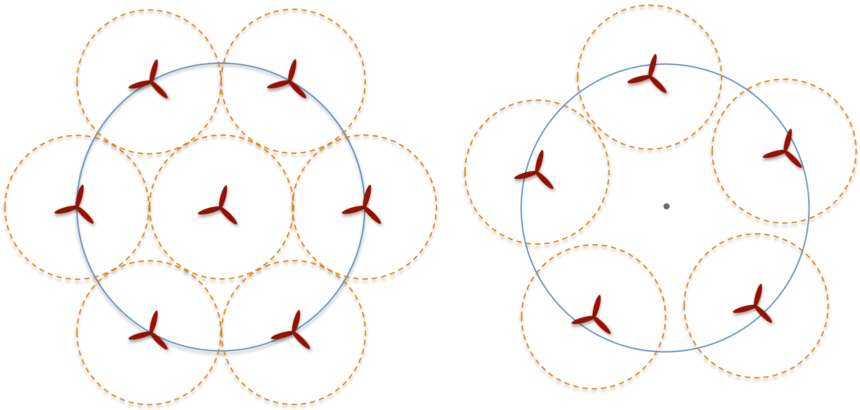


Figure 4.3: **By packing circles of half the minimum separation distance radius,** a maximum of five wind turbines strictly in the area less than the minimum distance of a node is achieved. The locations of the wind turbines are the same as in Figure 4.2.

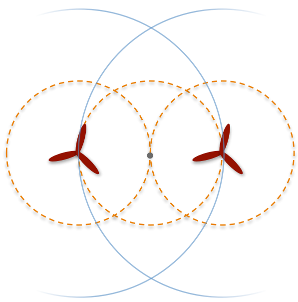


Figure 4.4: **A maximum of one active node in the circle area less than half the minimum separation distance of a node,** forces two active nodes to be separated by at least the minimum separation distance.

sets is particularly easily done in a rectangular or diamond grid, where one just substitutes the grid for that of twice the grid resolution. Minimum separation distance constraints are then created with the nodes in $\bar{V}$, after which the variables of the temporary nodes are deleted but the constraints retained.

4.2.3 Packing density in a discrete grid

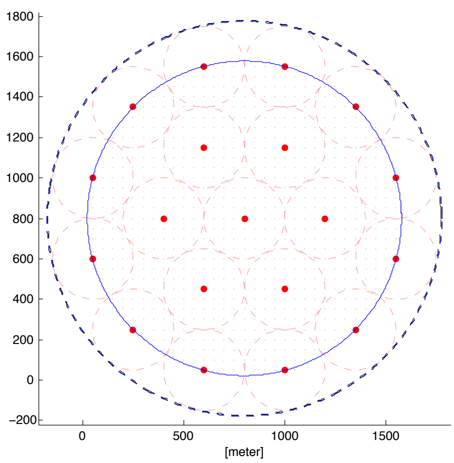
To allow for close to hexagonal packaging, a grid size of the minimum separation distance divided by eight or multiples of eight is recommended. On a bounded area, for which a hexagonal or a regularly spaced packing is optimal, optimization on a discrete grid may yield a solution close to the continuous optimum. Examples of this are shown in Figure 4.5, where the proven continuous optima are obtained from [11] and [10].

4.2.4 Production loss between wind turbines

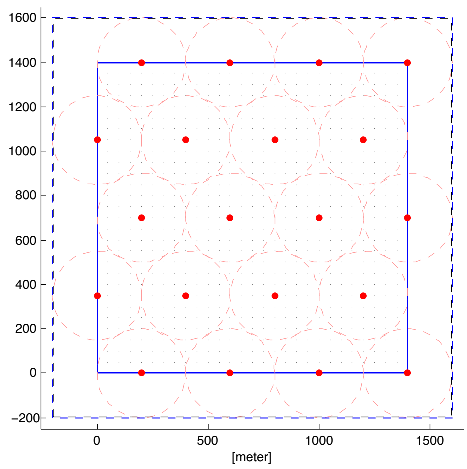
To account for wake effects, that is the loss of production when placing two wind turbines close to each other (see Section 2.1.4), a new variable $w_{ij} \in \{0,1\}$ is introduced. This variable, if equal to one, represents a production reduction at node j due to a wind turbine at node i. This constraint equals the one of the generalized vertex packaging problem [20] and is also used by Donovan in [4, 5] and Donovan et al. in [6]. The magnitude of the corresponding production loss is denoted d_{ij} MWh and must be calculated for every pair of nodes since it depends on the wind frequency and distribution for the specific location and direction of wake. Since the proximity range constraints of Section 4.2.2 prevent the use of edges shorter than αR, these can be excluded from the set E, defining the subset $\hat{E} \subseteq E$. The modified objective function is to maximize

$$\sum_{i \in V} p_i x_i \quad - \sum_{(i,j) \in \hat{E}} d_{ij} w_{ij},$$

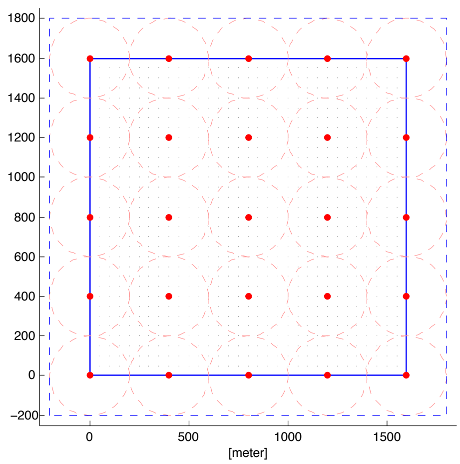
and the added constraints are



(a) 19 circles



(b) 20 circles



(c) 25 circles

Figure 4.5: **Circle packing.** Placement of 19, 20, and 25 wind turbines, respectively, in different geometric areas. The inner black dashed line represents the area required for packing the number of circles in a continuous manner, whereas the outer dashed line is the area required using the formulated linear program. The two dashed lines differ little or nothing in the three examples. The solid line is the project area used by the linear program.

$$
\begin{array}{rcll}
x_i + x_j - w_{ij} & \leq & 1, & (i,j) \in \hat{E}, \\
w_{ij} & \in & [0,1], & (i,j) \in \hat{E}.
\end{array}
$$

The variables $w_{ij} \in \hat{E}$ may take only binary values, but due to the formulation of the constraint relating two nodes and their edges this requirement will be fulfilled in an optimal solution also without integrality requirements on w_{ij}. The sum of the losses from wakes, $\sum_{(i,j)\in\hat{E}} d_{ij}w_{ij}$, is subtracted from the objective function, and the wake constraint forces w_{ij} to 1 if both $x_i = x_j = 1$ (i.e., the nodes i and j are active), otherwise it will take the value of 0.

The formulation of wake effects between two possible locations using linear relations is not restricted to any particular wake model. However, due to linearity the combination of multiple wake effects on a node equal the superposition of all wake effects from active locations on this node. If non-linear superposition is preferred, approximations by linearizations are addressed in Section 4.2.7.

4.2.5 Maximum sound levels

To deal correctly with maximum allowed sound levels one cannot simply define a minimum allowed distance to wind turbines from a noise sensitive area, denoted by an index $s \in \bar{S}$, corresponding to a certain sound level. This is because the sound pressure from two wind turbines will be superposed and the minimum allowed distance thus increased. The constraint regarding maximum sound levels is formulated as the superposition of the individual sound pressures a_{is} in s from active nodes i, which must not exceed a maximum allowed value, a_{max}. When using multiple types of wind turbines, their different sound characteristics must also be implemented in this constraint. Sound pressure (SP) and sound level (SL) relate as $\mathrm{SL} = 10\log_{10}(\mathrm{SP})$. The resulting constraint is given by

$$
\sum_{i\in V} a_{is}x_i \;\leq\; a_{max}, \quad s \in \bar{S}.
$$

4.2.6 The complete Production model

Combining the formulations in Sections 4.2.1, 4.2.2, 4.2.4 and 4.2.5 yields the complete *Production model*, which is here described and presented.

The *Production model* optimizes the total production when placing wind turbines in a project area. It can place at maximum a given number, n, or as many wind turbines as possible. The model respects the minimum separation distance between wind turbines, αR, and that the sum of the individual sound pressures, a_{is}, originating from the wind turbines, is held below the maximum level, a_{max}, for any noise sensitive area, a_s. The production, p_i, in each possible location, and the production loss, d_{ij}, between wind turbines i and j due to wake effects, are calculated from any production and wake effect model (see Sections 2.1.2 and 2.1.4), respectively. The graph $G = (V, E)$ consist of all the nodes V and the edges E connecting each pair of nodes. $\hat{G} = (\hat{V}, \hat{E})$ is the subgraph of G with edges shorter than the minimum separation distance, αR, removed, and $\bar{G} = (\bar{V}, \bar{E})$ is the graph with twice the grid resolution of G, used for implementing the minimum separation distances. The complete *Production model* is to

$$
\begin{array}{llcll}
\text{maximize} & \displaystyle\sum_{i \in V} p_i x_i - \sum_{(i,j) \in \hat{E}} d_{ij} w_{ij}, & & & \\
\text{subject to} & \displaystyle\sum_{i \in V} x_i & \leq & n, & \\
& \displaystyle\sum_{j \in Q_i} x_j & \leq & 1, & Q_i = \{j \in \bar{V} \mid |\bar{E}_{ij}| < \alpha R/2\}, \quad i \in \bar{V}, \\
& x_i + x_j - w_{ij} & \leq & 1, & (i,j) \in \hat{E}, \\
& \displaystyle\sum_{i \in V} a_{is} x_i & \leq & a_{max}, & s \in \bar{S}, \\
& x_i & \in & \{0,1\}, & i \in V, \\
& w_{ij} & \in & [0,1], & (i,j) \in \hat{E}.
\end{array}
$$

The model can also, through defining a revenue on production, be shifted to that of optimizing profit. Here, locally varying costs of foundations can also be incorporated. Additional formulations reducing the number of wake variables, reducing the MIP-gap and addressing linearization of non-linear combination of wake effects are described in Section 4.2.7, however not implemented.

4.2.7 Additional formulations

These suggested formulations or ideas have not, due to time restrictions, been implemented in this thesis. They seem however well worth to investigate in future work.

Dividing an area into sectors to reduce the number of wake vairables

A high resolution grid can be of importance in the tight packing of circles. The accuracy and level of detail of the wake effects need not however depend as greatly on this resolution. It would then mean a great improvement of problem size if the number of wake variables, w_{ij}, did not increase by the square of the number of nodes. The idea is to divide the entire area into square sectors with side lengths less than the minimum separation distance, αR, by $\sqrt{2}$ to contain a maximum of one active node. The number of possible locations within a sector could be large, but the wake effects are only to be calculated between active sectors, see Figure 4.6. Thus, the numbers of wake variables and corresponding constraints grow by a factor of $|V|$ instead of $|V|^2$. Now, moving two wind turbines relative to one another will affect the production in the locations picked, as well as their internal distance and bearing which wakes are calculated from. Examples of distributions in relative movement and bearing between two wind turbines when shifting their location to a sector centre is shown in Figures 4.7 and 4.8. The level of approximation seem reasonable and can be adjusted by shifting the sector side length. The modified wake constraint, if denoting the graph of sectors and interconnecting edges by $G^s = (S, E^s)$, is given by

$$\sum_{k \in S_i} x_k + \sum_{k \in S_j} x_k - 1 \quad \leq \quad w_{ij}, \quad (i, j) \in E^s_{ij},$$

where $S_i \subseteq V$, $i = 1, ..., |S|$, is the subset of the nodes within square sector i. The set E^s are composed by the edges connecting each pair of sectors.

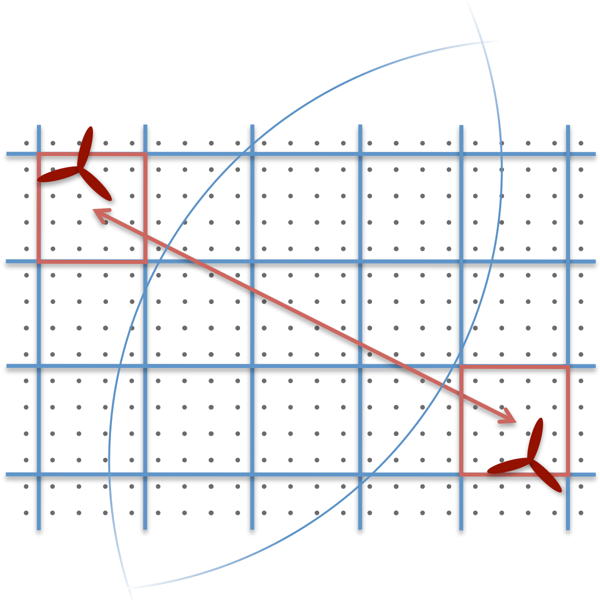


Figure 4.6: **Wakes are calculated between active sector centers,** whereas the resolution within a sector could be much higher.

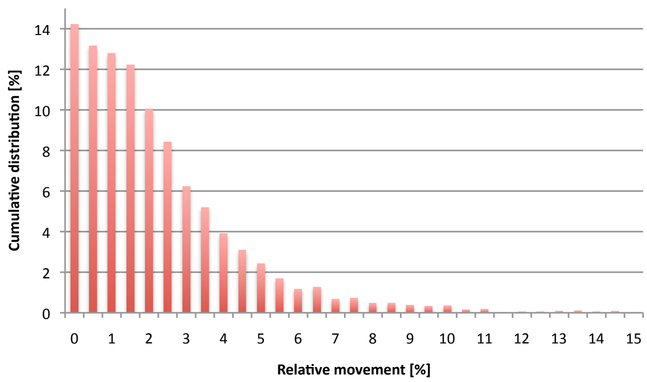


Figure 4.7: **Relative movement,** showing the cumulative distribution of the change in distance between two randomly placed wind turbines when shifting their locations to the nearest respective sector centers. One wind turbine is placed in the square sector with centre in the point $(x, y) = (0, 0)$, the other within $x, y \in [-2000, 2000]$, respecting the minimum separation distance of 400m. The sector side length is 100m.

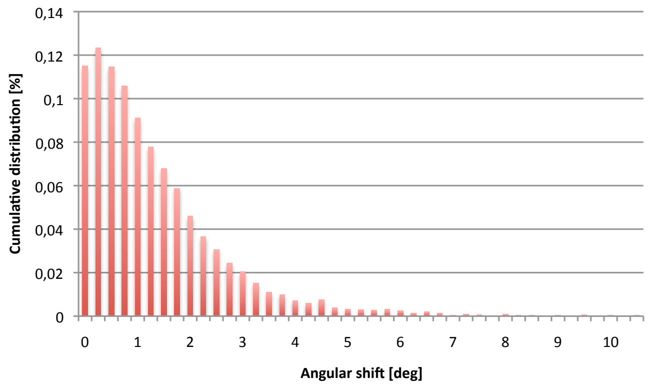


Figure 4.8: **Angular shift,** showing the cumulative distribution of the shift in bearing between two randomly placed wind turbines when shifting their locations to the sector centers. One wind turbine is placed in the square sector with centre in the point $(x, y) = (0, 0)$, the other within $x, y \in [-2000, 2000]$, respecting the minimum separation distance of 400m. The sector side length is 100m.

Reducing the MIP-gap by constraining the LP relaxation

In an optimal solution to the LP-relaxation of the production problem typically $x_i \leq 0.5$, $i \in V$, which implies that there are no active wakes between any nodes. Assuming a minimum allowed number of wind turbines, K, preferably the number of wind turbines used as the maximum number ($K = n$), the following constraints can enforce the proper total amount of wake effects to be activated and thus reduce the MIP-gap. The effect on the solution time due to the increased number of constraints has not been investigated. Using K as the minimum number of wind turbines, constraints forcing the total amount of incoming and outgoing wakes at an active node are given by

$$\begin{aligned}\sum_{j \in E} w_{ij} &\geq (K-1)x_i, \quad i \in \bar{V}, \\ \sum_{i \in E} w_{ij} &\geq (K-1)x_j, \quad j \in \bar{V}.\end{aligned}$$

These constraints do, however, only force one of the nodes to which the wake is connecting being active, and the use of small wakes to and from non-active nodes at a far distance is possible. Additional constraints guarantee that the nodes that an active wake is connecting to are also active, according to

$$\begin{aligned} w_{ij} &\leq x_i, \quad i, j \in \bar{V}, \\ w_{ij} &\leq x_j, \quad i, j \in \bar{V}. \end{aligned}$$

These four inequalities imply that the number of active wakes relates correctly to the minimum allowed number of active nodes, i.e., $\sum_i \sum_j w_{ij} = K(K-1)$. This does, however, not force the LP-relaxation to take on integer values. It is, for example, possible to split two quite separated wind turbines into two times two 0.5-valued ones next to each other. This allows for two times 0.5 valued wakes leaving each half wind turbine, resulting in a total wake value of two instead of one between the original wind turbine locations. The LP relaxation will by this splitting behavior yield large sums of wake variable values between distant nodes, and less between nodes close to each other.

Approximating other ways of combining wake effects

A linear superposition of wake effects will overestimate the combined production losses compared to, for example, the 'sum of squared velocity deficits'-method [14, 22, 9]. To compensate for this difference in a linear program, a positive contribution to the energy production would have to be added when a wind turbine is affected by two or more wakes simultaneously. To compensate for every possible combination of wakes affecting a wind turbine, too many variables would have to be added for the model to remain of reasonable size. However, if accepting the added positive production to be an average of the effect of, say, two arbitrary placed wind turbines placed within 1000m north of the affected wind turbine instead of the positive contribution from their exact locations, a somewhat simpler model can be formulated. Let the binary variable $K_{i,\theta}^{n}$ represent n number of wakes in the angular segment θ of node i affecting a wind turbine in node i. The added production compensation of this variable is $p_{i,\theta}^{n+1} \geq p_{i,\theta}^{n}$, meaning the more wakes combined the larger the production contribution. With $N = 2, 3 \ldots$ being the possible number of interfering wakes, and Θ the set of angular sectors around each node, typically the full circle divided by the angular spread of the wakes, and $S(i, \theta)$ being the set of nodes in the θ segment of node i, the problem is to

$$
\begin{aligned}
\text{maximize} \quad & \sum_{i\in V}\sum_{\theta\in\Theta}\sum_{n\in N} p^n_{i,\theta} K^n_{i,\theta}, \\
\text{subject to} \quad & \sum_{n\in N} K^n_{i,\theta} \;\leq \sum_{k\in S(i,\theta)} w_{ki}, \quad i \in V, \quad \theta \in \Theta, \\
& K^n_{i,\theta} \;\in\; \{0,1\}.
\end{aligned}
\tag{4.1}
$$

The formulation requires a number of added integer variables; this number is equal to $|V|\cdot|\Theta|\cdot|N|$. This is very many additional integer variables, and at the price of further approximations yet another approach is hereby proposed. Instead of correcting for multiple wakes in a number of sectors around each node, a positive production addition that is proportional to the total number of wind turbines aligned in a certain direction of the wind is suggested. This can be seen as sectors looking like to long stripes covering the project site. Sectors in a number of directions will be added, which can be divided into covering not the full stripe but a part of it. The formulation of the linear program will then be similar to that of 4.1, but the choice of $K^n_{l,\theta}$ for stripe l in angle θ will now depend on the number of active nodes within the sector instead of the number of wakes affecting a certain node. This approximate formulation is to

$$
\begin{aligned}
\text{maximize} \quad & \sum_{l\in L}\sum_{\theta\in\Theta}\sum_{n\in N} p^n_{l,\theta} K^n_{l,\theta}, \\
\text{subject to} \quad & \sum_{n\in N} K^n_{l,\theta} \;\leq \sum_{i\in S(l,\theta)} x_i, \quad l \in L, \quad \theta \in \Theta, \\
& K^n_{l,\theta} \;\in\; \{0,1\}.
\end{aligned}
\tag{4.2}
$$

The number of additional variables in this model is $|L|\cdot|\Theta|\cdot|N|$, where L is the set of stripes, Θ the set of directions of the stripes (which in this formulation only need to cover $[0,\pi)$) and $N = 2, 3, ...$ is the possible number of wind turbines within the stripe. This will generally result in a lot fewer variables compared to that of the model (4.1) but it is a more rough approximation and does not account for the wakes expanding. However, the two models (4.1) and (4.2) provide an approximate basis for correcting the linear superposition of wakes to more closely relate to a non-linear combination method.

4.3 Infrastructure formulation

Another optimization area of interest concerns the costs associated with the constructions of foundations as well as roads and cables necessary for the wind farm. In these formulations the locations of the wind turbines are considered fixed. However, the simultaneous solution of the *Production model* and the *Infrastructure model*, not requiring the locations to be fixed, is possible but seems to be a rather complex problem.

4.3.1 Foundation costs

The foundation costs of Section 2.2.1, c_i, are modeled simply as a negative revenue per active node i, added to the objective function. The objective function is then to

$$\begin{aligned} \text{maximize} \quad & -\sum_{i \in V} c_i x_i, \\ \text{subject to} \quad & x_i \in \{0, 1\}, \quad i \in V, \end{aligned}$$

which, if the locations of the wind turbines are fixed by the prior solution of the *Production model*, can only take on one value.

4.3.2 Formulating a Steiner tree

When connecting the wind turbine locations with roads and power lines, a spanning tree structure, as defined in Chapter 3.3, is to be constructed. An example of this is connecting the nodes of Figure 4.9 as shown in Figure 4.10, a *minimum spanning tree* solution obtained by the model described below. This example will be addressed again when discussing cables in Section 4.3.4.

Each edge is associated with a certain cost, d_{ij} defined by a number of factors such as length, topography and such. A linear programming formulation which will minimize the sum of the edge costs is sought. However, the formulation should also allow for, but not require, the tree to use paths involving nodes that are not active as locations of wind turbines. Such a formulation

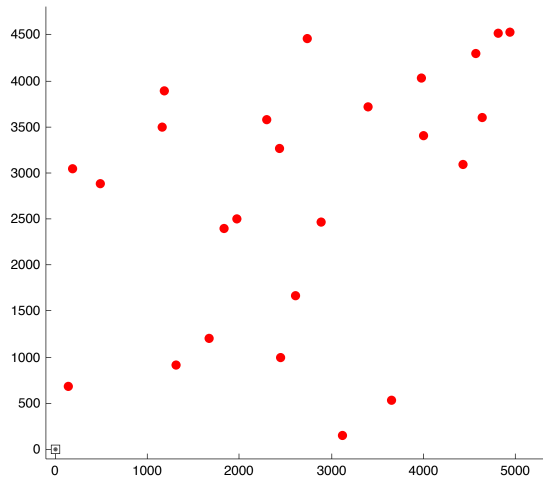


Figure 4.9: **Randomly scattered nodes,** to be connected by the cheapest possible tree structure. The bottom left square node is also to be connected, but will later be used as the point of connection to the external grid.

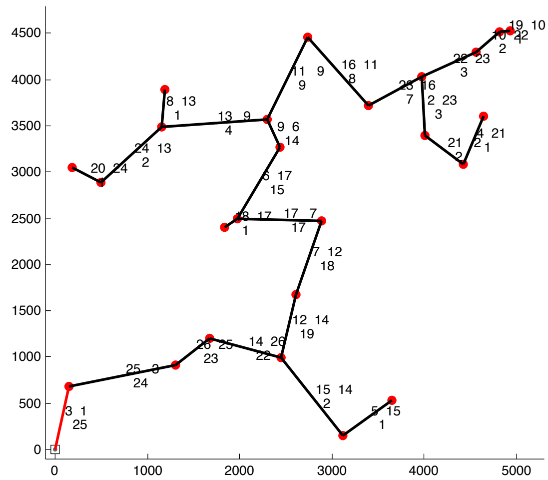


Figure 4.10: **A minimum spanning tree** considering only the layout of the edges and not minding that they are directed, connecting the nodes of Figure 4.9. There are no limits on the flow through the edges. The upper two figures of each triple of numbers state which two nodes the edge is connecting and the bottom figure state the flow through the edge.

could allow for a shorter tree (if defining an euclidean tree) or ease the process of formulating paths around obstacles (Figure 4.11), and results in a so called Steiner tree problem [3, 2]. Since the problem is that of electricity flow, at some central point of the tree the flow should be passed on to an external grid.

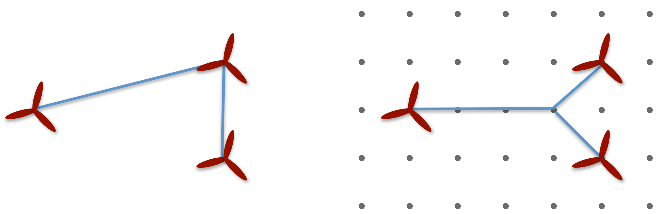


Figure 4.11: **Shorter paths are possible,** when using other nodes than just the active ones (Steiner tree formulation). The left path is 5% longer than the right one.

A positive revenue, r_i, per node $x_i \in V$, is defined such that the solution to the minimization of tree costs does not lead to only the trivial solution of all x_i being zero. When using this tree formulation in conjunction with the *Production model*, this revenue is replaced with the one from that model. The nodes with positive revenue are the ones previously selected as wind turbine locations. The graph of possible nodes and edges for the tree formulation is defined as $G_t = (V_t, E_t)$, where the set V from the *Production model* is a subset of V_t and E_t are the edges connecting each pair of nodes in V_t. The objective function of the tree formulation of this graph, using the binary variable $z_{ij} \in E_t$ to define the use of an edge from i to j, then is to

$$\text{maximize} \quad \sum_{i \in V} (r_i - c_i) x_i - \sum_{(i,j) \in E_t} d_{ij} z_{ij}, \tag{4.3}$$

where $x_i \in \{0, 1\}$, $i \in V$ and $z_{ij} \in \{0, 1\}$, $(i, j) \in E_t$.

To deal with the actual creation of the tree structure, and to prevent the occurrence of subtours, a new continuous variable $y_{ij} \in [0, f]$, representing the flow of electricity in the tree, is introduced. Here, $f \in [0, n]$, where n is the maximum allowed number of nodes connected at, or flow in, edge (i, j). Each active node creates one unit of flow and flow entering a node must also leave the node. The accumulated flow at the connection node (the external power grid), denoted $CN \in V_t$, must equal at least the number of active nodes minus one. We denote the flow on edge (i, j) by y_{ij}. This connection node constraint is not really necessary, since the flow must be exported somewhere and the connection node is the only option, but it was discovered having a slight effect on reducing the time of finding a feasible solution. The constraints read

$$\begin{aligned} \sum_{ij\in E_t} y_{ij} + x_j - \sum_{jk\in E_t} y_{jk} &= 0, && j \in V_t \backslash \{CN\}, \\ \textstyle\sum_{i\in V_t} y_{i,CN} &\geq \sum_{i\in V} x_i - 1, && \\ y_{ij} &\in [0, f_{ij}], && (i,j) \in E_t. \end{aligned} \tag{4.4}$$

Now, y_{ij} is the number of wind turbines connected on the branch from node i, including i, and it's flow can be limited either individually for each edge (i, j) or for all $(i, j) \in E_t$. The binary variable defining the activation of an edge is however z_{ij} and to induce the activation and cost of an edge we connect these two by replacing the last line of (4.4) by the constraints

$$0 \leq y_{ij} \leq f_{ij} z_{ij}, \quad (i,j) \in E_t. \tag{4.5}$$

Using the formulation illustrated in (4.4) with the modification of (4.5), isolated subtours are not possible due to the accumulative property of the variables y_{ij}, illustrated in example Figure 4.12. There exists feasible solutions with subtours connected to paths leading to the connection node, but it is then cheaper to break up the subtour, as illustrated in Figure 4.13.

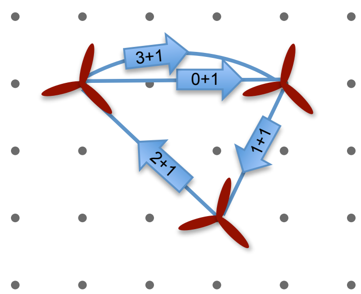


Figure 4.12: **Isolated subtours,** are not possible since flow is added in each active node and the flow along an edge can take on only one value.

The basic tree formulation can then be written as the linear program to

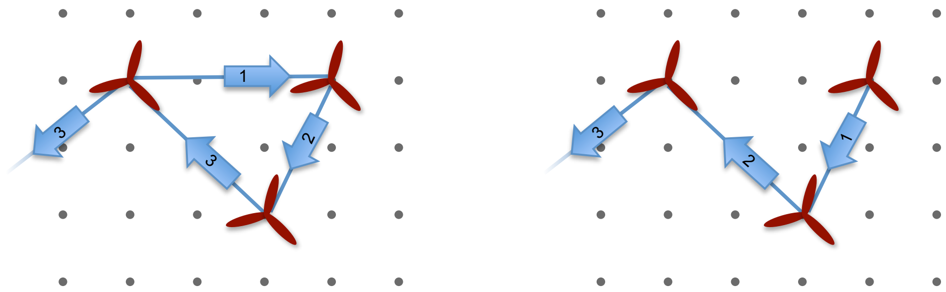


Figure 4.13: **Non-isolated subtours,** are possible, but breaking the subtour will yield a shorter path.

$$
\begin{array}{llcll}
\text{maximize} & \displaystyle\sum_{i\in V} r_i x_i - \sum_{(i,j)\in E_t} d_{ij} z_{ij}, & & & \\
\text{subject to} & \displaystyle\sum_{i:(i,j)\in E_t} y_{ij} + x_j - \sum_{k:(j,k)\in E_t} y_{jk} & = & 0, & j \in V_t \backslash \{CN\}, \\
& \displaystyle\sum_{i\in V_t} y_{i,CN} & \geq & \displaystyle\sum_{i\in V} x_i - 1, & \\
& 0 \leq y_{ij} & \leq & f_{ij} z_{ij}, & (i,j) \in E_t, \\
& x_i & \in & \{0,1\}, & i \in V, \\
& z_{ij} & \in & \{0,1\}, & (i,j) \in E_t.
\end{array}
$$

On top of this formulation, optional constraints are added. The following constraints do not alter the optimum solution, but have proven to reduce computation time (in the sense of improved solutions after any tested amount of time). The constraints

$$y_{ij} \geq z_{ij}, \quad (i,j) \in E_t,$$

relates the flow through and activation of an edge in the opposite way that (4.5) does. The need for at least one active edge variable, z_{ij}, per active node is expressed by

$$x_i - \sum_{j:(i,j)\in E_t} z_{ij} \leq 0, \quad i \in V.$$

A flow leaving a node requires either flow entering the node or the node itself being active, which is formulated by the constraint

$$\sum_{i:(i,j)\in E_t} z_{ij} + x_j - \sum_{k:(j,k)\in E_t} z_{jk} \;\; \geq \;\; 0, \quad j \in V_t \backslash \{CN\}.$$

Moreover, forcing flow in only one direction per edge is beneficial and can be expressed by

$$z_{ij} + z_{ji} \;\; \leq \;\; 1, \quad (i,j) \in E_t.$$

However, flow in multiple directions along an edge should be permitted but will be prohibited elsewhere and this latter formulation is therefore not an restriction. More on this in Section 4.3.7.

4.3.3 Minimizing the costs of roads

It is possible to limit the number of edges per node in a grid to only the shortest edge in, say, sixteen directions, resulting in a bit more costly spanning tree but reducing the number of variables z_{ij} and y_{ij} to $16|V_t|$ instead of $|V_t|^2$ each. If roads are to be connected along an existing road instead of at a single defined point, nodes with corresponding edges could be added along this road and the cost of these edges set to zero, as in the example illustration of Figure 4.15. Since roads and cables typically follow the same paths – due to the need of a road to place a cable – the cost of a road is merely an additional cost added to the cable cost, defined in Section 4.3.4, and no additional road formulations to the Steiner tree are thus given here. Two cables along the same path only requires one road to be built, which is solved in the same manner as for a single cable by defining double cables as a cable type in the mathematical formulation. This implemented model will be able to handle double cables when flows are parallel. However, in reality two cables with opposite flows will also be beneficial in terms of road costs, as per the rightmost illustration in Figure 4.14. Another formulation addressing this issue, where separate road variables are introduced, is given in Section 4.3.7, however not implemented.

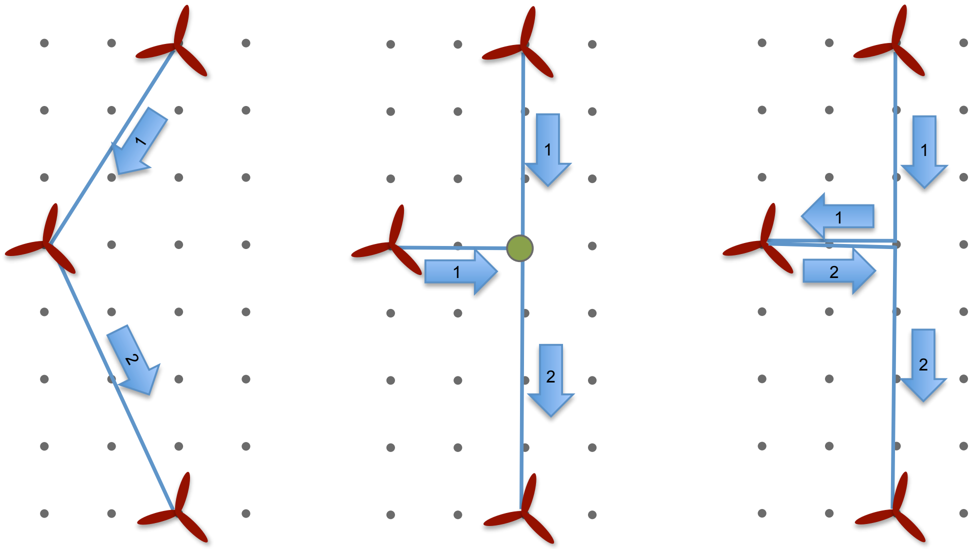


Figure 4.14: **Three different ways of connecting the wind turbines.** 1) The shortest spanning tree using only edges between wind turbine locations. 2) The cost of combining cables is low relative to the standard cable cost. 3) Cost of cable is low relative to the cost of combining cables and multiple cables on the same path yields only one road cost.

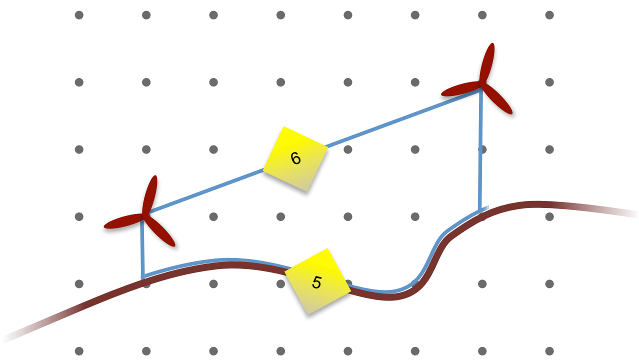


Figure 4.15: **Using existing roads** may reduce the cost of the path. Costs are given in the diamond symbols. The cable path following the existing road is longer but cheaper than the straight path.

However, if the tree of roads is to be connected at a different point than the connection node, where the electrical flow is exported, additional constraints are introduced. These constraints requires that at least one node in the spanning tree of edges – not necessarily a location of a wind turbine – is connected to any of K possible road connection points. The formulation starts by defining a cost C_{ki} for connecting point $k \in K$ to node i. The connection variable, R_{ki}, is forced to zero if no flow is sent from the node i through any of the edges (i, j), where $j \in V_t$, meaning that node i is not a part of the tree. With this variable connecting the external road points only to nodes being part of the Steiner tree, all that is left is to require at least one of the variables R_{ki} corresponding to these connection roads to take the value of one, and thus being active. This linear program formulation is to

$$
\begin{array}{lrcll}
\text{maximize} & -\sum\limits_{i \in V_t} \sum\limits_{k} C_{ki} R_{ki}, & & & \\
\text{subject to} & \sum\limits_{j \in V_t} z_{ij} - \sum\limits_{k \in K} R_{ki} & \geq & 0, & i \in V_t, \\
& \sum\limits_{i \in V_t} \sum\limits_{k \in K} R_{ki} & \geq & 1, & \\
& R_{ki} & \in & \{0, 1\}, & i \in V_t,\ k \in K.
\end{array}
$$

4.3.4 Cables

The capacity of a cable y_{ij} is upper bounded by the maximum number of connected wind turbines possible for that type of cable, f_{ij}. The node configuration of Figure 4.9, with only wind turbine locations as possible edge connection nodes, is solved with an upper limit of the number of wind turbines per edge, except for the unlimited export cable to node 1. The resulting layout is shown for a limit of three wind turbines per edge in Figure 4.16 and seven in Figure 4.17. Multiple cable types, $l = 1, 2, ...$, with different maximum capacity and costs, are introduced as the sets of variables $y_{ij}^l \in [0, f_{ij}^l]$ and $z_{ij}^l \in \{0, 1\}$, corresponding to the same edges (i, j) as for the original cable type. The cable types may include, for example, larger dimension cables or double low-capacity cables. Including these new variables in the flow forwarding and connection node constraints (4.4) will allow the model to choose between the cable types. The higher price of the high capacity cable will result in a solution that uses the cheapest cable that meets the capacity requirements. The previous example using both cable types mentioned, with different costs, is shown in Figure 4.18. Whether the solution time will benefit from an additional constraint restricting the use of only one z^l_{ij} per edge (i, j), which in practice will always be the case, has not been investigated.

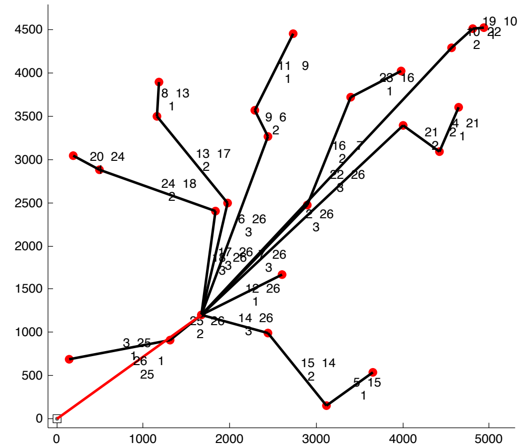


Figure 4.16: **An upper limit of three wind turbines per edge.** With a cable cost of 0.6 MSEK/km, and export cable cost of 3 MSEK/km, the total cost of connecting the nodes of Figure 4.9 is 22.7 MSEK. This configuration was found after about 25 s with a MIP-gap of 4% and was proven optimal after 243 s. The cable connecting to the external power grid in point (0,0) is the high capacity export cable originating from the transformer station.

Now, the transition between two cable types will occur either when passing a wind turbine, thus shifting the capacity demand, or when combining the flow from two different locations into one cable. The former case as well as a split at a wind turbine location is considered to come without extra costs, since the necessary equipment is already present. However, when the cables are to be combined at a location other than at a wind turbine, the cost of the required substation has to be accounted for. This is formulated by introducing a binary substation variable, s_j, where $j \in V_t$, with the associated cost g_j, which should be forced to a value of one if flows are combined outside of wind turbine locations. If the combination occurs at the location of a wind turbine, $s_j = 0$. This is achieved by requiring *either* s_j or x_j to take on a value of one if *more than* one flow enters the node j. The flow combination program is thus to

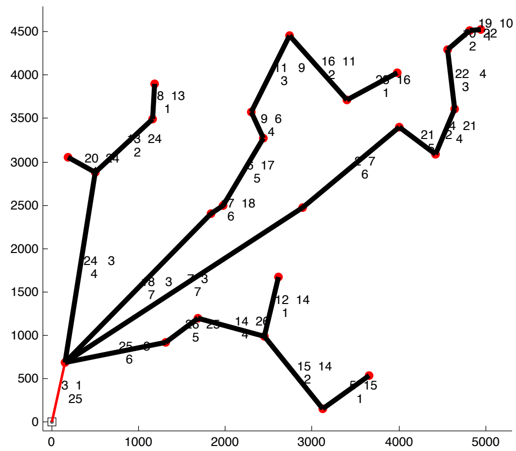


Figure 4.17: **An upper limit of seven wind turbines per edge.** With a cable cost of 1 MSEK/km, and export cable cost of 3 MSEK/km, the total cost of connecting the nodes of Figure 4.9 is 24.3 MSEK. This configuration was found and proven optimal after 11 s. The cable connecting to the external power grid in point (0,0) is the high capacity export cable originating from the transformer station.

$$
\begin{array}{llcll}
\text{maximize} & -\sum_{j\in V_t} g_j s_j, & & & \\
\text{subject to} & \sum_{i\in V_t} z_{ij} - 1 - M(x_j + s_j) & \leq & 0, & j \in V_t, \\
& s_j & \in & \{0,1\}, & j \in V_t,
\end{array}
\tag{4.6}
$$

where M equals the maximum number of incoming flows to a substation. It would be possible to relax the integrality requirement on s_j, and rewrite $-M(x_j + s_j)$ as $-(Mx_j + s_j)$, with an upper limit of s_j higher than one, but the cost of the substation would then be proportional to the number of incoming flows, which is typically not the case.

Flows can be combined, but are not permitted to be split, which calls for additional constraints. These constraints, limiting the number of flows out of a node i, are defined as

$$
\sum_{j\in V_t} z_{ij} \leq 1, \quad i \in V_t.
$$

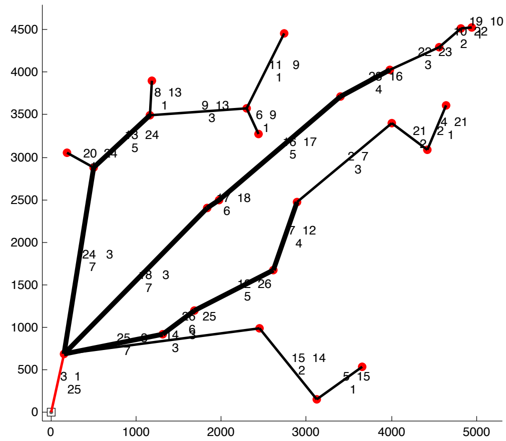


Figure 4.18: **Multiple cable variables with upper limits of three and seven connected wind turbines, respectively.** Using the prices from Figures 4.16 and 4.17 the total cost of connecting the nodes of Figure 4.9 is 20.4 MSEK. This configuration was found after about 50 s with a MIP-gap of 6% and was proven optimal after 47 min. The cable connecting to the external power grid in point (0,0) is the high capacity export cable originating from the transformer station.

4.3.5 Transformer station placement

The wind turbines of a wind farm generally need to be connected to the export cable through a transformer station. In Figures 4.16 – 4.18 the location of the transformer station is where the red export cable connects to the tree. Its optimal location will depend on the length and cost of the internal cables, the export cable and the foundation costs at the given location. If the latter does not vary locally, the cost of the transformer station could potentially be corrected for later, since the location of the export cable connecting to the root of the tree then only will depend on the relative prices of the cables. However, the use of transformer placement variables and costs has the additional function of restricting the number of transformer stations. Without these costs it could, for example, be beneficial to place two export cables at different locations and thus constructing two trees with the requirement of two transformer stations. This problem could of course be solved in other ways, for example, by restricting the number of export cables to one, $\sum_{i \in V_t} z_{i,CN} \leq 1$, but for this model the transformer placement, including locally varying costs, is included. The objective function is extended by the sum of the costs c_i^{tr} of the continuous transformer placement variables $t_i \in V_t \backslash \{CN\}$, which will take on the values only 0 or 1 due to the fact of $z_{i,CN}$ being binary. The linear programming formulation of the added transformer costs is to

$$
\begin{array}{llrcll}
\text{maximize} & - \displaystyle\sum_{i \; \in V_t \backslash \{CN\}} & c_i^{tr} t_i, & & & \\
\text{subject to} & & z_{i,CN} & \leq & t_i, & i \in V_t, \\
& & t_i & \in & [0,1]. &
\end{array}
$$

4.3.6 The complete Infrastructure model

Combining the above formulations yields the complete *Infrastructure model*, which is hereby described and presented.

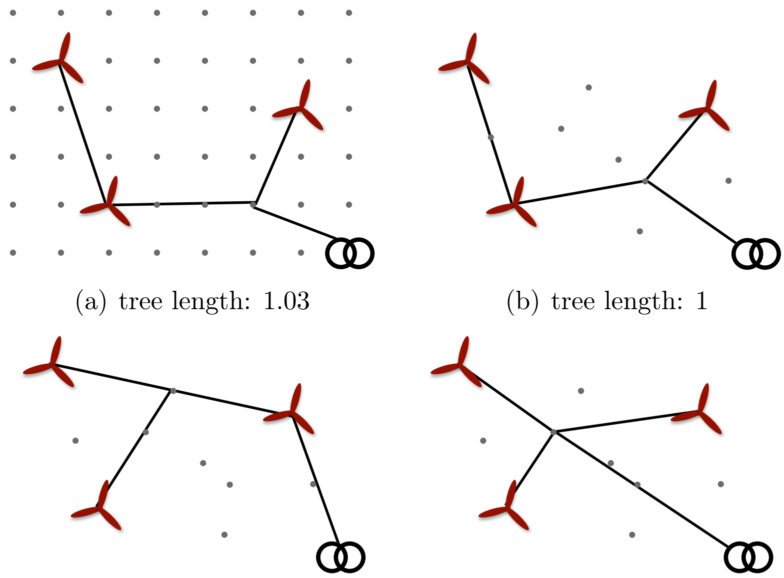
The *Infrastructure model* is that of connecting the wind turbines by roads and cables, and of locally varying foundation costs. If the locations of the wind turbines are fixed, the location variables x_i could be set to their defined values and no revenue, r_i, would be necessary. For the purpose of combining the two models, however, this revenue is included in the objective function. The graph $G_t = (V_t, E_t)$ contains the nodes and edges used to form the Steiner tree. The set, V, of possible or fixed locations of wind turbines is a subset of V_t. The nodes of V_t can be many more than the number of wind turbines, making cheaper connections possible, and includes the point of connection to the external grid, x_{CN}. Power flow along an edge is represented by y_{ij} and one unit of flow is added in each active node along a cable. The sum of the flow entering the connection node must not be less than the number of active nodes, or this number minus one if wind turbine placement is allowed in the connection node. Due to this accumulated flow and the need to export the flow, isolated subtours are prohibited. Several types of cables, y^l_{ij}, where $l = 1, 2, ...$, with different maximum flows, f^l_{ij}, can be used on an edge. The cost, d^l_{ij}, of cable and road type l on edge (i, j) is activated by the binary variable z^l_{ij}. Moreover, constraints requiring at least one road R_{ki} with cost C_{ki} be connected to the road tree are included. There are also constraints introducing a substation cost g_j through the combination variable s_j where flow is combined outside a wind turbine location. Flow splits are prohibited by another constraint, and the location specific cost of a transformer station, c^{tr}_i, is introduced at each transformer station location by the variable t_i. The complete *Infrastructure model* is to

$$
\begin{array}{lrcll}
\text{maximize} & \multicolumn{4}{l}{\displaystyle\sum_{i\in V} r_i x_i - \sum_{(i,j)\in E_t} d_{ij} z_{ij} - \sum_{i\in V_t}\sum_{k} C_{ki} R_{ki} - \sum_{j\in V_t} g_j s_j - \sum_{i\ \in V_t\setminus\{CN\}} c_i^{tr} t_i,} \\
\text{subject to} & \displaystyle\sum_{i:(i,j)\in E_t} y_{ij} + x_j - \sum_{k:(j,k)\in E_t} y_{jk} & = & 0, & j \in V_t\setminus\{CN\}, \\
& \displaystyle\sum_{i\in V} x_i - 1 & \leq & \displaystyle\sum_{i\in V_t} y_{i,CN}, & \\
& 0 \leq y_{ij} & \leq & f_{ij} z_{ij}, & (i,j) \in E_t, \\
& y_{ij} & \geq & z_{ij}, & (i,j) \in E_t, \\
& x_i - \displaystyle\sum_{j:(i,j)\in E_t} z_{ij} & \leq & 0, & i \in V, \\
& \displaystyle\sum_{i:(i,j)\in E_t} z_{ij} + x_j - \sum_{j:(j,k)\in E_t} z_{jk} & \geq & 0, & j \in V_t\setminus\{CN\}, \\
& \displaystyle\sum_{j\in V_t} z_{ij} - \sum_{k\in K} R_{ki} & \geq & 0, & i \in V_t, \\
& \displaystyle\sum_{i\in V_t}\sum_{k\in K} R_{ki} & \geq & 1, & \\
& \displaystyle\sum_{i\in V_t} z_{ij} - 1 - M(x_j + s_j) & \leq & 0, & j \in V_t, \\
& z_{i,CN} & \leq & t_i, & i \in V_t, \\
& x_i & \in & \{0,1\}, & i \in V, \\
& z_{ij} & \in & \{0,1\}, & (i,j) \in E_t, \\
& R_{ki} & \in & \{0,1\}, & i \in V_t,\ k \in K \\
& s_j & \in & \{0,1\}, & j \in V_t \\
& t_i & \in & [0,1], & i \in V_t.
\end{array}
$$

4.3.7 Additional formulations

Reducing the number of nodes when wind turbine positions are held fixed

If turbine positions are held fixed the number of nodes in the grid defining the tree structure can be reduced. It would, for example, generally not be of any use to keep nodes on the outside of the "outermost" wind turbines. Also, better geographic locations for combining flows, for example, at the centre of a group of wind turbine locations, can be found for various combinations of the wind turbines. Thus, the set of nodes can through a bit of smart preprocessing be redefined to a smaller but better placed set of nodes. The grid can



(a) tree length: 1.03

(b) tree length: 1

(c) tree length: 1.08

(d) tree length: 1.15

Figure 4.19: **Reduced number of nodes.** A principal example where 48 nodes are reduced to 11 consisting of nodes half way between active nodes and of nodes at a few cost weight centers. (c) and (d) only serve to illustrate possible new paths.

hence be transformed into a set of locations which suits the configuration of wind turbines even better. Figure 4.19 shows examples of these principles.

If the *Infrastructure model* is combined with the *Production model*, and thus, the locations of the wind turbines are not fixed, this type of rearrangement is not possible. However, a similar approach to that of defining wake effects between sectors (Section 4.2.7) may be adopted, creating the tree structure connecting geographical sectors instead of specific points (which are included in these sectors).

Separation of tree variables into cables and construction variables

Two or more cables on the same path is in the *Infrastructure model* formulated by an edge variable with the cost of multiple cables but the cost of only one road. However, it cannot handle, for example, opposite directions of flow, which would also lead to only one road having to be built. A simple solution to this problem, increasing the number of binary variables and constraints by the number of undirected edges, is to split the active path variable z_{ij} into cable variables z^c_{ij} and road construction variables z^r_{ij}. This does not call for creating two sets of trees – the cable variables are used as before in, for example, the flow forwarding constraints (4.4), but the cost of road construction is moved to the variable z^r_{ij} which is activated whenever a cable path is used. Since the road construction cost does not depend on the direction of the power flow, two nodes need to be connected by only the undirected z^r_{ij}, yielding the constraints

$$\begin{array}{rcll} z^c_{ij} + z^c_{ji} & \leq & 2z^r_{ij}, & \{j \in V_t \mid j > i\},\ i \in V_t, \\ z^r_{ij} & \in & \{0, 1\}. & \end{array}$$

This solves the problem of the costs not decreasing when using a path as in the rightmost illustration of Figure 4.14. However, due to the restriction of not allowing splits of flow, this type of path, where flow enters and leaves and then enters and leaves the same location again, is still not possible. A solution to this is to also implement the idea in the next section.

Multiple nodes per location to allow for additional paths

Even though it is sometimes economically beneficial to use a path as in the rightmost illustration of Figure 4.14, the restriction of not being able to split flow in a node also makes it prohibited to split flow in the same location. One solution to this problem would be to add an additional set of nodes in the same locations as the original set. Multiple nodes in the same location means that one flow could enter and leave one node and a second flow could enter and leave another node but in the same geographic location. Using the additional formulation above, with i_1 representing node 1 at location i and l being the number of nodes per location, the constraints read

$$\sum_{m=\{1,l\}} \sum_{n=\{1,l\}} z^c_{i_m,j_n} + z^c_{j_n,i_m} \leq M z^r_{ij}, \quad i, j \in V_t : j > i,$$

where $M \geq l$. This shift of nodes and edges requires changing the sets used in many of the other constraints too.

The problem size increases rather fast by the introduction of these multiple nodes, mainly because the number of edge variables y_{ij}^l and z_{ij}^l grows by the square of this multiple. The number of variables needed in different locations of the project area could however vary, as, for example, shown in Figure 4.20, which would result in the problem size growing at a slower rate.

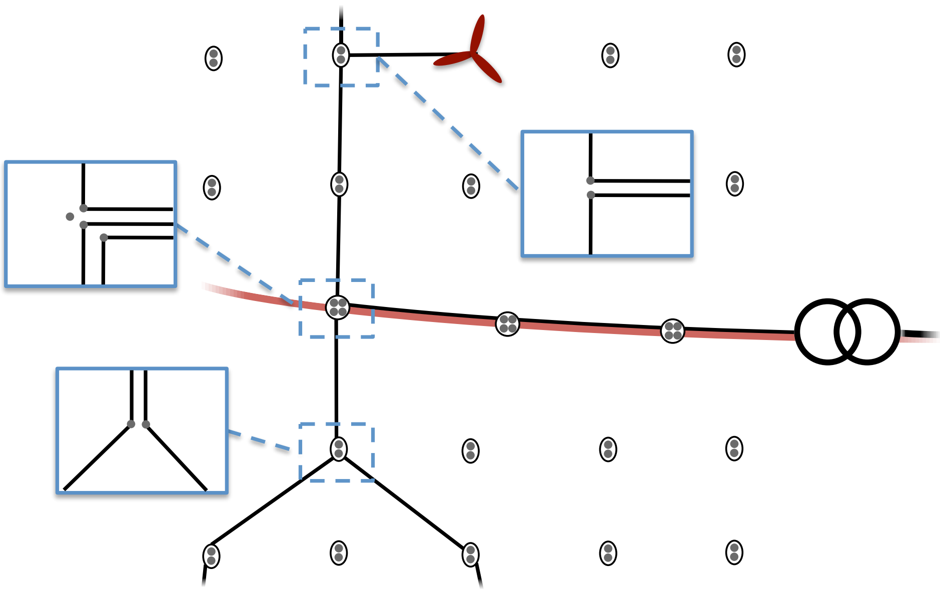


Figure 4.20: **Multiple nodes in each location.** The number of possible nodes can be determined locally depending on the characteristics of the setup.

Now, this formulation is preferably used when the locations of the wind turbines are fixed, whence multiple nodes are not needed at these locations, and the rearrangement of nodes to geographically superior locations, as in Figure 4.19, can be implemented.

Redefining the problem to that of connecting wind farms

The problem of connecting wind turbines could resemble that of connecting wind farms optimally to nearby high voltage export cables. These reformulated constraints aim to, as with the whole tree structure problem, modeling not all but some parts of the problem. An illustration of this problem is shown in Figure 4.21.

In this formulation, the values added to the flow through the cables of the

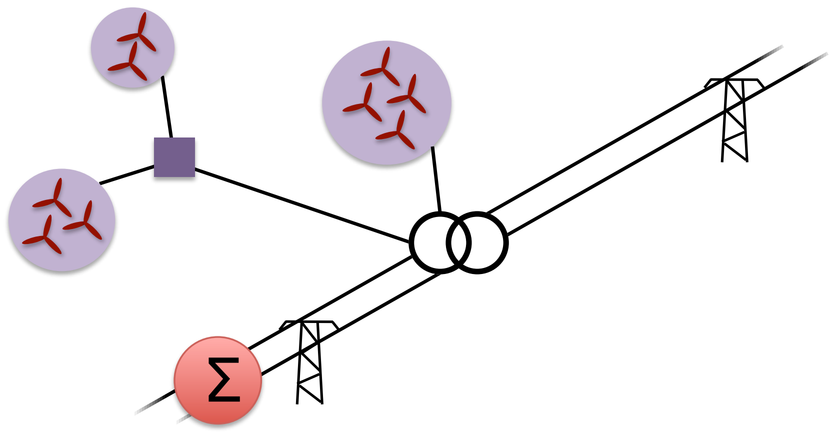


Figure 4.21: **Wind farms connected to a high voltage export line.** The farms can be combined in substations and must be connected with a transformer station somewhere along the export line. The node in which the incoming flow is accumulated does not have to be where the transformer station is placed.

Steiner tree are that of the number of wind turbines in the connected parks. Thus, the variable y_{ij} represents the power flow and the value of this flow in a point on the transmission line equals the total number of connected wind turbines. Two modifications to the original problem of substations (4.6) are addressed. First of all, the costs of substations combining flows is dependent on the magnitude of the entering flows. An additional constraint is introduced, which forces one of the binary substation capacity variables K_j^p, where p is a capacity in the set P_j, of node j, to take a value of one if the value of the substation variable, s_j, equals 1. The cost, g_j^p is now associated with K_j^p, where $g_j^{p_2} \geq g_j^{p_1}$ if $p_2 \geq p_1$. The new formulations for substation costs are to

$$\begin{array}{llcll} \text{maximize} & -\sum_{j \in V_t} \sum_{p \in P_j} g_j^p K_j^p, & & & \\ \text{subject to} & \sum_{i \in V_t} y_{ij} - M(1 - s_j) - \sum_{p \in P_j} p K_j^p & \leq & 0, & j \in V_t, \\ & K_j^p & \in & \{0, 1\}. & \end{array}$$

Now, this formulation generate costs when combining cables, thus requiring substations. To deal with the connection and transformation to a existing high voltage export line, a connection which could have only one input cable, an additional constraint is needed. If the connection could occur at any number of nodes along this line, this set of nodes, V_{el}, is subtracted from the additional substation constraint just formulated and is instead included in the constraint of the linear program to

$$\begin{array}{llcll} \text{maximize} & -\sum_{j\in V_{el}}\sum_{p\in P_j} g_j^p K_j^p, & & & \\ \text{subject to} & \sum_{i\in V_t} y_{ij} - \sum_{p\in P_j} pK_j^p & \leq & 0, & j \in V_{el}, \\ & K_j^p & \in & \{0,1\}. & \end{array}$$

To handle the arbitrary connection point on this export line, the flow gathering connection node of the tree structure is set as one of the nodes one the export line with flow variables y_{ij} and z_{ij} from the other nodes in V_{el} being without limit and cost respectively. Thus, even though summing the flow in one specific point, connection to the export line could occur anywhere along its path.

4.4 Combined revenue formulation

When investing in a wind farm an estimate of future electricity price is bound to be undertaken. A monetary value of the produced electricity over a certain length of time is then available, and the objective of the production optimization is changed into a revenue representation. This allows for combining the revenue and cost objectives. The temporary revenue formulation $\sum_{i\in V} r_i x_i$ of (4.3) is thus replaced by the expression of revenue from the production formulation. Introducing the constraints of the *Production model* then makes it possible to optimize both production and the costs at the same time, with a total profit objective. This combined problem does however seem to be very complex and since, for example, rearrangement of Steiner tree nodes (Section 4.3.7) cannot be used for this case, this is not recommended. The costs of roads and cables do also seem to be too small relative to the production profit to have any major effect on the wind turbine placements. This combined formulation was tested in this thesis but the problems possible to solve at a reasonable level and time were too small to be of interest in this report.

Another approach would be to define a multi-objective optimization model where one of the two objective functions is represented by a constraint, and by optimizing for different limits of this new constraint a trade-off curve, a so called Pareto effective front, between high production and low infrastructure costs can be created. It is not certain, however, how good such an approach would work, whence, for example, the wind turbine locations potentially may vary substantially between solutions. Multi-objective optimization is not tested in this thesis.

Chapter 5

Tests and results

This chapter starts off by listing optimization results for a real wind farm project site, shown in Figure 5.1, and the comparison of this optimization to the one of the commercial software *WindPRO 2.6 Optimize*, from now on referred to as *Optimize*. Further on, the performance of the *Production* and *Infrastructure model* in terms of speed and MIP-gap is expressed. The total production of any layout found will be calculated through the software *WindPRO 2.6 PARK*, from now on referred to as *PARK*. For references on MIP-gap and other mathematical terms, see Chapter 3.

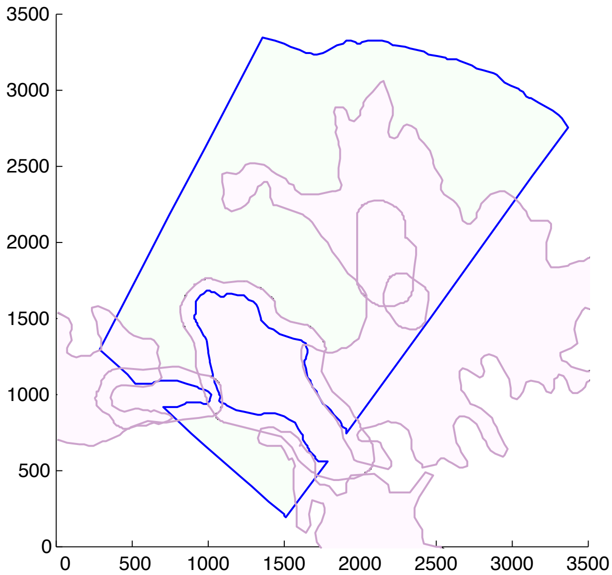


Figure 5.1: **Project area** of the verification examples. The green area is project area whereas the pink areas are wetlands.

As up to this point, the models have been invariant to how production, wakes and costs are calculated. For testing purposes however, a choice of these methods is needed. In the following results, production is calculated using a power curve for a Vestas V90 2MW wind turbine at 105m hub height, based on an average of the site atmospheric pressure. The wind energy content is calculated through WindPRO using the WindATLAS method [8, 18] and array losses are expressed through the Jensen/Risø single wake model [13, 14]. The sound pressure level propagation model used is the one recommended by Naturvårdsverket [1]. Furthermore costs of wind turbines, roads, power lines and the price of electricity are set at reasonable levels. The models are implemented as matrices in Matlab and optimization is carried out using the optimization package Cplex 12.1 with its Matlab application programming interface. The computers used are i686's with four Dual Core AMD

Opteron(tm) 270 processors, 4 GB of internal memory and running Linux 2.6. The Cplex optimization does however only make use of multithreading, that is, dividing an operation to run the parts on different processors instead of one, in parts of it's optimization process. When reaching certain problem sizes, Cplex Matlab API returns 'out of memory'-statuses. These problems could probably to some extent be corrected by formulating them through, for example, the AMPL modeling language where many more options are available. Computing large scale problems is however not the focus of this thesis and too large problem sizes will be omitted.

5.1 Verification example and comparison to commercial software

5.1.1 Optimization of an actual project

For these verification purposes the model is tested on a proposed project site with a site area according to Figure 5.1 and locally varying wind energy content, mainly due to topography, according to Figure 5.2. Optimization is performed using the two grid sizes 100m and 50m, with a minimum separation distance of 400m, translating to just below 4.5 rotor diameters for the selected wind turbine.

Examples of the layout is shown in Figure 5.3 for the 100m resolution with and without wake effects considered. Not modeling wake losses does not resemble reality but can be used as an aid in parts of the optimization process. It will, if wind energy content does not vary much or not at all within the project area, translate into fitting as many wind turbines as possible. This maximum number is valuable when trying to find the maximum total production of an area.

The maximum numbers of 28 and 30 wind turbines for the grid resolutions 100m and 50m, respectively, is also found when array losses decrease the production in tight placements. An example of this is shown for the 100m grid in Figure 5.3. In this example, the *PARK* total production is practically the same no matter whether the wake effects are considered in the optimization process or not, which is seen in Figure 5.4. A reason, other than the limited number of ways of fitting 28 wind turbines in the area, for the solution

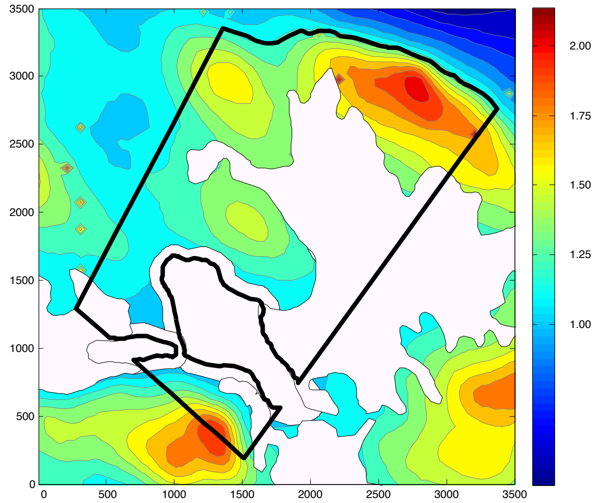
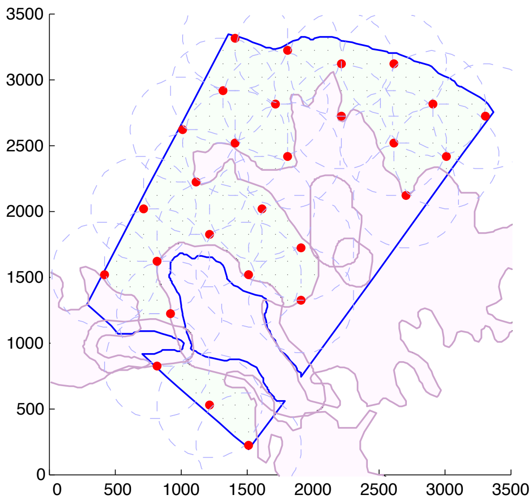
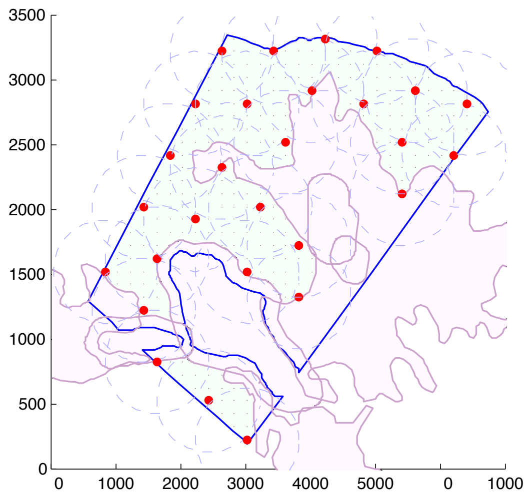


Figure 5.2: **Wind energy content.** Scaled to values of one being the minimum and two the maximum energy content within the project site area. The energy content is calculated using the WindATLAS method, which uses topography as well as other inputs.



(a) Without wakes



(b) With wakes

Figure 5.3: **The same number of wind turbines** are placed when accounting for wake effects as when not. Solution (a) is found after 0.2 seconds whereas solution (b) is the one obtained after 60 minutes at a MIP-gap of 13.5%.

considering the wake effects not yielding a better *PARK* calculation than the other is the fact that the *Production model* over-estimates the production losses when combining multiple wakes relative to the sum of squares of the deficits-method [21]. In the configuration of the Figure 5.3(b), the total losses by wake effects is calculated to be 27% larger than the losses calculated by *PARK*. In such tight placements the choice of wake combination model may thus yield significantly different results.

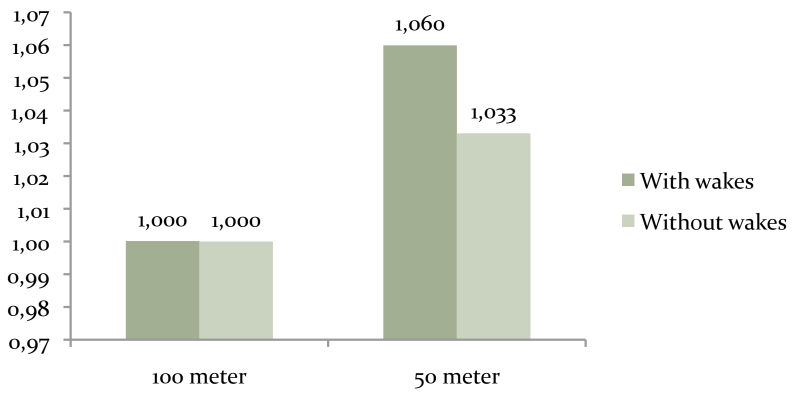
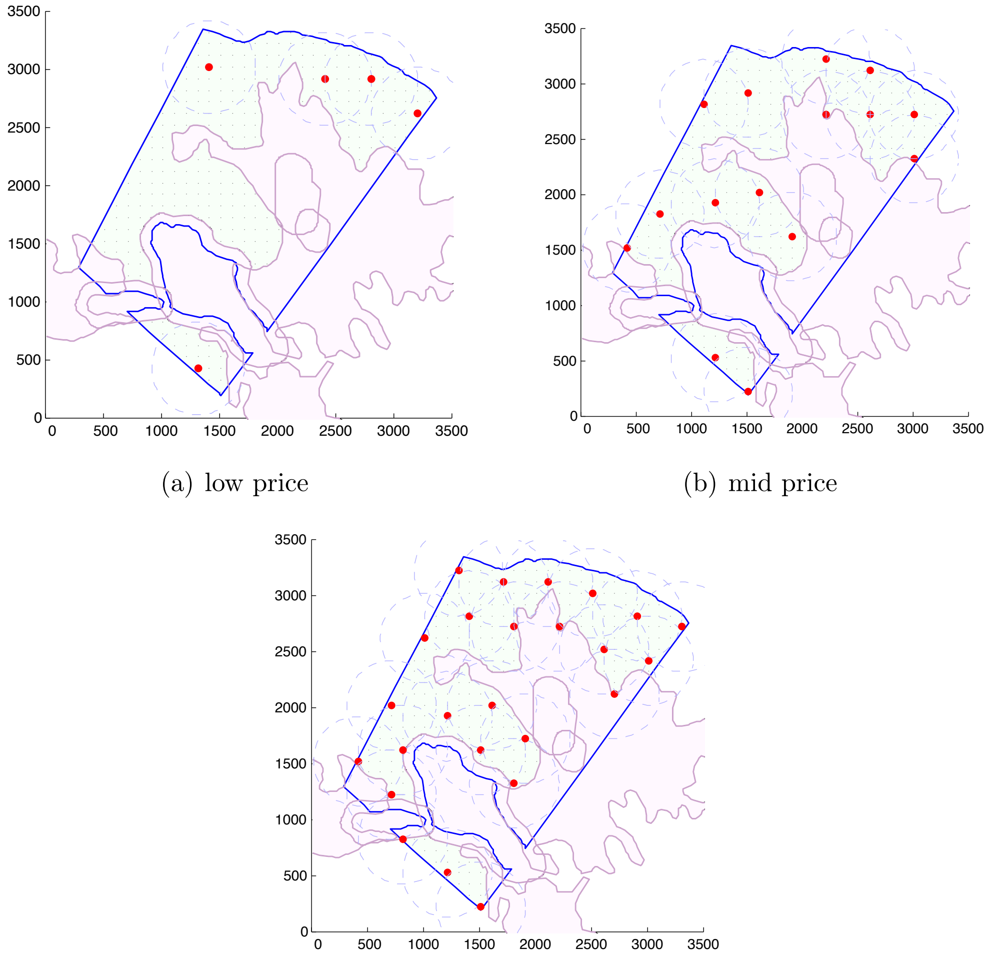


Figure 5.4: **Total production with and without wake effects considered.** The number of wind turbines in the 100m and 50m layout is 28 and 30, respectively. The values are scaled with the 100m layout without wakes.

Now, if only optimizing over the total production, placing as many wind turbines as possible into the area will generally be the case also when accounting for wake effects, since these are small relative to the undisturbed production. If the placement of yet another wind turbine in a small wind farm yields a total production gain of merely a fraction of a percent, it is still a gain. If, however, maximizing the profit, the construction cost of the additional wind turbine has to be accounted for, which would render the investment non-profitable. The *Production model* can be utilized in this manner, and solutions to the problem with shifting prices of electricity is shown in Figure 5.5.

A representation of the MIP-gaps when solving this example area at different prices of electricity is shown in Figure 5.6.

For the *Infrastructure model* optimization there is the option of including additional nodes outside the fixed wind turbine locations or not. Figure 5.7



(a) low price

(b) mid price

(c) high price

Figure 5.5: **Solutions for shifting prices of electricity.** At low prices only a few locations are profitable, whereas at high prices even tight placements are profitable.

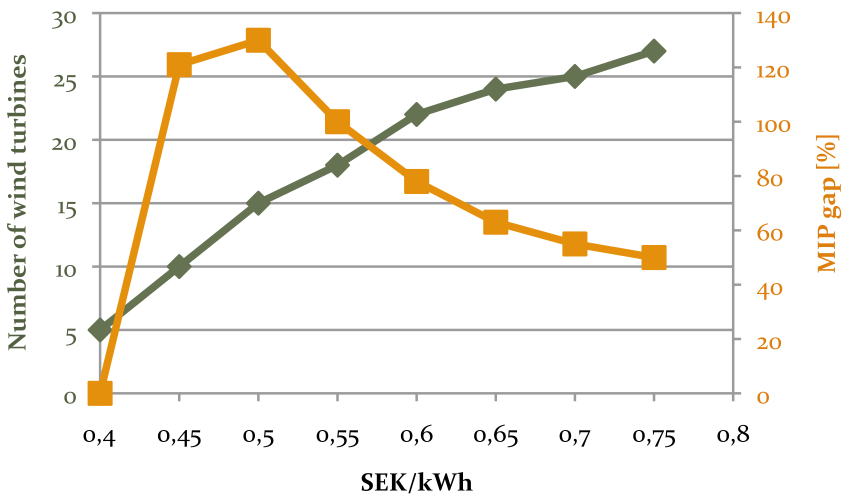


Figure 5.6: **The number of wind turbines and MIP-gap** for shifting prices of electricity. The computing time is 30 min.

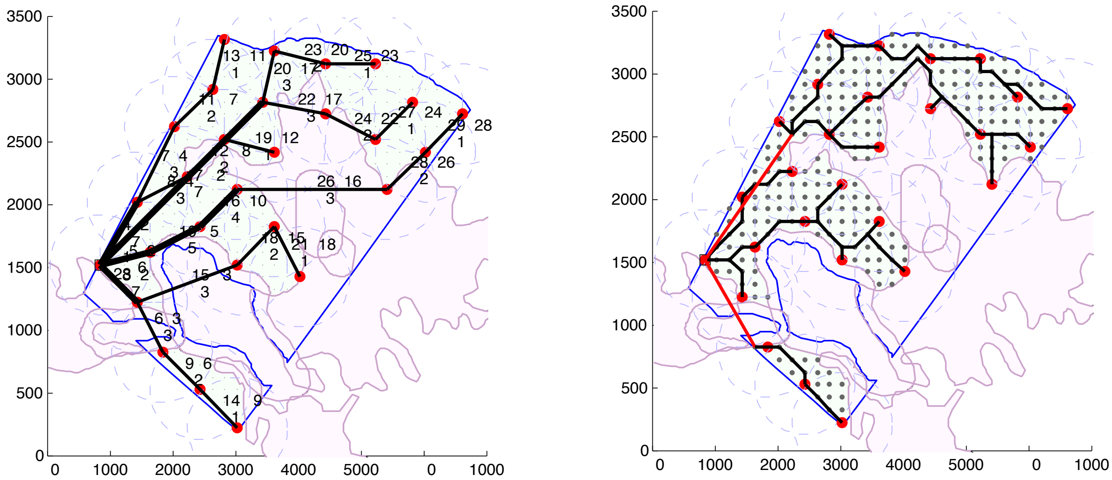
shows examples of both, but it is clear that the additional node formulation is to heavy for the problem to be solved in reasonable time - it had to be formulated using only one cable type and with edges in only eight directions to be able to yield any result in half an hour of computation time.

5.1.2 Comparison to commercial optimization software

To form an opinion of whether the optimization model developed is performing at commercial standards, a simple comparison to the module *Optimize* in *WindPRO 2.6* is made. The problem tested is the project area example from the previous section.

Problem modifications and the algorithm of the software *Optimize*

The software *Optimize* uses a heuristic placement algorithm (similar to the *greedy* algorithm) instead of a linear program formulation as in this thesis. Generally, if solving the same problem, a specialized algorithm has the advantage of possibly being faster than an integer linear programming solver



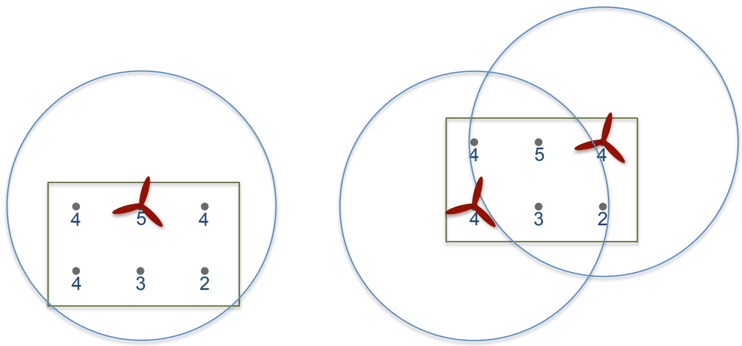
(a) Only edges between wind turbines

(b) Additional edges allowed

Figure 5.7: **Examples of the Infrastructure model output.** (a) A choice between two cables is possible with a maximum capacity of 3 and 7 wind turbines, respectively. The cable paths in the left part of the figure does not form a loop but the paths cross one another. (b) Edges to the eight closest nodes are allowed in every node and one cable with a capacity of 7 wind turbines is used. More edges per node are needed to allow for the transportation of flow even in the narrow passages and to reduce the zig-zag patterns. The computing time is 30 minutes per problem.

which, according to Chapter 3, grows rapidly in complexity with the size of the problem. However, the heuristic algorithm would normally be unable to find some of the solutions and is therefore not bound to find the optimal solution nor is it providing a measure of the quality of the solution. Solving exactly the same problem with the two different methods is desired, which in this comparison means reducing the problem to only maximizing the production without any formulations regarding infrastructure. The restrictions due to sound levels also need to be removed since *Optimize* model them merely as a minimum distance from each wind turbine and thus can not superimpose the effect from multiple wind turbines. Now, the formulations will still not be the same since - as discussed when formulating the wake effects - the *Production model* utilizes a linear superposition of the array losses, which is not possible to choose in *WindPRO*. Here, the sum of squares of the velocity deficit method [8] is used which, according to [14, 9], is a better representation of multiple wake combinations than is linear superposition, (although the reason for this is not absolutely determined [22]). Since the sum of squares-method is used both for the *Optimize* optimization and when *PARK* calculates the total production of the layout found [21], *Optimize* has the advantage of actually trying to optimize using the same wake formulation as *PARK* does, whereas the *Production model* would be optimizing a problem fundamentally different than this (i.e. using linear superposition). A discussion of the importance of this discrepancy is undertaken in the following chapter.

The principal heuristic algorithm of *Optimize* is presented as follows [8]. *Optimize* utilizes a grid of possible locations, preferably coordinates from the calculated wind resource map, within the project site area, and starts off by placing a wind turbine in the best of these locations. It thereafter tries placing a second wind turbine in any of the twenty or so best remaining locations, obeying minimum distances and in each test calculating the array losses, and chooses the location which yields the maximum combined production. It now tries moving the first wind turbine to possibly achieve a better total production of the two. Once these two are set, a third wind turbine is tested in various locations and positioned in the best found, then the two existing wind turbines are moved one at a time to try and find a better total production. This continues, with the possibility to fix the placements when a chosen number of wind turbines has been placed, until all desired wind turbines are placed or the algorithm cannot find space to introduce yet another one. The algorithm could be run quite fast, but has an essential drawback. Since it is only allowed to move one wind turbine at a time, it is easy to block a set of solutions which could be proven better than the



(a) The solution found by the software *Optimize*

(b) The solution found by the *Production model* (defined in Section 4.2.6)

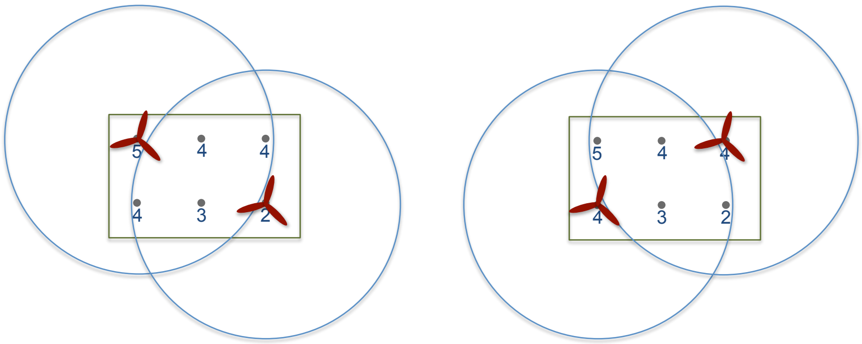
Figure 5.8: **Failing to place an optimum number of wind turbines.** (a) The heuristic algorithm of the software *Optimize* first chooses the best position for the first wind turbine. Next it tries placing another wind turbine, which is not possible given the position of the first and the objective value is 5. (b) The optimal value is 8.

solution found, as illustrated in figures 5.8 and 5.9. This is partly similar to, for example, placing billiard balls at random positions into the triangular rack. Soon you will not be able to place yet another ball without touching the ones already placed, and the rack is not filled.

There is also an option in the software *Optimize* called *auto fill*, which by no means performs any form of optimization but is a tool to just fit as many wind turbines as possible into the project site area. Neither this tool could, however, place the maximum number of wind turbines in any of the example problems, as seen in the 50m problem in Figure 5.10.

The comparison

The optimization is carried out using the same possible wind turbine coordinates, project site area, wind turbine type, minimum separation distance, and wind resource data as in the previous section. Due to the mentioned drawbacks of the heuristic algorithm, *Optimize* is not able to fit as many wind turbines into the project site area as the *Production model*. Here it is made sure the number of wind turbines placed before locking their positions is set higher than the maximum number that can be fitted in the area. For the 100m grid a total of 18 wind turbines are placed by *Optimize* and



(a) The solution found by the software *Optimize*

(b) The solution found by the *Production model* (defined in Section 4.2.6)

Figure 5.9: **Failing to place the wind turbines in optimum locations.** (a) The heuristic algorithm of the software *Optimize* first chooses the best position for the first wind turbine. Next it places another wind turbine in the best of the still available locations. It thereafter tries moving the first wind turbine to find a better total solution, which is not possible. The solution found has objective value 7 but the optimal value, shown in (b), is 8.

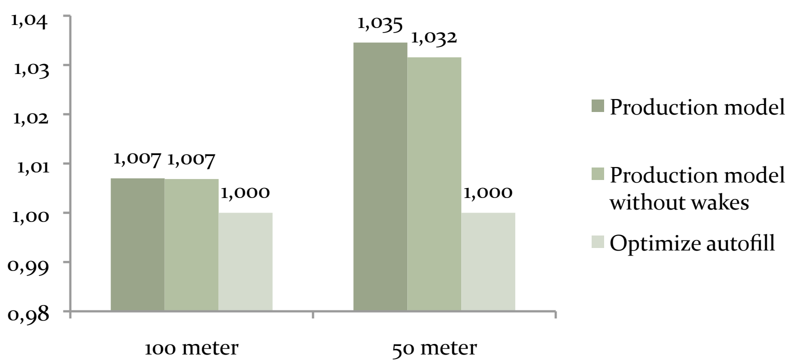
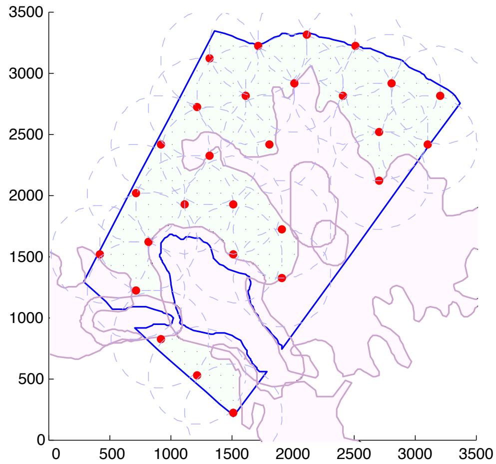


Figure 5.10: **Total production comparison between the *Production model* and the software *Optimize* autofill option.** The project area is the one of Figure 5.1 with two different grid resolutions. The number of wind turbines are [28 28 28] for the 100m case and [30 30 29] for the 50m case. The production values are normalized by the production values of the *Optimize* autofill layouts.

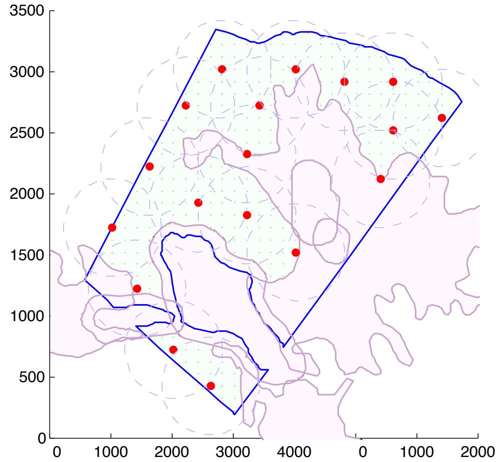
for the 50m grid a total of 20, to be compared to 28 and 30, respectively, by the *Production model*. The respective layouts can be viewed in Figure 5.11. Now, performing a PARK calculation for these layouts yields values of the total production to be compared to the results from the *Production model* solutions. The results indicate a 40% higher production for the layouts found by the *Production model*, shown in Figure 5.12. A tighter layout will of course yield a lower efficiency, but the aim of both optimization methods is to maximize the total production and not the efficiency.

Comparing the total production results of the *Optimize* locations with the ones obtained from the *Production model* when restricted to a maximum of 18 and 20 wind turbines, for the 50m and 100m grid respectively, the *Production model* proved to yield an equal or higher production. The total production of the 50m grid problem, where lack of memory restricted the maximum wake distance to 900m, was equal to the heuristic *Optimize* locations. In the 100m grid problem, the *Production model* yielded a solution about one percent better than that of the heuristic *Optimize*. Total production comparison for various maximum number of wind turbines is shown in Figure 5.13. The method of combining multiple wakes in a different way than the "true" *PARK* calculation model did thus not yield a worse solution.

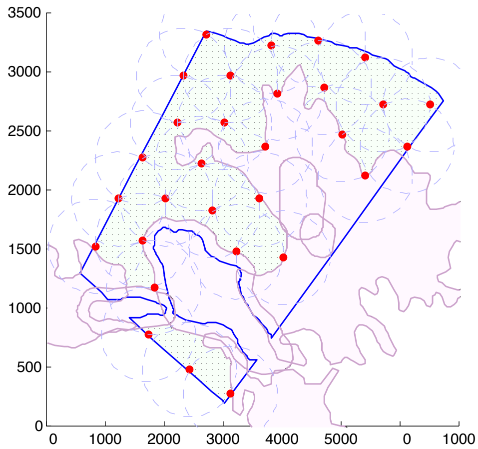
Now, it is proposed in the *Optimize* documentation that a grid of 10–25m is to be used. Running an optimization on a 10m grid using the earlier wind characteristics a solution containing 20 wind turbines is once again found, but this time better positioning led to over half a percent production increase compared to earlier simulations.



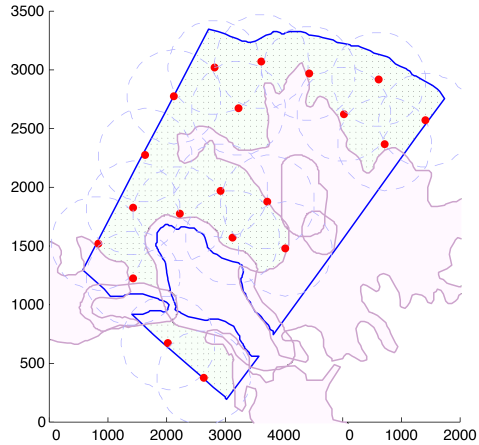
(a) *Production model* 100m



(b) *Optimize* 100m



(c) *Production model* 50m



(d) *Optimize* 50m

Figure 5.11: **Layout comparison.** The *Production model* places about 50% more wind turbines which yields about 40% higher total production, which is the objective of both optimization methods. The *Production model* computations were terminated after about one hour and the computations with the software *Optimize* took 25 minutes for the 100m problem and 80 minutes for the 50m problem.

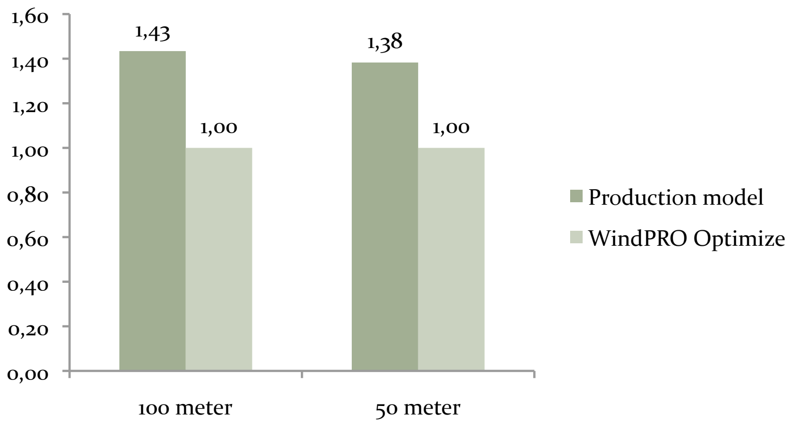


Figure 5.12: **Total production comparison between the *Production model* and the software *Optimize*.** The project area is the one of Figure 5.1 with two different grid resolutions. The number of wind turbines are [28 18] for the 100m case and [30 20] for the 50m case. The production values are normalized by the production values of the *Optimize* layouts.

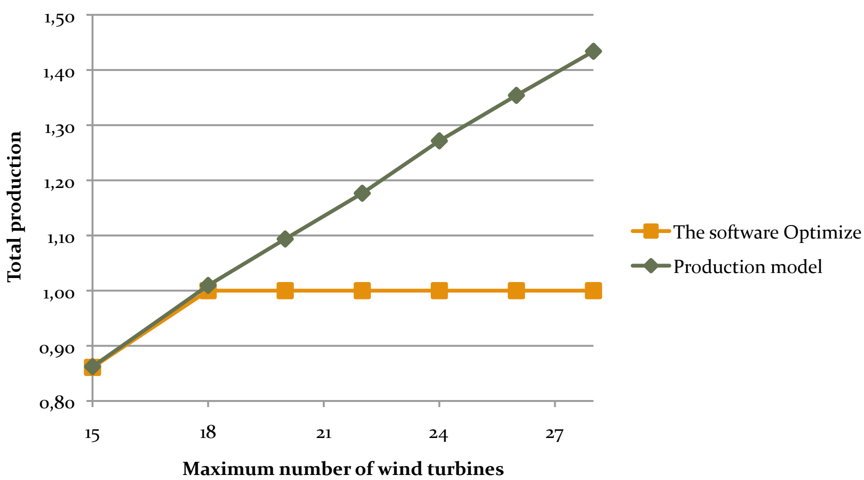


Figure 5.13: **Total production comparison with shifting maximum allowed number of wind turbines.** Project area is the one of Figure 5.1 with resolution 100m. For example, when trying to place at most 24 wind turbines, the *Production model* resulted in a 28% higher total production, and if restricted to a maximum of 18 wind turbines the gain was about 1%.

5.2 Performance

In this section the speed of the models, using the software and hardware setup described in the beginning of Chapter 5, are tested. The values of solution time or MIP-gap would probably change with many of the suggested further implementations, and should only bee seen as guidance.

The Production model

The use of wake effect variables highly increases the problem size and performance for varying project area and number of nodes is shown in Figure 5.14; the number of wake variables per number of nodes can be seen in Figure 5.15.

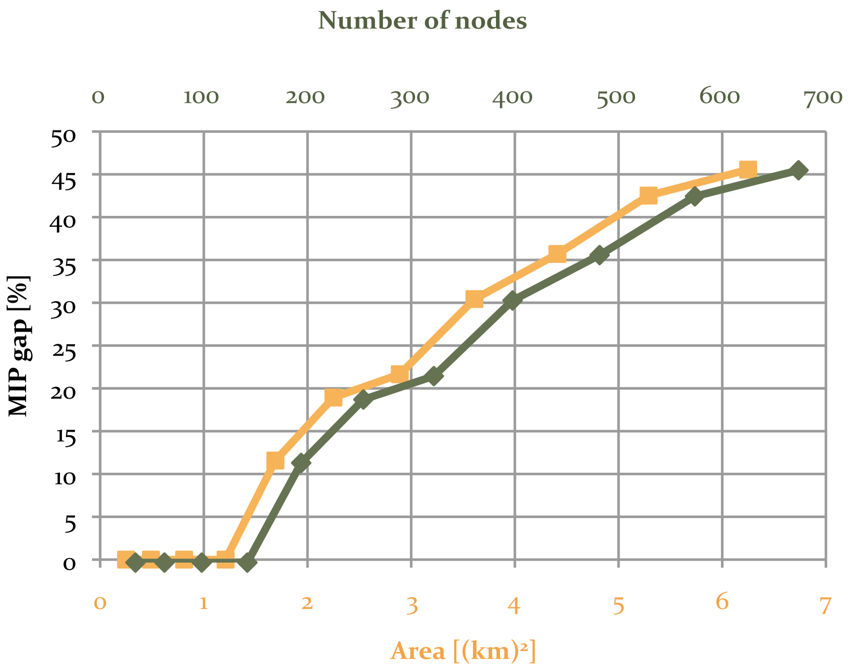


Figure 5.14: ***Production model.* MIP-gap with increasing number of nodes and area size.** 30 minutes of computations on a square area, where the grid size is 50m and the minimum separation distance is 400m.

When optimizing using a profit objective instead of a production objective the MIP-gap, which is the ratio between the best objective value of the

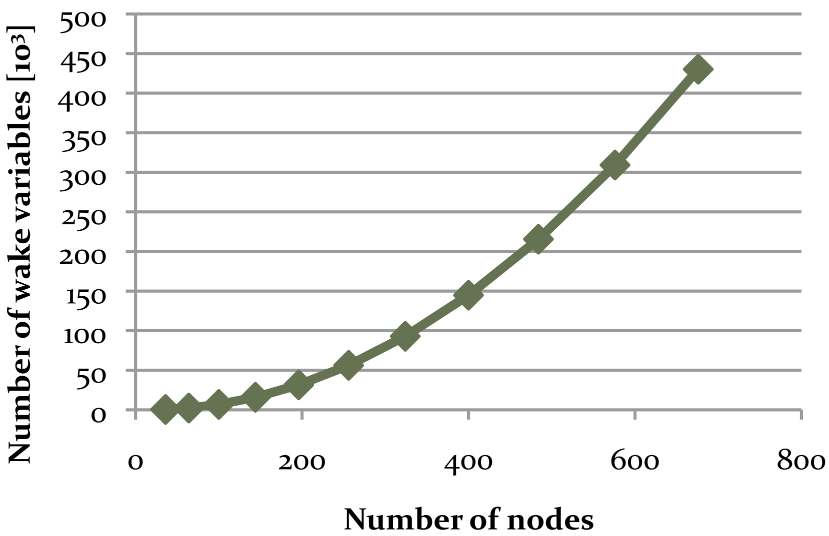


Figure 5.15: **Number of wakes variables as a function of the number of nodes.** The numbers correspond to the site areas in Figure 5.14.

integer problem found so far and the currently lowest upper bound of the continuous problem, is highly increased. This is due to the wake effects, not being present in the LP relaxation, being of greater importance for the number of wind turbines placed. Performance of the profit model can be viewed in Figure 5.16 as well as in Figure 5.6 for the verification example.

Optimization of the production problem without any wakes is merely the circle packaging problem for maximizing the production. This problem is solved faster if the individual production in each point is replaced by a unit value, but will then only represent the maximization of the number of wind turbines and not their optimal placement. The possibility of varying positions when the maximum number of wind turbines are placed is however highly restricted. The problem of placing almost the maximum possible number of wind turbines is harder to solve than placing the same number freely in an larger area, while a configuration where there exists only a few ways of placing the maximum number of wind turbines could be solved quite quickly. The performance of this problem relative to the number of nodes with and without the number of wind turbines restricted is shown in figures 5.17 and 5.18.

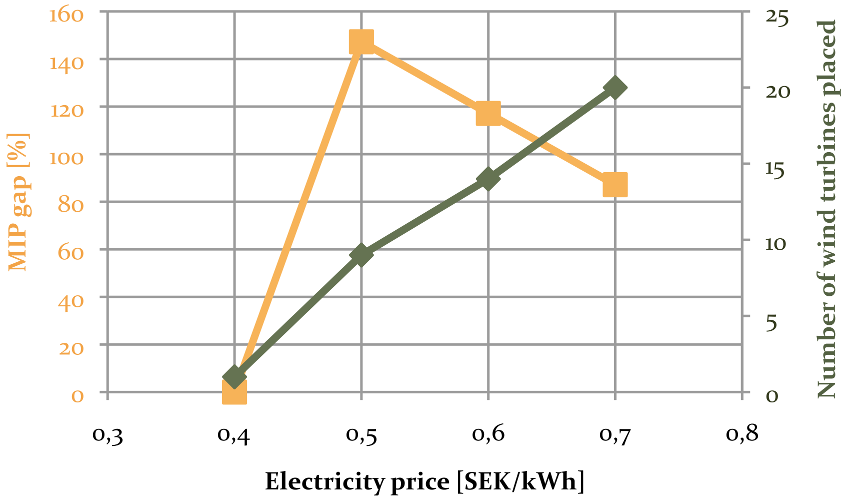


Figure 5.16: ***Production model.* The objective function expressed as the revenue instead of production**. The payback period and construction costs are held fixed, whereas the price of electricity varies. The values are obtained for a 2.89 $(\mathrm{km})^2$ area housing 324 nodes in a 100m grid run for 30 minutes. The maximum possible number of wind turbines of this configuration is 25, whereas a low electricity price will cause the wind turbines to be positioned further apart and thus fewer will be placed.

Figure 5.17: ***Production model.*** **Solution time as a function of the number of nodes and of the size of the area,** when the number of wind turbines is limited to 40 and the wake effects are not considered. Finding feasible locations for up to 40 wind turbines in the smaller areas means packing the circles tightly. The grid size is 50m and the minimum separation distance is 400m.

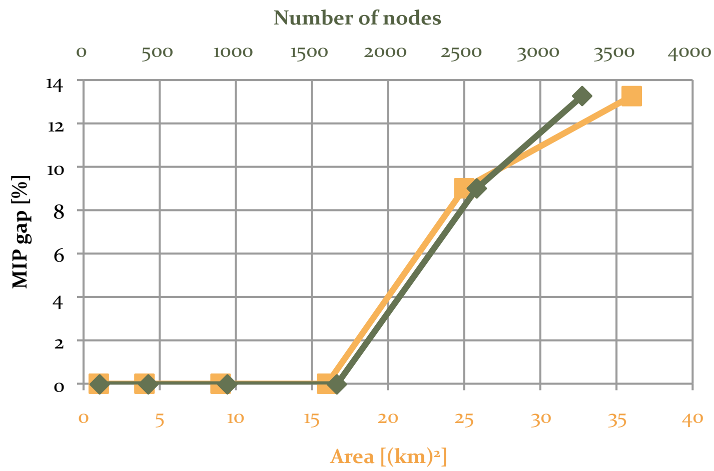


Figure 5.18: ***Production model.*** **The MIP-gap as a function of the number of nodes and of the size of the area,** when no limit on the number of wind turbines is used and wake effects are not considered. The grid size is 100m and the minimum separation distance is 400m.

The more number of wind turbines to place, the harder the problem to solve. Figure 5.19 displays the solution time increasing relative to the maximum number of wind turbines allowed in a 25 (km)^2 area.

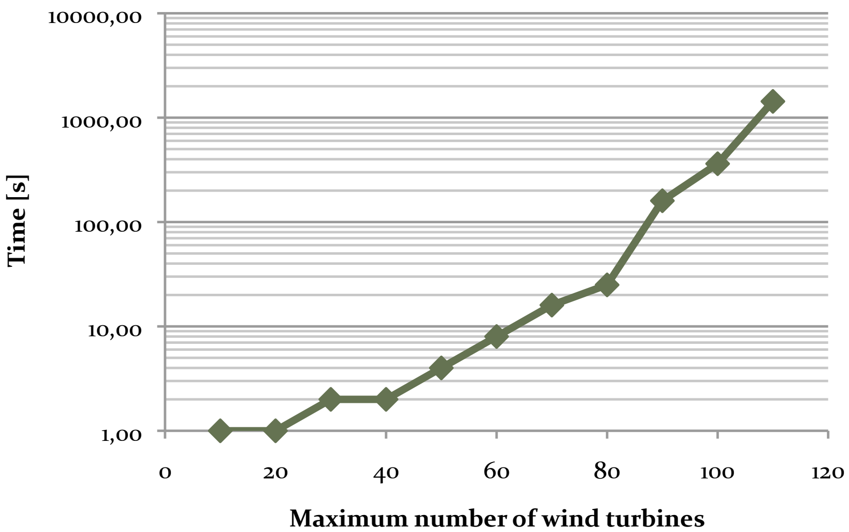


Figure 5.19: ***Production model.* Solution time as a function of the maximum number of wind turbines allowed.** On a 100m grid in a 25 (km)^2 area without the use of wake effects. The optimal value lies between 173 (from the best integer solution) and 188 (from the LP relaxation); this solution was found after three hours of computation.

The Infrastructure model

In the *Infrastructure model* cable and road paths can either be created between nodes containing wind turbines, or make use of surrounding nodes. Figures 5.20 and 5.21 present examples of both. Here, computer memory limitations highly restricted the possible problem sizes.

The combined production and infrastructure model

The simulations performed with the combined model were made — due to memory limitations — only for really small problem sizes, not interesting

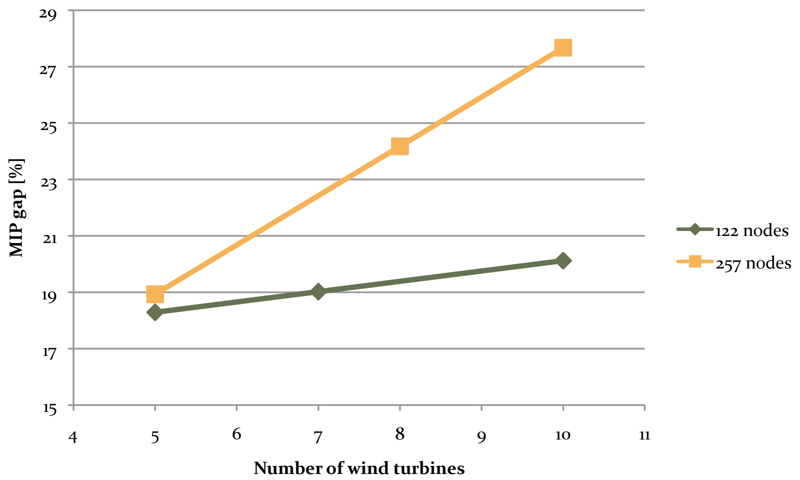


Figure 5.20: ***Infrastructure model.*** **MIP-gap when using the full grid of nodes for two problem sizes.** The number of paths per node is reduced to one per sixteen directions. The simulations use variables for two different cable capacities.

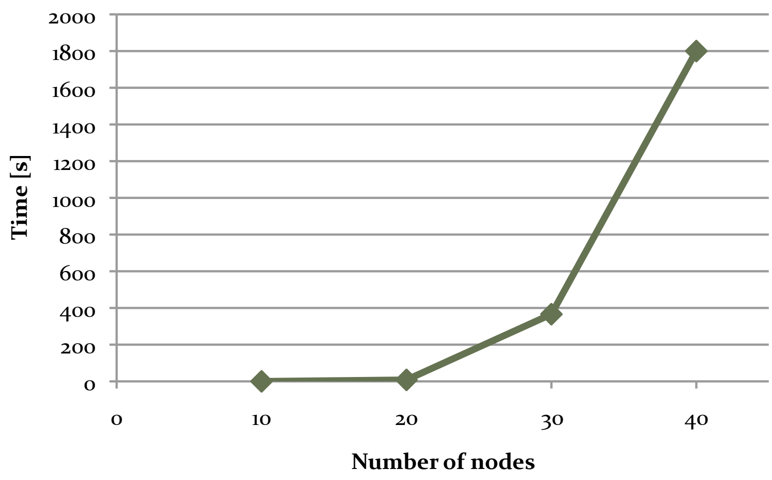


Figure 5.21: ***Infrastructure model.*** **Solution time as a function of the number of wind turbines with only the wind turbines as nodes.** The MIP-gap in the 40 nodes optimization is 4.89%. The simulations use variables for two different cable capacities.

enough for this report. It was, however, apparent that due to the small costs of roads and cables compared to the profit from production, improvements in the tree costs had a little effect on the MIP-gap in the optimization process.

Chapter 6

Conclusion and recommendations

The conclusions listed in this chapter will mainly regard the simulations and verifications made with the *Production* and *Infrastructure models*, whereas more of the applications and advantages of the models will be referred to in the discussion.

6.1 Conclusion

The mixed integer linear models derived in this thesis, the *Wind Farm Layout Production Optimization model* and the *Wind Farm Layout Infrastructure Optimization model*, can handle not only production optimization and a correct combination of noise propagation from multiple wind turbines, but also the minimization of costs of roads, interconnecting cables and placement of the transformer station.

The *Production model* places wind turbines, with respect to maximizing production, in an area. It does not miss feasible and good layouts as compared to the heuristic algorithm used by the commercial software (see Section 5.1.2). A high resolution is of more importance for finding the optimal production locations than for minimizing the wake effects. The number of continuous wake effect variables grows by the square of the number of possible locations; it can, however, be approximated to grow proportionally to this number by a suggested approximation of wakes between sectors. Due to the general formulation of sound pressure constraints, the model can incorporate advanced sound propagation models dependent on wind directions and wind intensities. The linearity of the model may impose a restriction when combining wake effects from several wind turbines, since other means of combination than linear superposition often are proposed. This issue can to some extent can be compensated by approximations suggested, but this will increase the problem size. Though not modeling combination of wake effects the same way as the commercial optimizer, *WindPRO 2.6 Optimize*, and the total production calculation software, *WindPRO 2.6 PARK*, used for finding the total production, the *Production model* yielded equal or better solutions than the commercial software in the verification examples when placing the same number of wind turbines. When optimizing the total production of the examples in this thesis, the objective of both the *Production model* and the software *Optimize*, without any limitations on the number of wind turbines, the *Production model* was capable of locating positions for about 50% more installed production capacity than the comparison software. This resulted in an about 40% higher total production.

The Infrastructure model is capable of connecting the wind turbines with cables and roads in a tree structure and place the transformer station in the most beneficial location. This is done either by cables directly from one wind turbine to another — with splits only occurring at these locations — or by the use of additional nodes or a grid for the cable paths to use. The latter has the preferred ability of allowing cheaper combinations but at the cost of many more nodes and edges, and thus, a higher complexity. With fixed wind turbine positions the problem would benefit greatly from redefining the set of nodes into fewer and better placed nodes. When using a grid of nodes the growth of the number of edges can, however, be approximated to be proportional to the number of nodes without any great loss of optimal objective value. This is done by only using the shortest edges in, for example, sixteen directions. Different cable diameters, double cables, and the power losses in a cable can all be modeled by the use of different sets of edge variables. The implemented tree structure resembles the reality when the costs of combining flows are low. If this is not the case, additional formulations suggested are required to be able to benefit from multiple cables along the same road.

The combined problem of the *Production model* and the *Infrastructure model* will, however, be large, especially since approximations in order to form a smaller tree structure cannot be applied here. For this combination to be of value, the costs of cables and roads should not be just a small fraction of the revenue of the electricity production. This is because changing the road and cable costs otherwise will not have much effect on the position of the wind turbines. A two-stage approach is therefore more likely to be applicable.

The mixed integer linear programs developed in this thesis are capable of addressing the defined problems of wind farm layout design. The programs can handle problems of realistic size, but would gain from further improvements. Some of these possible improvements are suggested in the thesis. The methods used seem to be applicable for use in the industry and compare well to the commercial software in the verification examples.

6.2 Discussion

Now, is it possible to gain from the use of computer optimization when designing a wind farm layout? Moreover, is linear programming a prominent tool for getting more bang for the buck? The answer to both of these questions would, according to the simulations and discussions in this thesis, be yes. There are of course many aspects to consider when designing a wind farm layout, and the human interaction in the workflow is not about to be replaced by simply pressing an optimization button. However, there seems to be room for rather large technology improvements regarding the use of automated optimization.

The integer linear programming models presented in this thesis do provide a framework for successfully solving both the problem of maximizing production in an area and the problem of minimizing costs of connecting the wind turbines by roads and power lines. There are issues restricting the possible site area or resolution of the placement grid, since the optimization problem grows rapidly in complexity with the number of possible locations, but possible techniques for addressing some of these issues are presented in Section 4.2.7 however not implemented. The integer linear programming is also restricted by not being able to properly model non-linear behaviour. The combination of wake effects is proposed to be better represented by a sum of squares combination than by linear superposition. Further, the power loss in a cable is proportional to the square of the current, which is not near linear at all. These problems can partially be corrected by approximations through added variables, at the cost of a larger optimization problem.

The comparison between the *Production model* and the commercial heuristic software *Optimize* was made for only a verification example with shifting grid resolutions, thus the outcome does not have to be general. However, it illustrates a serious drawback of an heuristic algorithm that places one wind turbine at a time, and which cannot adjust all of them simultaneously — it is easy to block combinations of locations resulting in non-optimal solutions. The integer linear programming approach used in this thesis not only gives a quality measure during the solution process — the *MIP-gap* — but is also generally guaranteed to converge to an optimal solution of the problem defined. Though it may take a tremendous time and may also require a very large amount of computer memory. Typically, the optimization process will be terminated after a certain time or when reaching a certain MIP-gap.

The *Infrastructure model* formulation was made under the assumption of the combination of flows being rather cheap. The implemented model therefore cannot benefit from placing cables with flows in opposite directions along the same road, and a closer resemblance to reality is obtained when instead using the additional formulations suggested. Using additional nodes outside the fixed wind turbine locations is necessary, forming a so called Steiner tree, but a full grid of nodes is too heavy to use and the added nodes should be placed in more motivated locations (as introduced in Section 4.3.7).

Application scenarios of the models presented

Differences between the integer linear programming approach, used in the *Production model*, and the heuristic method implemented in *Optimize* are concluded for the scenarios described below.

If wanting to yield as much production output as possible from a given area, the *Production model* will have a major advantage by not accidentally blocking solutions with many placed wind turbines which is done by *Optimize*, as described in Section 5.1.2. The *Production model* also has the advantage of offering other optimization objectives than just total production. One could maximize profit using a given payback period and a given electricity price. This could, for example, make it more beneficial to build only ten wind turbines rather than eleven, if the total production gain of the last added turbine is merely a fraction of a percent. This scenario of tightly placed wind turbines is perhaps more likely to occur in smaller wind farms than in larger ones, mainly due to the major losses caused by the wake effects when placing many wind turbines really tight. If designing a smaller park, an increased grid resolution will not yield an enormous problem size and approximations of other ways of combining multiple wake effects could be performed without the problem getting unreasonably large.

However, if the goal is not to produce as much output as possible from a given project area, but rather to place a given number of wind turbines as good as possible in a large area, the disadvantages of the *Optimize* heuristic placement algorithm does not need to be as crucial. For this scenario, the ability to calculate the combination of wake effects correctly may be used to judge which method manages the best. The *Production model* would have to be made more complex by compensating for the combination of wake effects, if a sum of squares method is decided to have better relevance. Fewer and more separated wind turbines does, however, reduce the error of combining the wake effects linearly.

When sound levels in surrounding areas are to be considered, the *Production model* has the advantage of being able to superimpose the sound pressures from several wind turbines correctly, whereas *Optimize* only adopts a minimum distance to these areas. The *Production model* can also handle more advanced sound propagation models, which depend on the wind speed and direction as well as other meteorological factors.

Now if the project at hand is one where wind turbines are only allowed to be placed at certain given locations, for example, at shallow sections and reefs in an offshore project, the *Production model* avoids the risk of blocking locations as mentioned above. The exact locations are also easily entered into the model. Such a problem is, due to the limited number of possible locations, generally small and additional variables approximating non-linear combination of wake effects can be added without solution times increasing to unreasonable levels.

Other types of offshore scenarios, where placement is done more freely but foundation costs vary greatly with the depth and the type of sea bed, can also be run in the *Production model* by optimizing over revenue.

The software *Optimize* cannot, according to its manual, perform its normal optimization algorithm on an offshore project or on any flat area at all. This is because it needs local variations in wind characteristics not to block too many solutions. Flat areas do not pose a problem for the *Production model*. To be able to handle flat areas, *Optimize* has a mode where it places wind turbines in a regular grid and then tries varying parameters such as row spacing to find better solutions. Due to these defined straight or curved rows, this option is useful for creating visually appealing layouts. This approach would be possible and fast to implement also in the *Production model* by iterating the optimization with changing row parameter values. Due to the highly restricted number of possible locations in each iteration, the problem can be solved quickly also with extra variables approximating non-linear combination of wakes added.

On the comparison with human optimization

Comparing an optimization algorithm to a human, placing the wind turbines continuously, the need for a discrete grid of possible locations is a drawback, but a small one. This means restricting the possibility of placing a wind turbine in the exact right windy spot, and reduces the ability of packing wind turbines tightly. When using a grid fine enough, as suggested in this thesis, possibly refined on and around especially prominent locations, the derived models would however be able to place the wind turbines efficiently.

The benefits of a computer software being able to calculate the production and the wake effects in a complex terrain with locally varying conditions, and not, for example, just placing wind turbines five or seven rotor diameters apart, highly seem to outweigh the problem of discrete placement. Combining this information advantage with the variety in using different wind turbine models or noise levels, a proper implementation of sound level restrictions and the need of connecting the wind turbines with roads and the right cables, could make a computer optimization model hard to beat. In addition to this, valuable time could be saved in the actual layout process. There is thus great potential in using mixed integer linear programming in the pursuit of production and financial gain, and the techniques should definitely be further developed.

6.3 Recommendation for future work

The suggested sector based wake formulation should be implemented and evaluated at different sector sizes. Approximations of the combination of wake effects should be applied to be able to handle other than linear superposition of multiple wakes. A wake originating from a far distant wind turbine will have a negative impact on another wind turbines production, but it will not have a major effect on the positioning of the same. It would be interesting to try and adopt this behavior in the model, making the positioning depend on mainly close wind turbines. Regarding the *Infrastructure model* the models with multiple nodes in each location and the separation of road and cable variables should definitely be implemented to resemble a realistic problem. Whether the tree nodes should be rearranged into better positions or a formulation similar to the sector based wake formulation recommended depends on whether the locations of the wind turbines are fixed, and both of these complexity reducing methods ought to be tested.

This thesis has not focused on solution methods, and apart from the problem specific approximations utilized and suggested, more general means of relaxations and time reducing techniques should be considered. It may be possible that Lagrangian relaxation of the constraints relating the wind turbine positions to the tree structure would ease the solving of the combined problem of the *Production model* and the *Infrastructure model*. These constraints are however quite complex and other relaxations may prove to be of greater importance. The choice of branching strategy in the branch and bound method could also be considered, perhaps as referred to in [5].

Bibliography

[1] Boverket, Energimyndigheten, and Naturvårdsverket. *Ljud från vindkraftverk, Rapport 6241*. Naturvårdsverket, Stockholm, Sweden, 2001.

[2] D. Cieslik. *SHORTEST CONNECTIVITY An Introduction with Applications in Phylogeny*. Springer Science+Business Media, Inc, New York, USA, 2005.

[3] R. Courant and H. Robbins. *What is mathematics? An elementary approach to ideas and methods*. Oxford University Press, London, Great Britain, 1941.

[4] S. Donovan. Wind farm optimization. In *Proceedings of the 40th Annual Conference of the Operations Research Society*, Wellington, 2005.

[5] S. Donovan. An improved mixed integer programming model for wind farm layout optimization. In *Proceedings of the 41th Annual Conference of the Operations Research Society*, Wellington, 2006.

[6] S. Donovan, G. Nates, H. Waterer, and R. Archer. Mixed integer programming models for wind farm design. In *Slides used at MIP 2008 Workshop on Mixed Integer Programming*, Columbia University, New York City, 2008.

[7] C.N. Elkinton, J.F. Manwell, and J.G. McGowan. Offshore wind farm layout optimization (OWFLO) project: An introduction. In *Proceedings of Copenhagen offshore wind conference 2005*, Copenhagen, Denmark, 2005.

[8] EMD International A/S. *WindPRO 2.6 Help*.

[9] S. Frandsen, L. Chacón, A Crespo, P. Enevoldsen, R. Gómez-Elvira, Hernández J, J.Højstrup, F. Manuel, K. Thomsen, and P. Sørensen.

Measurements on and modelling of offshore wind farms. Technical Report Risø-R-903(EN), Risø, 1996.

[10] E. Friedman. Circles in circles. http://www2.stetson.edu/~efriedma/cirincir/. 2010-02-10.

[11] E. Friedman. Circles in squares. http://www2.stetson.edu/~efriedma/cirinsqu/. 2010-02-10.

[12] S.A. Grady, M.Y. Hussaini, and M.M. Abdullah. Placement of wind turbines using genetic algorithms. *Renewable Energy*, 30(2):259–270, 2005.

[13] N.O. Jensen. A note on wind generator interaction. Technical report, Risø, 1983.

[14] I. Katić, J. Højstrup, and N.O. Jensen. A simple model for cluster efficiency. In *European Wind Energy Association Conference and Exhibition*, pages 407–10, Rome, Italy, 1986.

[15] A. Kusiak and Z. Song. Design of wind farm layout for maximum wind energy capture. *Renewable Energy*, 35(3):685–694, 2010.

[16] J. Lundgren, M. Rönnqvist, and P. Värbrand. *Optimization*. Studentlitteratur, Lund, Sweden, 2010.

[17] G. Mosetti, C. Poloni, and B. Diviacco. Optimization of wind turbine positioning in large windfarms by means of a genetic algorithm. *Journal of Wind Engineering and Industrial Aerodynamics*, 51(1):105–116, 1994.

[18] Risø National Laboratory. *WAsP 8 Help Facility*.

[19] R. Aghabi Rivas. Optimization of offshore wind farm layouts. In *Slides used at MIP 2008 Workshop on Mixed Integer Programming*, Columbia University, New York City, 2008.

[20] Hanif D. Sherali and J. Cole Smith. A class o web-based facets for the generalized vertex packing problem. *Discrete Applied Mathematics*, 146(3):273–286, 2005.

[21] M. Lybech Thøgersen. *WindPRO/PARK, Introduction to Wind Turbine Modelling and Wake Generated Turbulence*. EMD International A/S, Aalborg, Denmark, 2005.

[22] L.J. Vermeer, J.N. Sørensen, and A. Crespo. Wind turbine wake aerodynamics. *Progress in Aerospace Sciences*, 39(6–7):467–510, 2003.

[23] C. Wan, J. Wang, G. Yang, and X. Zhang. Optimal siting of wind turbines using real-coded genetic algorithms. In *Proceedings of European Wind Energy Association Conference and Exhibition*, Marseille, France, 2009.

[24] T. Wizelius. *Vindkraft i teori och praktik, 2:a uppl.* Studentlitteratur, Lund, Sweden, 2007.