Modelling and Optimisation of Renewable Energy Systems
MODELLING AND OPTIMISATION
OF
RENEWABLE ENERGY SYSTEMS

By

Jeanne Andersen

Supervisor: Kim Allan Andersen

A PhD thesis submitted to the School of Business and Social Sciences, Aarhus University, in partial fulfillment of the PhD degree in Economics and Business

March 2015
Summary

This thesis consists of three chapters, each of which constitutes a self-contained research paper. The three papers are all related to the modelling of optimisation problems within energy systems.

In the first two papers, we look at electricity system operations within the hour, where supply and demand of electricity have to be balanced. In the papers, we present two proactive models. Based on forecasted imbalances between supply and demand, the models aim at reducing intra-hour balancing cost by optimally adjusting the production level before real-time operation by utilising manual reserves. In both papers, we see that taking a proactive approach entails substantially lower cost than letting automatic reserves handle all the imbalances real-time.

In the first paper, we propose a mixed integer deterministic model for the intra-hour balancing problem. We investigate the Danish system, where wind power is highly used and a major source for intra-hour imbalances between supply and demand. We find that balancing cost do not outweigh the benefits of the fluctuating inexpensive wind power.

The second paper is an extension of the first paper. Here, we present a two-stage stochastic mixed integer model for the intra-hour balancing problem. The model captures the uncertainty in the forecasts of wind power production by generating scenarios for the prediction errors. We compare the stochastic and the deterministic solutions to the solution of perfect foresight and find that prediction errors entail huge balancing cost. Furthermore, we see that the stochastic solution incorporates a buffer when activating manual reserves compared to the deterministic solution. Incorporation of this buffer results in higher expected cost, but the actual cost incurred is lower in most of the investigated cases.
In the last paper, we look at an environmental supplement or alternative to wind power. Here, we present a mixed integer linear model, which can be used to design the supply chain network for biomass. Improving the logistics of biomass-to-energy supply chains is a requirement for enabling biomass as an economical and environmentally sound additional long-term source of energy. The design of the supply chain network is important for achieving efficiency in logistics operations, and it involves long-term decisions regarding transport flows, capacities, as well as the number and types of facilities used in the network for providing biomass and transforming it into energy. In our case the particular biomass is straw that may be converted into compressed briquettes. In addition to other models for biomass supply chain design, we optimise the network design together with the truck routes. The paper is a preliminary version as we have not yet received all the necessary data needed to run the model on our full Danish case study (design of the supply chain network and transportation routes for the new bioethanol plant at Maabjerg). For now, we propose a Lagrangian relaxation by variable splitting as the solution method. On small test instances, this method provides reasonable gaps between upper and lower bounds on the objective function value. Standard solvers as CPLEX are not able to solve some of these small instances, which indicate a need for a tailor-made algorithm to solve the problem.
Resumé

Denne ph.d.-afhandling består af tre uafhængige artikler i hvert sit kapitel, som alle er relatert til modelleringen af optimeringsproblemer inden for energisystemer.

I de to første artikler ser vi på elsystemets drift inden for den enkelte time, hvor udbud og efterspørgsel af elektricitet skal balanceres. Vi præsenterer to proaktive optimeringsmodeller, som, baseret på forudsigelser af ubalancer mellem udbud og efterspørgsel, prøver at reducere balanceringsomkostningerne ved at justere produktionsniveauet med manuelle reserver før selve drifttidspunktet. Vi ser i begge artikler, at den proaktive tilgang resulterer i betydeligt lavere omkostninger end hvis vi venter med at håndtere ubalancerne til selve drifttidspunktet, hvor automatiske reserver benyttes.

Vi foreslår i den første artikel en deterministisk model til at beskrive balanceringsproblemet inden for timen. Vindenergi er en kæmpe kilde til ubalancer inden for timen mellem udbud og efterspørgsel af elektricitet. I den første artikel undersøger vi derfor effekten balancemæssigt af vindenergi i det danske system, hvor det dækker en stor del af elforbrug. Det viser sig, at selvom fluktuerende vindenergi skaber større ubalancer i elsystemet jo mere vind der inkluderes, så er fordelene ved den billige vindkraft rent omkostningsmæssigt langt større end ulemperne.

Den anden artikel ligger i forlængelse af den første og her præsenterer vi en stokastisk model til at beskrive balanceringsproblemet inden for timen. Modellen fanger usikkerheden i forudsigelserne af produktionen af vindenergi ved at generere scenarier for fejlen i vindenergi-prognoserne. Vi sammenligner den stokastiske løsning og den deterministiske løsning med løsningen givet ved perfekt fremsynethed og ser, at prognosefejlene medfører betydelige balanceringsomkostninger. Endvidere ser vi, at den stokastiske løsning, set i forhold til den determi-
nistiske løsning, inkorporerer en buffer, når den aktiverer manuelle reserver. Inkorporeringen af denne buffer resulterer i højere forventede omkostninger, men de aktuelle omkostninger er til gengæld lavere i de fleste af de undersøgte test instanser.

Preface

I have prepared this thesis during my enrollment as a PhD student at the Department of Economics and Business, Aarhus University, in the period from January 2011 to March 2015. My PhD project has been funded by CFEM (Center for Foundations of Electronic Markets) which in turn is funded by the Danish Council of Strategic Research.

During my PhD studies, I had two small internships and a temporary employment at the Danish transmission system operator, Energinet.dk. The internships were in 2011 and the temporary employment was in the period from September 2012 to February 2013. Even though my employment at Energinet.dk was not directly related to my PhD studies, I learned a lot about electricity systems and markets which I think has been a huge asset for me throughout the rest of my studies. However, it has also been a challenge to combine knowledge of real world problems and approaches with the academic discipline. It has especially been difficult to find the balance where a real world problem was described good enough, while an academic approach still was maintained.

On the personal level I was fortunate to become mother to a wonderful boy in late 2013. Even though it has not been easy to manage both motherhood and writing this thesis, it has also brought structure and happiness into the writing process; it has just been lovely to have these small joyful moments of pause that naturally occur when spending time with a small child.

Acknowledgement

First and foremost, I want to express my gratitude to the Department of Economics and Business for providing an excellent research environment and to CFEM for providing financial
support.

I want to thank my supervisor Kim Allan Andersen for encouraging me to write a PhD, and for being a moral support throughout the years. Even though the project did not follow the original plan, he never doubted me, and he always protected my interests for which I am truly grateful.

I want to give a special thanks to my co-supervisor, Nina Kildegaard Detlefsen. She has always been there when I needed support as well as when I just needed to get frustrations out in the open. Furthermore, we have had many fruitful academic discussions which definitely has improved the content of this PhD thesis.

I want to thank my dear friend and collaborator, Ditte Mølgård Heide-Jørgensen, who is a co-author on my first paper. Even though we have had some challenges, pregnancies among other things, we have kept our friendship and learned a lot from each other. It has been a privilege to work with her!

I also want to give a special thanks to Trine Krogh Boomsma, who is also a co-author on my first paper. She did not have an easy task of supervising the paper, since she where not involved from the beginning, but she never gave up on us. I also want to thank her for all the work she has put in to commenting on every thing I have sent her - even though it came back more red than black I am truly grateful, and I think it has made me a better writer.

From March to June 2012, I visited Asgeir Tomasgard at the Department of Industrial Economics and Technology Management at NTNU, Trondheim. It was an interesting period, where I had the opportunity to be in a research environment with people who were interested in problems similar to mine. I am truly grateful to Asgeir for hosting my stay and for our collaboration on my second paper. During my stay in Norway, Asgeir introduced me to Michal Kaut at SINTEF, Trondheim. I want to thank Michal for our collaboration on my second paper.

I also want to thank my co-authors on the third paper, Andreas Klose and Claus Aage Grøn Sørensen. I think it has been an interesting project and a very good collaboration.

Another person I want to give a special thank is Stephan Wöllner, who I met during my stays at Energinet.dk. Stephan has been a good colleague and he has made a lot of effort in making SimBa, a model from Energinet.dk, deliver the necessary data in the correct format, such that it could be used in my second project. I also want to express my gratitude to En-
erginet.dk for providing data, lending models, for their hospitality during my two internships.

Also, I want to thank my colleagues at CORAL and my fellow PhD students. I specially want to thank Camilla Pisani, David Sloth Pedersen, Lene Gilje Justesen, Lukas Bach, Maria Elbek Andersen, Marie Herly, Reza Pourmoayed, Samira Mirzaei, Sune Lauth Gadegaard, Tine Louise Mundbjerg Eriksen, Tue Rauff Lind Christensen, and Viktoryia Buhayenko. The last four years would not have been the same without you guys!

Last but not least, I want to give a special thank to my family and especially my boyfriend, Morten Leander Petersen, for believing in me and for all the support he has given me. I can only imagine how it must have been for him throughout the last couple of years, and especially the last couple of months, while I have been writing this thesis. I really appreciate all his support and his efforts to make my life easier. I love him with all of my heart!
# Contents

Summary ........................................................................................................... i

Resumé ............................................................................................................ iii

Preface ................................................................................................................ v

Acknowledgement ............................................................................................. v

Introduction ....................................................................................................... xi

Electricity markets and systems ....................................................................... xiii

Biomass .............................................................................................................. xvi

1 Short-term balancing of supply and demand in an electricity system: Forecasting and scheduling ................................................................. 1

1.1 Introduction .................................................................................................. 2

1.2 Related literature ........................................................................................ 5

1.3 From hour-by-hour to intra-hour scheduling .............................................. 7

1.4 Wind power forecasts ................................................................................ 9

1.5 The intra-hour model .................................................................................. 13

1.6 Case study .................................................................................................. 21

1.7 Results and discussion ............................................................................... 25

1.8 Conclusion .................................................................................................. 32

Appendices ......................................................................................................... 33

1.A Wind power forecasts: Parameters and correlations ................................. 33
Introduction

An energy system can be thought of as a network consisting of production from energy sources, storage, transmission, distribution, and consumption of energy. The most important types of energy in an energy system are found within the transportation, heating, and electricity sectors. These sectors often rely on traditional fossil fuel based energy sources such as coal, oil, and gas. However, with the alertness of climate changes in the world, it is vital to replace fossil fuels with significant amounts of renewables, such as solar, hydro, and wind power, in the energy systems.

The European Union directive 2009/28/EC states that renewable energy should comprise a 20% share of the total energy production in EU in 2020. In order to reach this overall target, each European member state has been given mandatory national targets. The national target for Denmark is 30%. In addition, the Danish Government aims to have a share of 35% renewable energy in the total energy system, and 50% of the electricity consumption covered by wind power in 2020. Furthermore, it aims at 100% renewable energy by 2050 (Danish Ministry of Climate, Energy and Building (2012)). Denmark will accommodate these targets and reduce emissions by, among other things, integrating considerable amounts of wind power in the electricity system, and by utilising biomass in the heating and transportation sectors.

Including the required amount of renewables in the Danish energy system will be challenging. In this thesis, we address some of these challenges and present models that can be used to analyse them. For instance, integration of large amounts of wind power can challenge the balancing of supply and demand of electricity at the time of operation due to the fluctuating and unpredictable nature of wind. Therefore, in the first two chapters, we look further into wind power and its short-term effects on electricity systems and use Denmark as a case study. These
first two chapters show that wind power highly impact the short-term balancing of electricity systems with a large amount of installed wind capacity due to the unpredictable nature of wind. Furthermore, we see that acting on predicted imbalances between supply and demand before real-time operation can reduce balancing cost for the system operators, both by being proactive and also by taking uncertainty of wind power predictions into account.

As a contrast to the fluctuating renewables, biomass is one of the few renewable energy sources that can actually be stored and generate different types of energy on demand. Therefore, biomass is a good supplement to uncertain wind power. As a supplement to the wind, biomass can be utilised in power plants when the wind is not blowing and for providing balancing reserves. However, the amount of biomass available is limited by its growth rate, and in order to exploit the full potential of biomass it is essential to obtain the biomass from a local level. In the last chapter, we look further into the utilisation of biomass and investigate the network design of biomass supply chains. The network design of supply chains for biomass is very important, since the utilisation is extremely vulnerable to the high cost arising when setting up the supply chains. Overcoming these high logistic costs can maybe lead to an expansion in the utilisation of biomass in the future.

The main contribution of this thesis is the development of new optimisation models. Each of the models in the chapters contributes to enlighten further integration of wind power or biomass in energy systems similar to the Danish energy system. The first two chapters can be helpful to system operators when analysing electricity systems. They can use the models to get a better understanding of how small changes will influence the system. The last chapter can be helpful to companies that want to use biomass as input to their production, e.g. for producing biofuel or heat. These companies can use the model to investigate the supply chain of biomass and use it as support for decision making.

In order to establish a fundamental basis for the reader of this thesis, we will in the following give a short introduction to electricity markets and systems and how they are influenced by wind power, and finally, we will briefly discuss biomass utilisation in heating and transportation sectors.
Electricity markets and systems

Electricity has a major impact on the social and economic developments of nations around the world since it is an essential ingredient of modern society. The modern society depends on constant accessibility of this commodity in order to maintain the present way of living. We are used to a constant supply of electricity and as consumers we believe that we are entitled to buy (consume) this commodity whenever we want to. Electricity can be traded like all other commodities which is challenging since it cannot be stored in any feasible way with present technology, not even with batteries which are too expensive in the scale needed for electricity systems. Since the electricity cannot be stored, it has to be balanced such that supply equals consumption in every micro second of the day. This makes it a unique and interesting commodity with its very own characteristics.

When balancing the system, the philosophy so far has been to let the production follow the consumption. This philosophy has been possible since the energy sources for the production traditionally have been fossil fuels which can be stored and employed on demand. However, with the climate changes we are looking at a future with a huge reorganisation of the energy systems, and thereby also the electricity systems, where utilisation of renewable energy sources is in focus. In Northern Europe, wind power is one of the most promising large-scale renewables that can replace parts of the conventional supply. However, if large amounts of wind power are installed in the electricity systems, the characteristics of the electricity production will change and thereby the historical philosophy will be challenged. New methods to obtain balance between supply and demand have to be explored. One such method can be to let the consumption follow the supply as attempted through smart grid initiatives.

In smart grid initiatives, intelligent technologies for the consumption can react based on information about the electricity system and prices in an automatic way to improve the economics and reliability of the system. Some initiatives look at how to reward households if they consume when the consumption is low and the production high, for instance, during the night. One way is to implement technologies in electricity devices that can react to the prices and automatically turn on when the price is low. Since normally, the price is low when there are high production and low consumption, these technologies will help direct the consumption to fol-
low the supply. However, many smart grid incentives are not yet incorporated in the electricity markets even though the technology is ready. It takes time to adjust the market structure to the developed technologies and to change the mind-set of the consumers.

Today, electricity is traded on many different markets. In the Nordic countries, most of the electricity is traded through the platform Nord Pool Spot. At Nord Pool Spot, they first have a day-ahead market, Elspot, which is sometimes also called a spot market. As the name suggests, the day-ahead market opens the day before the energy has to be delivered. After gate closure of the day-ahead market, the intraday market, Elbas, opens, and it closes just before the hour of operation (where the electricity has to be delivered). The intraday market is used for buying or selling electricity if unforeseen events occur after the gate closure of the day-ahead markets. In this manner, the participants can reduce their expected imbalances for which the system operator otherwise holds them responsible. In the case of wind power, for instance, wind forecasts are more accurate closer to the hour of operation, which could be a reason for trading on the intraday market. In both the day-ahead and intraday markets, the participants are either buyers or sellers of electricity.

After the closure of the intraday market, it is no longer possible for the participants to change their positions and it is up to the system operators to maintain the balance between supply and demand. If, for some reason, one of the participants in either the day-ahead or intraday market does not fulfil their obligations by deliver or consume the promised amount of electricity, the system operator either has to activate or deactivate reserved electricity in order to maintain balance. When the system operators activate or deactivate electricity, it is referred to as activating reserves.

The following distinctions are made when talking about reserves: primary reserves, secondary reserves, and tertiary reserve. The first two belong to the category of frequency control, and the last one is a slower manual reserve. Primary reserves are automatically activated when the frequency deviates from the set point value by a predetermined amount in order to maintain the balance between supply and demand. In the Nordic countries, primary reserves have to be fully operational within 30 seconds after a disturbance has occurred. In places where secondary reserves are used, it is used to restore the primary reserves and is also automatically
activated. The tertiary reserves have slower response times and are manually activated. In the Nordic countries, the tertiary reserves have to be fully activated within 15 minutes after activation. It restores the primary and secondary reserves. In the Nordic countries, the system operators have a common platform called NOIS for buying the tertiary reserves. The NOIS system belongs to the cross-border reserve trading category and is a list containing all up and down regulation offers in the Nordic countries.

When handling imbalances in electricity systems after the gate closure of the intraday market, the normal procedure is first to let the imbalances be handled by the automatic reserves which are then restored with the manual reserves. However, in Denmark, the transmission system operator tries to predict anticipated imbalances in the system based on forecasts for the consumption and wind power as well as production schedules from power plants and transmission lines. Based on the expectation of the system imbalance, they activate the manual reserves in order to minimise the use of the automatic - and more expensive - reserves.

Measurement of electricity

Regarding the production of electricity two things have to be separated: power and energy. Power is measured in megawatt (MW) or gigawatt (GW) and denotes the instant electricity production. The maximum power a given unit can produce is denoted the capacity of the unit and is also measured in watt. Energy, on the other hand, is the production over time and is measured in megawatt hours (MWh) or gigawatt hours (GWh). If a unit with capacity of 1 MW produces power at full capacity during 1 hour it will produce 1 MWh of energy during this hour. However, it is possible to measure energy in other time resolutions, and then the measure denote the average power production throughout a specific time interval.

Wind power

Political visions for more sustainable energy systems have led to a tremendous development of wind power during the last decade. According to Leung and Yang (2012), the global installed wind capacity was 59.1 GW in 2005, in 2010 it was 175 GW, and the expected level for 2015 is 425 GW. Nations with large amounts of installed wind capacity are China, The United States,
Germany, Spain, and India. However, if we also look at the share of the national electricity consumption wind power comprises, Denmark is leading the field. According to the Danish Energy Agency, Denmark had 4.8 GW of installed wind capacity in 2014, and the share of the national electricity consumption that was comprised by the wind power production in 2013 was 32.7%, whereas it was 39.1% in 2014 (Energinet.dk (2015)).

If the growth of wind power continues, it will comprise a substantial part of the electricity system both in Denmark and in many countries around the world. In many years, system operators have operated, as a minimum, according to the N-1 security criterion, where they have enough system reserves within their own control area to at least cover an outage of the largest generator, transmission line, transformer, or reactor. With higher and higher amounts of installed fluctuating renewables, such as wind power, and less controllable production in the systems, the security criterion is under pressure and new ways to estimate the needed level of reserves are necessary. Furthermore, in electricity systems with high amounts of installed wind capacity, it is essential to have well performing wind power forecasting tools to predict the wind power production and thereby reduce the system imbalances. Weber (2010) shows that the total system forecast error asymptotically converges to the wind forecast prediction error as the proportion of wind capacity installed in the system increases. Thereby, a system with a substantial share of wind power depends on the accuracy of the forecast models since their precision directly affects the level of reserves needed to maintain balance. Many system operators or other balance responsible parties use both wind prediction models and meteorological models to forecast the wind power production, since combining models reduces prediction errors (Nielsen et al. (2007)). For an overview of forecasting models, see Lange and Focken (2008) and Foley et al. (2012).

**Biomass**

Biomass used for energy production often originates from plant based material. However, biomass covers essentially both animal and vegetable organic material. The interesting aspect of biomass in energy production is that it takes carbon out of the atmosphere while it is growing and returns it as it is burned, which is the key element for promoting biomass as replacement...
for fossil fuels. In this context, it is important to stress that the time that goes by between harvesting and burning the biomass is essential. Oil consists of fossilised organic materials, but it is so long ago that it absorbed the carbon, which it now releases when it is burned, that it has a negative influence on the emissions. Hence, when defining biomass as a substitute for fossil fuels, the biomass needs to absorb the carbon it releases in the present such that there is no net increase in carbon emissions. If managed properly, biomass can be an environmentally sound and essential counterpart to fluctuating renewables in energy systems. It can be used to produce electricity, heat, and fuel for transportation and it can be stored and thereby utilised on demand, which makes it a flexible commodity.

Even though biomass is expected to play an important role in sustainable energy systems, it is per se not a true renewable resource such as solar and wind since it is not an inexhaustible resource of energy. It is only possible to utilise biomass at the rate of biomass production, e.g., in the form of growing crops, straw, and trees. Furthermore, biomass is sustainable but only if it is managed in a sustainable way, where it is part of a continuous cycle of replanting. In that case, new growth can take carbon out of the atmosphere at the same time as new carbon is released by burning the previous harvest. In addition, in order to evaluate the environmental impact of the biomass, all stages of its life cycle have to be evaluated including the transportation and distribution steps. If not managed carefully, biomass can create huge environmental damages. If harvested in unsustainable rates, biomass increases net carbon emissions, or if it is transported over long distances, it consumes more energy than it produces and again leads to a net increase in carbon emissions. If the fields where the biomass is grown normally is used for food production, it can increase food prices with huge impact on societies.

In Denmark, biomass is one of the environmentally friendly options we have to supplement the fluctuating wind power, and therefore, it is expected to play an essential role in the further development of the heating sector in combined heat and power plants and of the transportation sector.
The heating sector

District heating is the most widespread heating method in Denmark and it is an important element in the Danish energy system, where it has two main advantages. First, heat for the district heating system and electricity for the electricity system are produced in a joint process on combined heat and power plants. Second, the heat can be stored. When electricity is produced without heat on a conventional power plant, it is possible to utilise around 40% of the energy stored in the fuel. The relatively low extraction of the energy is due to not exploiting the heat, which is a by-product of electricity production. If, on the other hand, electricity is jointly produced with heat, the extraction of the energy can be more than 90% since the waste heat from the electricity production can be captured and used for district heating (Danish District Heating Association (2015)).

In a future energy system, where the use of fossil fuels has to be minimised or completely avoided, there are at least two different views on how the heating of buildings will develop (Lund et al. (2010)). One view states that we will have low-energy buildings that do not require any heating. The other view states that this scenario will be far out in the future since the buildings today are too expensive to rebuild, and therefore we need a mix of alternative and renewable resources that can complement each other.

In the second view, the role of district heating and the relationship between the electricity system and the heating system becomes important. An example of how electricity and heating systems can be even more integrated could be installment of large electrified heat pumps in the electricity systems. Then the heat pumps can be used to produce heat when there is a lot of excess electricity. By letting the heat pumps use the excess electricity, we can store the energy as heat and save the fuel otherwise needed to produce heat for the heating sector. Then, when electricity prices are high, the combined heat and power plants can produce heat and electricity to support the electricity system.

Even though biomass has a lower efficiency as fuel in combined heat and power plants than coal, it is necessary to transform these plants to be able to only utilise biomass when we want a more sustainable energy system, especially when the heat and electricity systems are able to support each other. Denmark has one of the biggest biomass fired combined heat and power
plants in the world, called Avedøre 2, and many of the coal fired combined heat and power plants are under consideration for rebuilding such that they can use biomass.

**The transportation sector**

Biofuels, processed from biomass, are suggested to replace fossil fuels in the transportation sector in the future (Antizar-Ladislao and Turrion-Gomez (2008)). One of the products belonging to the category of biofuels is bioethanol. Since cars today can drive on a blend of petrol and small amounts of bioethanol without problems, bioethanol is expected to play an essential part in the reduction of carbon emissions. The hope is to gradually increase the amounts of bioethanol in petrol over time. However, if large amounts of bioethanol are blended with the petrol, most of the conventional petrol engines will have to be adjusted.

Bioethanol is produced as sugar is fermented. The sugar used for the production of bioethanol usually comes from two different sources. First, if the sugar is derived from corn, wheat, or sugar canes and beets, that contain large amounts of starch and sucrose (sugar), the bioethanol belongs to the category of first generation bioethanol. Second, if cellulosic biomass, such as straw and organic waste, is used, the bioethanol belongs to the category of second generation bioethanol. The production of second generation bioethanol requires advanced processing techniques.

The United States and Brazil produces a substantial amount of first generation bioethanol. In the United States, much of the bioethanol is derived from corn, while Brazil produces large amounts of bioethanol using sugar canes. The bioethanol produced in Denmark is second generation bioethanol derived from straw. At the moment it is planned to install a large second generation bioethanol plant at Maabjerg in Denmark. In the last chapter, we will provide additional information about the plant since it will be our case study for the supply chain network design for biomass.
Chapter 1

Short-term balancing of supply and demand in an electricity system: Forecasting and scheduling

Jeanne Andersen · Ditte Mølgård Heide-Jørgensen
Nina Kildegaard Detlefsen · Trine Krogh Boomsma

J. Andersen*: CORAL, Department of Economics and Business, Aarhus University, Fuglesangs Alle 4, 8210 Aarhus V, Denmark

D. M. Heide-Jørgensen* and T. K. Boomsma: Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, 2100 København Ø, Denmark

N. K. Detlefsen: Dansk Fjernvarme, Merkurvej 7, 6000 Horsens, Denmark

* Contributed equally to this work.

This paper has been submitted to “Annals of Operations Research”. The version presented here has been modified such that notation follows the rest of the thesis.
Abstract
Until recently, the modelling of electricity system operations has mainly focused on hour-by-hour management. However, with the introduction of renewable energy sources such as wind energy, fluctuations within the hour result in imbalances between supply and demand which are undetectable with an hourly time resolution. Ramping on production units and transmission lines contribute further to these imbalances. In this paper, we therefore propose a model for optimising electricity system operations within the hour. Taking a social welfare perspective, the model aims at reducing intra-hour cost by optimally activating so-called manual reserves based on forecasted imbalances between supply and demand. Since manual reserves are significantly less expensive than automatic reserves, we expect a considerable reduction in total cost of balancing. In addition to providing guidelines for current electricity system operation, the model can be used for analysing the management of a future electricity system, and in particular the cost of balancing supply and demand with a high penetration of wind energy. We illustrate our model in a realistic Danish case study and investigate the effect of an expected increase in installed wind capacity. We find that the balancing cost do not outweigh the benefits of the inexpensive wind power, and that the savings from activating manual reserves are even larger for the high wind capacity case.

Keywords: OR in energy, Scheduling, Forecasting, Power system balancing

1.1 Introduction
For more than a century, electricity supply has been highly dependent on fossil fuels such as coal and oil, and even today, fossil fuels continue to make up a large part of the global supply chain for electricity. Due to depletion of fossil fuels and looming climate changes, development of sustainable energy, not least electricity, has received considerable attention over the last two decades. Sustainable electricity development concentrates on three major aspects: Savings on
the demand side, efficiency improvements on the supply side, and replacement of fossil fuel based production by various sources of renewable production (Lund (2007)).

With the introduction of renewable production, security of supply has become an issue of utmost importance. Unlike most commodities, electricity cannot be efficiently stored. Furthermore, major imbalances between supply and demand are extremely undesirable, since deficit or excess in production may result in blackouts. In fact, production and consumption have to balance in the very short term such as minute-by-minute or second-by-second basis. However, whereas conventional production can be controlled, renewable production is often intermittent and non-controllable. Therefore, increasing shares of renewables significantly challenge the ability of an electricity system to balance supply and demand.

In addition to the fluctuating nature of intermittent production, several other aspects affect the balancing of supply and demand in the electricity system. Substantial transmission distances between regions, and different ramping speeds to and from a region may result in large imbalances between supply and demand. Ramping restrictions on conventional power units, failures of power units and transmission lines, and fluctuations in the demand for electricity likewise impact the balance of the system.

The monitoring and maintenance of the balance between supply and demand is controlled by System Operators (SOs), who each have their own control area. To maintain the balance in the area, an SO manages so-called reserves. In the situation of deficit in production, the SO increases planned production by activating reserved capacity. In the opposite situation of excess production, the SO deactivates planned production. In general, reserves can be categorised into automatic and manual reserves. Automatic reserves are fast (operational within seconds) and flexible, hence expensive, and used for controlling the system frequency. Manual reserves are slower (operational e.g. within minutes) and less flexible, hence less expensive, and usually activated in order to release automatic reserves. With the integration of renewables, a major problem for the SO is how to manage the system in the most cost efficient way.

Common practice of many SOs for managing supply and demand is indeed a reactive approach in which they first leave the observed imbalance to be handled by the fast automatic reserves, and second let the slower manual reserves take over. Nevertheless, by activating slow
reserves prior to the observation of the imbalance and rather on the basis of an expected imbalance, the SO could utilise slower reserves much more efficiently. In particular, such proactive activation of manual reserves would reduce the need for automatic reserves and thereby the cost of balancing.

This paper establishes a framework for short-term management of electricity system operations, including the balancing of supply and demand by proactive activation of power systems reserves. We consider a system in which wind energy is the renewable source. Still, the methodology applies equally well to many other renewable sources such as solar energy and run-of-river hydro-power. We assume that the associated electricity market includes a day-ahead market for hourly dispatch and an intra-hour balancing market like the electricity markets in the Nordic countries. However, the set-up of the paper can easily be adapted to other market designs. Our contribution is twofold:

• We develop a mixed integer linear programming model for the management of a future electricity system with significant shares of wind energy. We use it to analyse the system, although it can likewise provide guidelines for current system operation. This model is a multi-area economic dispatch model that proactively activates manual reserves such as to minimise balancing cost, while taking operational constraints into account, including detailed ramping restrictions on transmission lines and conventional generation units. To appropriately represent intra-hour system operations, the model has a time resolution of a few minutes. However, to ensure consistency with hour-by-hour day-ahead planning, it is integrated with a unit commitment (UC) model that has hourly time resolution.

• We generate a representative forecast of wind power production that will serve as input to the intra-hour model. To use a suitable marginal distribution for wind power at a given point in time, while also capturing the temporal correlation structure, we employ a copula-based approach.

In the following, we refer to the intra-hour model and the accompanying framework with the unit commitment and forecasting as OPTIBA (OPTimisation of the Intra-hour BAlance).

We illustrate the intra-hour model with a number of case studies based on the Danish elec-
tricity system for varying seasons in a year and for the current and planned future level of wind power penetration. Furthermore, in each case we run our model in a rolling planning fashion to incorporate updated information and thereby simulate the results of optimal intra-hour balancing over time.

1.2 Related literature

The hour-by-hour management of an electricity system is done by deciding on the start-ups, shut-downs and production levels of the power generating units, that is, by solving the UC problem. The first mixed integer linear programming approach to this problem dates back to Garver (1962), and a review of the early works on UC models and corresponding solution methods can be found in Sheble and Fahd (1994). Many UC models incorporate uncertainty in some form. Already in Dillon et al. (1978), the problem of unit outages and the resulting capacity shortage is addressed in a stochastic programming setting. More recent contributions also consider uncertain demand, e.g. see Bunn and Paschentis (1986); Takriti et al. (1996); Carøe and Schultz (1998); Nowak and Römisch (2000).

With the improvement of demand forecasts and the large scale penetration of sustainable energy, renewable production has become the major source of uncertainty and new models have emerged. One such example is found in Weber et al. (2009). This UC model is likewise a stochastic programming model, and it accounts for the variability and unpredictability of wind energy. To reflect occasional updating of wind power forecasts, limit problem size, and reduce computation time, the model is run in a rolling planning fashion with re-optimisation every few hours.

Others handle the problem of uncertainty by solving both the UC and the subsequent intra-day economic dispatch problem in two-stage stochastic mixed integer models. One example is Papavasiliou and Oren (2013) who handles unit commitment of slow units in the first stage and commitment of fast units as well as dispatch of all units in the second stage. This model takes into account ramp rates, contingencies, and transmission network constraints. A wind scenario generation and scenario selection algorithm is included as well as a subgradient algorithm to solve the problem. Similar two-stage models are found in Bouffard and Galiana (2008); Zheng
et al. (2013); Morales et al. (2009); Pritchard et al. (2010). However, in all of these references, the time resolution is hourly, and being only two-stage models, they do not capture the frequent update of wind forecasts carried out in reality. Moreover, the models do not include the technical constraints on reserve power, that are presented in this paper.

For UC models in general, it is the mixed integer nature of the problem combined with a larger number of generating units and time intervals throughout the optimisation horizon that cause problem complexity and computation times to increase drastically. Several attempts to tighten the constraints and limit the number of binary variables have therefore been made, e.g. see Carrión and Arroyo (2006); Ostrowski et al. (2012); Morales-España et al. (2013). Furthermore, considerable attention has been devoted to specifically designed solution methods, most of them based on a decoupling of units by Lagrangian relaxation, see Takriti et al. (1996); Nowak and Römisch (2000). The inclusion of uncertainty makes computational intractability even worse. In spite of this, to fully account for variations in renewable production, there is a need for models with higher time resolutions.

Our model attempts to overcome this need by optimising balancing decisions with a time resolution of a few minutes. To make the problem computationally feasible, we decouple hourly and intra-hourly scheduling and carry out the optimisation in a sequential fashion (which provides an upper bound on the cost of a joint optimisation). Since UC decisions are often made ahead of operation, this decoupling allows us to formulate the intra-hour optimisation as an economic dispatch problem and thereby facilitates the use of a higher time resolution. By considering a short time horizon of only one or two hours, predictability increases and we are therefore able to formulate a deterministic problem.

Related models for intra-hour management include Lindgren and Söder (2008), who presents a mixed integer multi-area optimisation model based on the Northern European system. This paper simulates wind forecasts over time and model the reserve market with re-optimisation when new forecasts become available, taking frequency controls and transmission into consideration, but ignoring operational constraints such as ramping. Reserve activation is done via pre-determined bid lists. Other models replace such exogenously given bid lists by endogenous activation of reserve power, see Jaehnert and Doorman (2012); Farahmand
and Doorman (2012). As in our model, these models are based on sequential optimisation of
day-ahead and balancing operations. Focus is likewise on cost savings in an integrated Northern
European power market, but with emphasis mainly on hydro-power resources. Furthermore,
the authors do not account for ramping restrictions and market rules for the activation
of reserves. Finally, Ela and O’Malley (2012) addresses the sequence of day-ahead planning,
 intra-hour balancing, and second-to-second automatic generation control. Focus is on how
variability and unpredictability of wind power affect system cost and reliability. Although this
model includes technical constraints such as ramp rate restrictions, it still does not account for
the reserve activation rules of the market.

Just as the intra-hour models mentioned above, our model has a sufficiently high time res-
olution to represent the variability of wind, while it in addition considers the complex market
rules and restrictions on ramping and activation of reserve power. Receiving output from a UC,
model it optimises the operation of the system with a two-hour time horizon based on adjust-
ments to the day-ahead schedule via activation of reserves. We set it up in a rolling planning
fashion to accommodate hourly updates of wind forecasts.

The rest of the paper is divided into sections as follows: The interface between hour-by-hour
and intra-hour scheduling is outlined in Section 1.3. The generation of wind power forecasts
can be found in Section 1.4, and our model for optimisation of intra-hour balancing is presented
in Section 1.5. The data and assumptions of the case study are provided in Section 1.6, and the
results are discussed in Section 1.7. Finally, a conclusion is provided in Section 1.8.

1.3 From hour-by-hour to intra-hour scheduling

We aim at short-term balancing of supply and demand through forecasting of wind power pro-
duction and scheduling of conventional generation. In doing so, we assume that an hour-by-
hour production plan has already been made, for example by day-ahead market clearing. On
the basis of this plan, our intra-hour model updates the wind power forecast and reschedules
the generating units with an intra-hour resolution.

The development of an intra-hour model requires detailed information. In generating input
to this model, we wish to replicate the real-life situation in which power producers provide
Intra-hour model OPTIBA

Figure 1.1: Our modelling framework, including the link between the hour-by-hour UC model (run a day ahead of operation) and our intra-hour model (run an hour ahead of operation).

generation schedules while the SO updates consumption and wind power production forecasts close to real-time operation. We generate the input on the basis of an hour-by-hour UC model. The time horizon may be different depending on the market design, but here we refer to the hour-by-hour unit commitment as day-ahead planning. The link between this model and our intra-hour model, OPTIBA, is illustrated by the framework in Fig. 1.1.

The first module, UC, collects hourly values of predicted consumption, forecasted wind power production, planned production schedules for conventional units, and allocated flow on transmission lines from the UC model. OPTIBA reads this data into three modules that convert it to $\tau$-minute time resolution\(^1\), with $\tau$ dividing the 60 minutes of an hour into $60/\tau$ time intervals. The resulting $\tau$-minute values are meant to represent real generation schedules and forecasts provided close to real-time.

The module HA_cons (Hour Ahead consumption) converts the predicted hourly consumption level to an hour-ahead prediction with $\tau$-minute time resolution. For simplicity, and assuming demand is a smooth function of time, we do this using a third order spline interpolation between the hourly values.

HA_prod (Hour Ahead conventional production) converts the schedule for hourly pro-

\(^1\)Depending on the frequency with which activation of manual reserves is permitted by the market rules or technical restrictions of the units, and taking into account the running time of the model, $\tau$ can for example be taken to be five or ten minutes.
duction, and transmission between internal areas as well as to and from external areas (import/export) into τ-minute data. We refer to the change in production or transmission level over time as the ramp rate. For the conventional power generating units, ramping is scheduled in the first and last τ_{max} time intervals of an hour, where 2 · τ_{max} is the number of time intervals a unit can ramp according to the technical restrictions on the generating unit or the power system. The same ramping patterns are used for the transmission lines to both internal and external areas.

HA_wind (Hour Ahead wind power production) collects the hourly values of wind power production and updates the forecast in τ-minute intervals. A more detailed description can be found in Section 1.4.

The information generated by the three modules, HA_cons, HA_prod, and HA_wind, is used as input to the intra-hour model. The model then re-schedules the generating units to cover imbalances between supply and demand, which is what we refer to as the activation of manual reserves.

Note that some information is sent directly from the UC module to the intra-hour model without conversion and is therefore not shown in Fig. 1.1. This includes information about the generating units such as minimum and maximum production levels, variable cost, etc.

1.4 Wind power forecasts

For scheduling in power systems with high wind penetration, an accurate wind power forecast is crucial. The aim of this section is therefore to generate a representative wind power production forecast that will serve as input to the intra-hour model presented in Section 1.5.

For both forecasting and scheduling, we use the following notation. We discretise the scheduling horizon [0, T] into τ-minute intervals [τ(t − 1), τt], t = 1, . . . , T/τ and let \( \mathcal{T} = \{1, \ldots, T/\tau\} \). In the case of hourly intervals, we use the notation [60(h − 1), 60h], h = 1, . . . , T/60 and \( \mathcal{H} = \{1, \ldots, T/60\} \). For the correspondence between these discretisations, we let \( \mathcal{T}_h = \{t \in \mathcal{T} : |τ(t − 1), τt| \subseteq [60(h − 1), 60h]\}, \ h \in \mathcal{H} \).

We assume that an hour-by-hour forecast \( \{\bar{w}_h\}_{h \in \mathcal{H}} \) is already known, but we wish to make a forecast with higher time resolution. As can be seen from Fig. 1.2 on the next page, variations
within the hour are significant. Thus, we aim to generate a $\tau$-minute forecast $\{w_t\}_{t \in T}$. Without further intra-hour information, one could use linear interpolation between the hourly values $\bar{w}_h$ and $\bar{w}_{h+1}$ to determine $w_t$ for $t \in T_h$. Here, we use a slightly different approach.

The idea is to approximate the linear interpolation forecast by a simulation mean. The advantage of the simulation approach is its ability to describe the entire distribution of wind power production and use this as input to a stochastic extension of the intra-hour model. The simulation mean may not fully agree with the values obtained by linear interpolation. In reality, however, the SO usually receives more detailed information, including an updated wind power forecast, close to real-time operation. As a result, deviations occur between the hour-by-hour and intra-hour forecasts. We assume that the simulation mean represents an updated forecast, and interpret the deviations between the simulation and interpolation as forecasting errors.

The simulation approach is based on the use of copulas to describe the temporal dependence structure of the stochastic process, $\{W_t\}_{t \in T}$, of wind power production. Whereas a low-dimensional joint distribution may be explicitly specified by a parametric model, joint parametric modelling becomes cumbersome in higher dimensions. The use of copulas allows us to independently model the marginal distributions and the multivariate dependence structure. We assume marginal Beta distributions to capture the lower fat tail of wind power production.
error, and choose the Gaussian copula, which offers a simple sample procedure to capture the dependence structure.

### 1.4.1 Marginal Beta distributions

Whereas many forecasting studies are based on the assumption of a (symmetric) Normal distribution, e.g. Jordà and Marcellino (2010), we assume marginal Beta distributions to appropriately reflect that high wind speeds are less frequent than low wind speeds. Further justifications of the choice of Beta distribution can be found in Bludszuweit et al. (2008).

For wind power production at a given point in time, we let

$$W_t \sim Be(\alpha_t, \beta_t),$$

where $Be(\alpha_t, \beta_t)$ is the Beta distribution with parameters $\alpha_t > 0, \beta_t > 0$. The distribution function of $W_t$ is then

$$F_t(w; \alpha_t, \beta_t) = \frac{\Gamma(\alpha_t + \beta_t)}{\Gamma(\alpha_t)\Gamma(\beta_t)} \int_0^w v^{\alpha_t-1}(1-v)^{\beta_t-1}dv, \quad 0 \leq w \leq 1,$$

where $\Gamma(\cdot)$ is the gamma function. In the following, we denote this function simply by $F_t(\cdot)$. For how to determine the parameters, see Appendix 1.A on page 33.

### 1.4.2 Gaussian copula

Although wind power production at a given point in time can be described by a Beta distribution, we have disregarded the temporal correlation structure in the specification of the marginal distributions. The Gaussian copula allows us to construct a joint distribution that captures this dependence.

Let $\Sigma$ be a $(T/\tau) \times (T/\tau)$-correlation matrix. For how to find this, see Appendix 1.A on page 33. The Gaussian copula is then

$$F(w_1, \ldots, w_{T/\tau}; \Sigma) = \Phi_{T/\tau}(\Phi^{-1}(F_1(w_1)), \ldots, \Phi^{-1}(F_{T/\tau}(w_{T/\tau})); \Sigma),$$

where the term $\Phi^{-1}(\cdot)$ is the inverse distribution function of a standard Normal distribution and the term $\Phi_{T/\tau}(\cdot; \Sigma)$ is the distribution function of a $(T/\tau)$-variate Normal distribution with correlation matrix $\Sigma$.

Now, let $(U_1, \ldots, U_{T/\tau})^T \sim N_{T/\tau}(0, \Sigma)$, where $N_{T/\tau}(0, \Sigma)$ is the $(T/\tau)$-variate Normal distribution, and for $t = 1, \ldots, T/\tau$, let

$$W_t = F_t^{-1}(\Phi(U_t)),$$
where $F_t^{-1}(\cdot)$ is the inverse of the marginal distribution function $F_t(\cdot)$. We then obtain that $W_t \sim \text{Be}(\alpha_t, \beta_t)$ with marginal distribution function $F_t(\cdot)$ and that $(W_t, \ldots, W_{T/\tau})^\top$ has joint distribution function $F(\cdot; \Sigma)$.

1.4.3 Simulation

We can use the above observations for sampling. If we generate a number of sample paths $(u_1^s, \ldots, u_{T/\tau}^s), s = 1, \ldots, S$ from the $(T/\tau)$-variate Normal distribution $N_{T/\tau}(0, \Sigma)$, and for each such sample path compute

$$w_t^s := F_t^{-1}(\Phi(u_t^s)),$$

then $(w_1^s, \ldots, w_{T/\tau}^s)^\top, s = 1, \ldots, S$ can be viewed as samples from a joint distribution with marginal Beta distributions $\text{Be}(\alpha_t, \beta_t)$ and temporal correlation matrix $\Sigma$.

To generate samples from the $(T/\tau)$-variate Normal distribution $N_{T/\tau}(0, \Sigma)$, we make the following observations. Let $(V_1, \ldots, V_{T/\tau})^\top \sim N_{T/\tau}(0, I)$, where $I$ denotes the $(T/\tau) \times (T/\tau)$-identity matrix. The marginals $V_t$ are independent and all have a standard Normal distribution. Apply Cholesky decomposition of the correlation matrix such that

$$\Sigma = LL^\top,$$

where $L$ is a lower triangular $(T/\tau) \times (T/\tau)$-matrix and $L^\top$ its transpose. We now have that $(U_1, \ldots, U_{T/\tau})^\top := L(V_1, \ldots, V_{T/\tau})^\top \sim N_{T/\tau}(0, \Sigma)$. Hence, if we generate independently a number of samples $v_t^s, s = 1, \ldots, S$ from the standard Normal distributions, then

$$(w_1^s, \ldots, w_{T/\tau}^s)^\top := L(v_1^s, \ldots, v_{T/\tau}^s)^\top, s = 1, \ldots, S$$

can be viewed as samples from the $(T/\tau)$-variate Normal distribution with correlation matrix $\Sigma$.

The generated sample paths $(w_1^s, \ldots, w_{T/\tau}^s)^\top, s = 1, \ldots, S$ describe the distribution of wind power production and may serve as input to a stochastic scheduling model. However, since our model is currently deterministic, we compute the simulation mean

$$w_t = \frac{1}{S} \sum_{s=1}^{S} w_t^s, \quad t \in T,$$

and use $\{w_t\}_{t \in T}$ as the wind power forecast.
1.4.4 Forecasting results

In the above, we work on wind power production normalised by capacity. In our case study, we therefore multiply the forecast with a given wind power capacity for each geographical area. Fig. 1.3 illustrates 15 of 5000 sample paths used for computing the simulation mean. The data is based on expected wind penetration in Western Denmark in 2020, and the generated forecast has a five minute time resolution. Clearly, wind power production shows large variations over time. Moreover, Fig. 1.3 and Fig. 1.4 on the following page display the simulation mean computed on the basis of 500 and 5000 samples, respectively, and the linear interpolation forecast. It is likewise clear that our updated forecast shows significant intra-hour variations around the linear interpolation between the hourly values, as would be the case in reality.

1.5 The intra-hour model

The aim of the intra-hour model is efficient short-term balancing of supply and demand in the electricity system. We consider balancing close to real-time operation, when a production plan has already been made by day-ahead market clearing. On the basis of this plan, our model re-dispatches the generating units, which is what we refer to as intra-hour scheduling of manual reserves. In doing so, we assume that the UC schedule has been fixed, and consider only the
re-dispatch of already committed units\(^2\). As already mentioned, short-term balancing is the responsibility of the SO, and therefore we base our model on social welfare maximisation, or equivalently, assuming inflexible demand, minimisation of balancing cost.

We take the scheduling horizon to be a few hours, which we discretise into \(\tau\)-minute intervals. For the notation, see Section 1.4. To formulate the intra-hour model, we further introduce the following notation.

The transmission grid is modelled as a network \((\mathcal{N}, \mathcal{A})\) with finitely many nodes \(\mathcal{N}\) and arcs \(\mathcal{A} = \{a : a = (n, n'), n, n' \in \mathcal{N}, n < n'\}\) representing transmission lines. We denote by \(\delta^{\text{out}}(n) = \{a : a = (n, n'), n' \in \mathcal{N}\}\) and \(\delta^{\text{in}}(n) = \{a : a = (n', n), n' \in \mathcal{N}\}\) the sets of arcs originating from and terminating in node \(n \in \mathcal{N}\), respectively. For \(a \in \mathcal{A}\), we let the capacity of transmission line \(a\) be \(L_{\text{max}}^a\) and the maximum ramping rate be \(R_a\). Moreover, we let the flow allocated day-ahead to line \(a\) in time interval \([\tau(t - 1), \tau t]\) be \(L_{\text{at}}\). Furthermore, there is a net import from \(n\) to \(n'\), or equivalently a net export from \(n'\) to \(n\), if \(a = (n, n')\) and \(L_{\text{at}} > 0\). To account for loss, we adjust the available capacity on the line by a small factor. We represent the intra-hour scheduled import on the transmission line in the same time interval by the variable

\(^2\)This assumption can be justified for power generation units with start-up times in excess of a few hours. However, some power generation units may have start-up times less than an hour, in which case this is a simplifying assumption.
\( \Delta l_{at} \) using the same conventions regarding its sign.

The set of conventional units is denoted by \( \mathcal{I} \), the set of units online in the time interval \([\tau(t-1), \tau t]\) is denoted by \( \mathcal{I}_t \), and the units located in node \( n \) by \( \mathcal{I}_n \). For \( i \in \mathcal{I} \), we let the minimum and maximum generation level of unit \( i \) be \( P_{i}^{\text{min}} \) and \( P_{i}^{\text{max}} \), respectively, and the maximum rates for ramping up and down be \( R_{i}^{+} \) and \( R_{i}^{-} \). We denote by \( C_i \) the variable generation cost of unit \( i \in \mathcal{I} \). Note that whereas activation of reserves generates an additional cost, deactivation of reserves results in cost savings. We therefore let \( C_{i}^{+} := (1 + \gamma) C_{i} \) and \( C_{i}^{-} := (1 - \gamma) C_{i} \) be the cost of activating manual reserves and savings of deactivating manual reserves, respectively, with \( \gamma \in [0,1] \) being a mark-up or mark-down. The idea is to allow the cost and savings to reflect the additional stress imposed on the unit when using it for balancing purposes. We assume that the mark-up and mark-down are the same, though this is not always the case in the market. The costs and savings of activating or deactivating automatic reserves in node \( n \) are denoted by \( C_{n}^{+} \) and \( C_{n}^{-} \), respectively, where most likely, \( C_{n}^{+} > \max_{i} C_{i}^{+} \) and \( C_{n}^{-} < \min_{i} C_{i}^{-} \).

We now let day-ahead planned generation on unit \( i \) in time interval \([\tau(t-1), \tau t]\) be given by the parameter \( P_{it} \). The variables \( \Delta p_{it}^{+} \geq 0 \) and \( \Delta p_{it}^{-} \geq 0 \) represent the intra-hour activation and deactivation of manual reserves on the unit, respectively. Likewise, the variables \( q_{nt}^{+} \geq 0 \) and \( q_{nt}^{-} \geq 0 \) represent the generation shortage and surplus in node \( n \) during time interval \([\tau(t-1), \tau t]\). In order to keep track of whether reserves are activated or deactivated compared to the original production, we introduce two binary variables, \( z_{it}^{+} \) and \( z_{it}^{-} \). Hence, \( z_{it}^{+} = 1 \) if we have activated reserves in time interval \( t \) and zero otherwise. Likewise, \( z_{it}^{-} = 1 \) if we have deactivated reserves and zero otherwise.

To model detailed ramping restrictions, we let the binary variables \( u_{it}^{+} \) and \( u_{it}^{-} \) be one if unit \( i \) is ramping towards or from a new generation level, respectively, in time interval \([\tau t, \tau (t+1)]\) and zero otherwise. The binary variable \( v_{it} \) is set to one when the ramping towards a new generation level has finished. Newly activated and deactivated reserves at time \( t \) are represented by the variable \( g_{it} \geq 0 \), and the minimum level of new reserves that can be activated is given by the parameter \( G_{i}^{\text{min}} \). This resembles the bidding practice of the balancing market, but allows for optimising the use of reserves as opposed to using bid lists.

Finally, we assume that wind power production and demand is inflexible, and denote their
values by the parameters $w_{nt}$ and $D_{nt}$ in node $n$ and time interval $|\tau(t-1), \tau t|$.

### 1.5.1 Economic dispatch

We schedule the activation of manual reserves to cover any imbalances between supply and demand. Occasionally, this may be technically infeasible or it may be feasible only at very high cost, in which case imbalances are left to automatic reserves. The optimal schedule is therefore determined by a trade-off between the activation cost of manual and automatic reserves. The objective is

$$\sum_{t \in T} \left( \sum_{i \in I} \left( C_i^+ \Delta p_{it}^+ - C_i^- \Delta p_{it}^- \right) + \sum_{n \in N} \left( C_n^+ q_{nt}^+ - C_n^- q_{nt}^- \right) \right),$$

which is minimised subject to the following constraints.

The balancing constraint ensures system balance between supply and demand. According to this constraint, if, at any point in time, scheduled production exceeds predicted consumption or vice versa, we experience generation surplus or shortage, which will be left to automatic reserves. We assume that it is always possible to provide sufficient automatic reserves. Production includes day-ahead planned generation on conventional units, intra-hour activation and deactivation of manual reserves, forecasted wind power production, and finally day-ahead and intra-hour net import/export on the transmission lines. Thus, we have that

$$\sum_{i \in I} \left( P_{it} + \Delta p_{it}^+ - \Delta p_{it}^- \right) + \sum_{a \in A(n)} (L_{at} + \Delta l_{at}) - \sum_{a \in A(n)} (L_{at} - \Delta l_{at}) + q_{nt}^+ - q_{nt}^- = D_{nt} - w_{nt}, \quad n \in N, t \in T. \quad (1.1)$$

Transmission flow is restricted by the available line capacity. In particular, intra-hour scheduled import on the transmission lines is bounded above by the line capacity less the capacity allocated day-ahead. Thus, we have that

$$-(L_{a}^\text{max} - L_{at}) \leq \Delta l_{at} \leq L_{a}^\text{max} - L_{at}, \quad a \in A, t \in T.$$

Activation of reserves is bounded above by the capacity that has not already been dispatched day-ahead, whereas deactivation is bounded by the dispatched capacity in excess of
the minimum capacity. Formally,
\[ \Delta p^+_i \leq (P^\text{max}_i - P_{it})z^+_it, \quad i \in \mathcal{I}_t, t \in \mathcal{T}, \]
\[ \Delta p^-_i \leq (P_{it} - P^\text{min}_i)z^-_it, \quad i \in \mathcal{I}_t, t \in \mathcal{T}. \]

Finally, to ensure that a unit does not activate and deactivate reserves at the same time, we have that
\[ z^+_it + z^-_it \leq 1, \quad i \in \mathcal{I}_t, t \in \mathcal{T}. \]

The above constraints are very similar to those of the hour-by-hour economic dispatch problem. In our formulation of the intra-hour balancing problem, however, we include much more detailed ramping restrictions.

**Ramping**

We assume simple ramping constraints on the transmission lines. These are restrictions on the change in allocated transmission flow from one time interval to another and apply to net import. Thus,
\[ -R_a - L_{a(t+1)} + L_{at} \leq \Delta l_{a(t+1)} - \Delta l_{at} \leq R_a - L_{a(t+1)} + L_{at}, \]
\[ a \in \mathcal{A}, t \in \mathcal{T} : t \leq T/\tau - 1. \]

For the generation units, we include detailed ramping restrictions.
\[ -(R^-_i + P_{i(t+1)} - P_{it})u^-_it \leq \Delta p^+_i - \Delta p^-_i \leq (R^+_i + P_{i(t+1)} + P_{it})u^+_it, \]
\[ i \in \mathcal{I}_t, t \in \mathcal{T} : t \leq T/\tau - 1, \]
\[ -(R^-_i - P_{i(t+1)} + P_{it})u^-_it \leq \Delta p^-_i - \Delta p^-_i \leq (R^-_i + P_{i(t+1)} - P_{it})u^-_it, \]
\[ i \in \mathcal{I}_t, t \in \mathcal{T} : t \leq T/\tau - 1. \]

Here we also record if a unit is ramping towards \((u^+_it = 1)\) or back from \((u^-_it = 1)\) a previous generation level. To prevent the ramping variables from being one when no ramping occurs,
we assume a minimum ramp rate of \( \varepsilon \).

\[
\varepsilon u^+_{it} + M (u^+_{it} - 1) \leq \Delta p^+_{i(t+1)} - \Delta p^+_{it} + \Delta p^-_{it},
\]

\[i \in \mathcal{I}_t, t \in \mathcal{T} : t \leq T/\tau - 1,
\]

\[
\varepsilon u^-_{it} + M' (u^-_{it} - 1) \leq \Delta p^+_{it} - \Delta p^-_{it(t+1)} + \Delta p^-_{i(t+1)},
\]

\[i \in \mathcal{I}_t, t \in \mathcal{T} : t \leq T/\tau - 1,
\]

where \( \varepsilon \) is a sufficiently small number and \( M \) and \( M' \) are sufficiently large numbers.

To reduce the stress imposed on the unit when using it for balancing purposes, ramping is further constrained as follows. Ramping in consecutive time intervals is restricted to a maximum time of \( \tau_{\text{max}} \), which we enforce by the constraints

\[
\min\{T/\tau, t + \tau_{\text{max}}\} \sum_{t' = t}^{t + \tau_{\text{max}}} (u^+_{it'} + u^-_{it'}) \leq \tau_{\text{max}}, \quad i \in \mathcal{I}_t, t \in \mathcal{T}. \tag{1.3}
\]

**New generation level of a unit**

Upon activation of reserves, a unit must be scheduled to operate at a minimum level for a fixed time. However, by consecutively activating or deactivating reserves at the same level, this is equivalent to considering it as a minimum activation time. Additional levels of reserves may be activated or deactivated on the same unit at a given point in time. This we refer to as new reserves\(^3\). For an illustration of the resulting operation schedule of a unit, see Section 1.5.2.

To ensure a minimum activation level, we have

\[
G_{\text{min}}^{i\min} v_{it} \leq g_{it} \leq M v_{it}, \quad i \in \mathcal{I}_t, t \in \mathcal{T}, \tag{1.4}
\]

where \( M \) is a sufficiently large number, e.g. \( p_{\text{max}}^i \). The minimum generation level is activated upon ramping towards a new level, hence,

\[
u^+_{i(t-1)} - u^+_{it} \leq v_{it}, \quad i \in \mathcal{I}_t, t \in \mathcal{T}. \tag{1.5}
\]

\(^3\)For example, assume that 5 MWh (per \( \tau \)-minute time interval) is activated at time 12:00 for 20 minutes. At 12:10, we may additionally activate 5 MWh for 20 minutes. Total amount of reserves provided is then 5 MWh at 12:00, 10 MWh at 12:10, and 5 MWh at 12:20.
The ramping towards a new generation level, the ramping back from a new generation level, and the activation of a new generation level cannot happen simultaneously, so

\[ u_{it}^+ + u_{it}^- + v_{it} \leq 1, \quad i \in I, t \in T. \]

Finally, the total amount of reserves provided is the sum of previously activated (respectively de-activated) reserve levels with a fixed time of \( \tau_{\text{res}} \). Hence,

\[
\sum_{t' = \max \{1, t - \tau_{\text{res}} + 1\}}^{t} g_{it'} \leq \Delta p_{it}^+ + \Delta p_{it}^- \leq \sum_{t' = \max \{1, t - \tau_{\text{res}} + 1\}}^{t} g_{it'} + M(u_{it}^+ + u_{it}^-),
\]

(1.6)

\( i \in I, t \in T, \)

for \( M \) large enough, e.g. \( P_{i}^{\text{max}} \).

### 1.5.2 Example of activation of reserves

The activation of reserves abides by a number of rules originating from the technical restrictions on the units and the rules of the balancing market. The following provides an example of how to manage reserves in the model. For simplicity, we confine ourselves to intra-hour scheduling and activation of reserves.

Consider the model with a time resolution of \( \tau = 5 \) minutes, a fixed activation time of \( \tau_{\text{res}} = 6 \) time intervals, a maximum ramping time of \( \tau_{\text{max}} = 2 \) time intervals, a minimum ramping level of \( \varepsilon = 1 \) MW, and finally a minimum reserve activation level of \( G_{i}^{\text{min}} = 2 \) MWh/12, corresponding to produce at 2 MW during each five minute time interval. In the balance equation (1.1) the following deficit of energy compared to the day-ahead planned production is detected:

<table>
<thead>
<tr>
<th>t</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy (MWh/12)</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

This deficit can be covered by a reserve activation of length \( \tau_{\text{res}} = 6 \) with \( \Delta p_{it}^+ = 2 \) MWh/12 for some unit \( i \) in the time intervals \( t = 3, \ldots, 8 \). Assuming that the unit starts to ramp towards this
value at time $t = 1$, and ramp in a maximum of $\tau_{\text{max}} = 2$ intervals, the ramping constraint (1.2) will set $u_{i1}^+ = 1$. At $t = 2$, we further ramp towards the new volume, setting also $u_{i2}^+ = 1$. Since we can only ramp for two consecutive time intervals, $u_{i3}^+ = 0$. With the minimum ramping of $\varepsilon = 1$ MW, the unit can cover the deficit in the ramping period $t = 2$. As the unit reaches the desired level of production, the ramping ends and (1.5) ensures that $v_{i3} = 1$. This means that a new activation of reserves has taken place. Equation (1.4) now forces the reserve level to be above the minimum value of $G_{i\text{min}} = 2$ MWh/12, i.e. $2 = g_{i3} \geq G_{i\text{min}} = 2$. The reserve activation variable $\Delta p_{i8}^+$ is maintained at 2 MWh/12 by (1.6) during the six time intervals; see the dark grey reserve activation in Fig. 1.5 on the next page. In case no additional reserves are activated, (1.6) forces $\Delta p_{i8}^+$ to ramp back to the day-ahead level when the activation ends, which again forces $u_{i8}^- = u_{i9}^- = 1$ by (1.2). Since we can only ramp for two consecutive intervals, $u_{i10}^- = 0$ by Equation (1.3). As before, the unit can cover the deficit in the ramping period $t = 9$. Finally, Equation (1.6) ensures that we are back at the day-ahead planned production level at time $t = 10$. In the case of further need of reserves, e.g. in time interval $t = 6$, we can activate new reserves on top of the initial activation. Let the total deficit compared to the day-ahead planned production be:

<table>
<thead>
<tr>
<th>t</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy (MWh/12)</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3.5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>1.5</td>
</tr>
</tbody>
</table>

Then the unit starts ramping towards the new level in time interval $t = 4$, forcing $u_{i4}^+ = 1$ by (1.2) and still ramps in time interval $t = 5$, so that also $u_{i5}^+ = 1$. Note that we could not have started ramping to the new reserve level at $t = 3$ due to the maximum ramping restriction (1.3). This new activation is shown as the light grey area of Fig. 1.5 on the facing page. Again by (1.6) we must ramp back from the initial reserve activation in time intervals $t = 8$ and $t = 9$, but preserving our latest and still running reserve activation. This will continue to run at the level $g_{i9} = 3$ MWh/12 for the whole of the activation period.

Should there be a need for a longer reserve activation, by (1.6) it is possible to continue at the same reserve level directly after the previous has ended, without ramping back, by letting $g_{i9} = g_{i3}$. 

20
1.6 Case study

In our case studies, we consider a time horizon of two hours with a five minute time resolution, corresponding to $\tau = 5$ and $T = 120$.

1.6.1 Rolling planning

We generate an hour-by-hour production plan by running the UC model WILMAR from Weber et al. (2009) with data from the Danish electricity system of 2010$^4$. When the UC model has run for a day, the intra-hour model is launched. Initially, it updates the wind forecast and converts the hourly output from the UC model (production, consumption, etc.) to intra-hour values, see Section 1.3. To reflect the frequent updating of wind forecasts, we run the model in a rolling planning fashion. Hence, the intra-hour model reschedules the generation units for the first two hours of the day, discards the second hour, and is rerun for the second and third hour of the day, etc. By using a two-hour time horizon, we avoid undesired end effects in the intra-hour optimisation such as the tendency of production units to decrease the generation level towards the end of the hour. The end values from the hour previous to the two-hour horizon serve as input to the optimisation in order to accommodate reserve activation across the shift between hours. The model is run for each hour of the day, with the final run corresponding to

---

$^4$System data has kindly been provided by Energinet.dk, which is the Danish transmission system operator (TSO).
the time period from 10 p.m. until midnight. In this run, both the first and the second hour of
the optimisation horizon are used. Finally, the UC model is run for another day. The process is
repeated for each day of a week.

1.6.2 Cases

We consider a winter week in January 2010 as our base case, and compare this to a week in
May, July, and October of 2010 corresponding to Spring, Summer, and Fall cases. To comple-
ment the 2010 cases, we increase the capacity of wind power in the data to fit the Danish 2020
goal according to guidelines from the Danish TSO. This results in four cases corresponding to
the same weeks but with a significantly higher wind penetration. For our case studies, this
represents an increase in the wind power penetration from an average of 22.4% in the current
wind penetration cases to 30.4% in the high wind penetration cases.

In all eight cases, the ramping on transmission lines to areas outside the two balancing areas
is handled as for the production units, see Section 1.3. We consider $\tau^{\max} = 3$ time intervals,
resulting in a ramping pattern where the transmission lines are ramping from one hourly level
to the next within the last 15 minutes of the previous hour and the first 15 minutes of the current
hour. However, according to the Danish TSO, this is only the case for transmission lines to the
Nordic countries. Ramping on transmission lines to Central Europe is done within the last five
minutes of the previous hour and the first five minutes of the current. We investigate how this
inconsistency, that is, different ramping patterns, leads to additional system imbalances in the
case of Winter 2010.

1.6.3 Model complexity and running times

The model is run with an Intel Core 3.10GHz processor and 4 GB RAM. The model is imple-
mented in the 64 bit GAMS framework version 24.1.3 for Windows using the CPLEX 12.5.3.0
solver.

For each case, the rolling planning setup of OPTIBA with a total time horizon of a week
leads to seven runs of the UC model and a total of 168 runs of the intra-hour model. For the
eight 2010 and 2020 cases, running times vary between 11 hours and 31 minutes (Winter 2010)
and 22 hours and 31 minutes (Summer 2010). Considering the relatively long running times for a whole week, recall that the purpose of our model runs is analysis of the system rather than actual rescheduling. The average running time for the two hour horizon is less than six minutes which would easily be applicable for an SO wishing to reschedule proactively. Finally, note that the Winter 2010 case with different ramping patterns is significantly slower than the standard 2010 case: For the whole week, the running time is 24 hours and 45 minutes.

Out of the 168 model runs for each of the eight current and high wind penetration cases, at least 80.9% solve with a gap less than 2%, 89.9% solve with a gap less than 5%, and 95.2% with a gap less than 10%. For the Winter 2010 case with different ramping patterns, the gaps are slightly larger: 79.8%, 88.1%, and 92.3% of the model runs solve with a gap less than 2%, 5%, and 10%, respectively.

1.6.4 System characteristics

The system consists of two balancing areas with a total of 118 power producing units. For each of the two balancing areas the wind farms are grouped into offshore and onshore wind power producing units. Aggregated characteristics of hourly power system operations from the modules HA_cons, HA_prod, and HA_wind are displayed in Table 1.1 on the next page. As expected, the higher the wind power share is, the lower the conventional production and import are, and the higher the export is.

1.6.5 Intra-hour balancing characteristics

Current guidelines from the Danish TSO, based on physical restrictions on the generating units, lead to the following assumptions. Each reserve activation or deactivation lasts $\tau_{\text{res}} = 6$ time intervals. Longer activation times are handled in the model as a multiple of individual reserve activations. The minimum activation or deactivation level is $G_i^{\text{min}} = 10 \text{ MWh}/12$ corresponding to a production level of 10 MW in the activation or deactivation period. The maximum ramping time $\tau^{\text{max}} = 3$ time intervals, and the minimum ramping limit $\varepsilon = 1 \text{ MW}$.

Regarding the cost of reserves, we assume that the cost of manual activation or deactivation of reserves is the marginal cost of the unit multiplied by a mark-up or mark-down, respectively.
Table 1.1: The table shows aggregated characteristics of hourly power system operations for a selected week in each season with current and high wind power share. The aggregated characteristics consist of the means (standard deviations) of demand, conventional production, conventional capacity, wind power production, wind power capacity, import, and export. Capacities are in MW and all other values in MWh/12 corresponding to the five minute intervals. Note that the wind capacity does not vary over time as we assume turbines are never shut down.

<table>
<thead>
<tr>
<th></th>
<th>All areas – 2010 cases – current wind penetration</th>
<th>All areas – 2020 cases – high wind penetration</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Winter</td>
<td>Spring</td>
</tr>
<tr>
<td>Demand (MWh/12)</td>
<td>4183(884)</td>
<td>3554(667)</td>
</tr>
<tr>
<td>Conv. prod. (MWh/12)</td>
<td>3110(483)</td>
<td>2857(301)</td>
</tr>
<tr>
<td>Online capacity (MW)</td>
<td>4397(492)</td>
<td>4015(361)</td>
</tr>
<tr>
<td>Wind power prod. (MWh/12)</td>
<td>1880(878)</td>
<td>251(242)</td>
</tr>
<tr>
<td>Installed wind capacity (MW)</td>
<td>4752</td>
<td>4752</td>
</tr>
<tr>
<td>Import (MWh/12)</td>
<td>1438(493)</td>
<td>2222(610)</td>
</tr>
<tr>
<td>Export (MWh/12)</td>
<td>2227(475)</td>
<td>1778(534)</td>
</tr>
</tbody>
</table>

Furthermore, we assume system-wide cost of the automatic reserves. As mentioned above, usually $C^+ > \max_i C^+_i$ and $C^- < \min_i C^-_i$. However, as we may need to dispatch units with very high marginal cost in some time intervals, the difference between the lowest and highest marginal cost of the dispatched units in our cases is relatively large. Hence, we estimate the cost of automatic reserves from the historical price of manual reserves. In particular, we assume that the cost of activating automatic reserves is above the historical price of manual reserves 95% of the time, and that of deactivating automatic reserves is below the historical price 5% of the time. In other words, we set the cost of activating and deactivating automatic reserves to the 95% and the 5% fractile of the historical distribution of reserve prices in the Danish system, respectively. The cost can be seen in Table 1.2 on the facing page.

In future power systems with even higher wind power penetrations, the SOs may have to install further reserves to maintain system reliability. This would be at an additional cost, which is not present in our model. However, the cost could be included as a higher cost on
Table 1.2: The cost of reserves. The cost of manual reserves depends on the marginal cost of the units and a mark-up or mark-down for activation and deactivation, respectively. The cost of automatic reserves is a system-wide cost.

<table>
<thead>
<tr>
<th></th>
<th>Marginal cost</th>
<th>Mark-down</th>
<th>Mark-up</th>
<th>Down</th>
<th>Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manual reserves (€)</td>
<td>3.8-216.2</td>
<td>0.9</td>
<td>1.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Automatic reserves (€)</td>
<td>20</td>
<td>95</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

automatic reserves. We have investigated the influence of a change in the cost of automatic reserves in Appendix 1.B on page 34.

1.7 Results and discussion

In this section, we present numerical results from our intra-hour model and thereby aim to justify the need for such models.

1.7.1 Cost savings from using proactive reserve activation

From Table 1.3 on the following page, it is evident that hourly scheduling without proactive re-dispatch leads to extensive use of the expensive automatic reserves. By proactively activating and deactivating reserves, these reserve cost can be reduced. For the 2010 cases, savings vary between 2.1 – 4.9‰ of total system cost for different seasons. In spite of the small share of total cost, the absolute savings of €391,601 – 628,070 are not insignificant. For the 2020 cases, the same pattern shows, savings varying between 2.4 – 9.3‰ of total system cost corresponding to €408,758 – 738,137. These savings will increase if the cost of automatic reserves rises in the future.

Comparing the 2010 cases with the 2020 cases, it is clear that higher wind penetration results in lower total system cost, including reserve cost, both for the proactive and the reactive strategy. Thus, the benefits from the inexpensive wind power is not outweighed by the balancing cost. Furthermore, the savings from using the proactive approach are larger for the 2020 case, and hence, future wind penetrations of even greater magnitude will only make the intra-hour model increasingly relevant.
Table 1.3: Weekly cost of operating the system proactively and reactively. For each strategy, cost are divided into manual activation and deactivation of reserves and automatic activation and deactivation of reserves. Moreover, the table shows total balancing cost and total system cost including day-ahead cost. All numbers listed are in €.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>All areas – 2010 cases – current wind penetration</th>
<th>Winter</th>
<th>Spring</th>
<th>Summer</th>
<th>Fall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proactive</td>
<td>Man. act.</td>
<td>414 128</td>
<td>283 371</td>
<td>279 582</td>
<td>396 637</td>
</tr>
<tr>
<td></td>
<td>Man. deact.</td>
<td>−427 554</td>
<td>−224 730</td>
<td>−206 994</td>
<td>−309 489</td>
</tr>
<tr>
<td></td>
<td>Auto. act.</td>
<td>506 66</td>
<td>328 00</td>
<td>429 83</td>
<td>534 28</td>
</tr>
<tr>
<td></td>
<td>Auto. deact.</td>
<td>−620 15</td>
<td>−580 30</td>
<td>−706 73</td>
<td>−630 94</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>−14 560</td>
<td>56 118</td>
<td>69 820</td>
<td>86 188</td>
</tr>
<tr>
<td></td>
<td>Total system</td>
<td>126 668 489</td>
<td>186 520 657</td>
<td>143 835 549</td>
<td>180 548 204</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Strategy</th>
<th>All areas – 2020 cases – high wind penetration</th>
<th>Winter</th>
<th>Spring</th>
<th>Summer</th>
<th>Fall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proactive</td>
<td>Man. act.</td>
<td>616 352</td>
<td>299 673</td>
<td>270 315</td>
<td>395 581</td>
</tr>
<tr>
<td></td>
<td>Man. deact.</td>
<td>−337 719</td>
<td>−246 961</td>
<td>−289 132</td>
<td>−331 126</td>
</tr>
<tr>
<td></td>
<td>Auto. act.</td>
<td>14 159</td>
<td>409 38</td>
<td>438 00</td>
<td>666 99</td>
</tr>
<tr>
<td></td>
<td>Auto. deact.</td>
<td>−15 128</td>
<td>−659 85</td>
<td>−653 66</td>
<td>−666 66</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>123 947</td>
<td>50 207</td>
<td>−20 972</td>
<td>64 458</td>
</tr>
<tr>
<td></td>
<td>Total system</td>
<td>78 989 767</td>
<td>175 514 299</td>
<td>124 950 092</td>
<td>161 777 073</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Strategy</th>
<th>All areas – 2020 cases – high wind penetration</th>
<th>Winter</th>
<th>Spring</th>
<th>Summer</th>
<th>Fall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reactive</td>
<td>Man. act.</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Man. deact.</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Auto. act.</td>
<td>1 090 133</td>
<td>599 378</td>
<td>540 179</td>
<td>806 828</td>
</tr>
<tr>
<td></td>
<td>Auto. deact.</td>
<td>−228 049</td>
<td>−125 144</td>
<td>−152 392</td>
<td>−173 694</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>862 084</td>
<td>474 234</td>
<td>387 786</td>
<td>633 133</td>
</tr>
<tr>
<td></td>
<td>Total system</td>
<td>79 727 904</td>
<td>175 938 326</td>
<td>125 358 850</td>
<td>162 345 747</td>
</tr>
</tbody>
</table>

1.7.2 Imbalances, updates of wind power forecast, and ramping on transmission lines

Table 1.4 shows that total balancing cost depend on total imbalances in the system. More specifically, total cost of activation and deactivation of reserves can in most cases be explained by the corresponding total deficit and surplus in the system. In all cases, except Fall 2020, deficit of power incurs positive total reserve cost, whereas surplus of power incurs negative overall reserve cost because of the need for activation and deactivation of reserves, respectively. The standard deviation of the imbalances in Table 1.4 on the next page indicates that there are more
Table 1.4: Weekly averages (standard deviation) of total imbalances and wind forecast errors in the system before re-dispatch in each five minute interval. Moreover, the table shows the average positive and negative imbalances and the average positive and negative forecast errors. The values are in MWh/12 corresponding to the five minute intervals. When the forecast error is positive, the updated forecast is higher than the original forecast.

<table>
<thead>
<tr>
<th></th>
<th>Winter</th>
<th>Spring</th>
<th>Summer</th>
<th>Fall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deficit</td>
<td>-52.6</td>
<td>-36.4</td>
<td>-36</td>
<td>-50.8</td>
</tr>
<tr>
<td>Surplus</td>
<td>69.2</td>
<td>34.8</td>
<td>33.8</td>
<td>47.9</td>
</tr>
<tr>
<td>Imbal.</td>
<td>16.6(189.6)</td>
<td>-1.6(104.5)</td>
<td>-2.3(110.9)</td>
<td>-2.9(142.3)</td>
</tr>
<tr>
<td>Neg. error</td>
<td>-1.7</td>
<td>-5.5</td>
<td>-19.6</td>
<td>-9.2</td>
</tr>
<tr>
<td>Pos. error</td>
<td>8</td>
<td>2.4</td>
<td>3.8</td>
<td>6.1</td>
</tr>
<tr>
<td>Wind error</td>
<td>6.3 (9.3)</td>
<td>-3.1 (9.1)</td>
<td>-6.8 (16.7)</td>
<td>-3.1 (17.9)</td>
</tr>
</tbody>
</table>

All areas – 2020 cases – high wind penetration

<table>
<thead>
<tr>
<th></th>
<th>Winter</th>
<th>Spring</th>
<th>Summer</th>
<th>Fall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deficit</td>
<td>-67.9</td>
<td>-37.6</td>
<td>-33.8</td>
<td>-50.5</td>
</tr>
<tr>
<td>Surplus</td>
<td>67.5</td>
<td>37.2</td>
<td>45.4</td>
<td>51.6</td>
</tr>
<tr>
<td>Imbal.</td>
<td>-0.4(217.3)</td>
<td>-0.3(107.8)</td>
<td>11.5(118.1)</td>
<td>1.1(153.2)</td>
</tr>
<tr>
<td>Neg. error</td>
<td>-18.1</td>
<td>-4.7</td>
<td>-8.8</td>
<td>-7.2</td>
</tr>
<tr>
<td>Pos. error</td>
<td>5.7</td>
<td>3.2</td>
<td>15.7</td>
<td>8.9</td>
</tr>
<tr>
<td>Wind error</td>
<td>-12.4 (28.4)</td>
<td>-1.6 (10)</td>
<td>6.9 (31)</td>
<td>1.7 (21.8)</td>
</tr>
</tbody>
</table>

frequent and/or larger imbalances in some cases than in others. This may increase the cost of balancing the system since activation of automatic reserves is more expensive than deactivation. Thus, the large standard deviation of imbalances in the Fall 2020 case leads to positive total balancing cost, even though there is a total excess of power before the intra-hour model is run. The standard deviation of imbalance can also explain some of the differences in total balancing cost for the proactive approach in Table 1.3, between cases with average imbalance of the same sign. For instance, for the 2010 cases, the standard deviation is smallest in the spring, followed by summer and fall, which corresponds to the balancing cost being smallest for the spring case, followed by summer and fall.

Table 1.4 also shows wind forecast errors. In all cases, the total system imbalance is negative (positive) when the total wind forecast error is negative (positive) as expected. In fact, the average wind forecast errors contribute substantially to the total imbalances, cause activation and deactivation of reserves, and thereby account for a significant share of balancing cost.
1.7.3 Imbalances and ramping on production units and transmission lines

The updates of wind power forecasts, the ramping on power producing units, the ramping on transmission lines, and the higher time resolution of demand data all affect the imbalances in the system. The imbalances are reduced when manual reserves are proactively deployed. However, they are not entirely eliminated, which is evident from Fig. 1.6. This is due to the ramping restrictions on production units that impose bounds on the use of reserves, making it impossible to meet demand precisely. It would be very expensive to meet imbalances at the spikes, as this would require many of the slower units to ramp up and thereby incur large imbalances in the time intervals before and after the spikes. Hence, this would never be optimal from a cost minimisation point of view.

For the Winter 2010 case with different ramping patterns on transmission lines, the imbalances are larger than for the Winter 2010 base case. The imbalances are depicted in Fig. 1.7 on the next page which shows how the imbalances exhibit additional spikes, that will not be captured in hourly UC models. These substantially larger imbalances increase balancing cost by €57,909 corresponding to 0.5‰ of the total system cost. As in the base case, the imbalances are reduced when reserves are proactively deployed. However, the imbalances are not entirely eliminated, and a higher frequency of spikes remain after re-dispatch than in the base case.
Although not cost efficient, the SO may wish to fully cover imbalances to ensure system reliability. After the proactive reserve activation/deactivation, the minimum available deactivation reserves vary between 427 MW/12 and 1662 MW/12 in each five minute interval for the four 2010 cases, and between 339 MW/12 and 1597 MW/12 for the 2020 cases. The minimum available activation reserves vary between 127 MW/12 and 645 MW/12 in each five minute interval for the four 2010 cases, and between 110 MW/12 and 456 MW/12 for the 2020 cases. The winter 2010 case, with different ramp rates, also has values in these ranges. Evidently, the increase in wind power reduces the reserve capacity available and thereby the flexibility of the power system. Nonetheless, we find that the available reserve power can cover the imbalances after the proactive reserve activation/deactivation in all time intervals in all cases except for one five minute interval in the Winter 2020 case. In this five minute interval the system has a surplus of 213.57 MWh/12 without the possibility to deactivate more reserves, but it can be covered by turning off a power plant. Thus, the SO does have options to cover the remaining imbalances in the system.
Figure 1.8: Two-hour schedule for activation of reserves on a selected unit, showing five separate reserve activations. Each column represents the amount of energy activated in a five minute interval, e.g. 70 MWh/12, corresponding to adding 70 MW to the production level in that interval.

1.7.4 Reserve activations

A closer look at the Winter 2010 base case shows additional results from the intra-hour model. First, Fig. 1.8 shows a given power generating unit for a two-hour time horizon. It can be seen how reserve power is activated in several overlapping blocks of the same length, i.e. $\tau^{\text{res}} = 6$ time intervals. Moreover, the figure shows how the unit ramps. This unit can ramp 70 MW from one five minute interval to the next and is seen to ramp to the limit several times. Deactivation of reserves is likewise utilised, as seen in Fig. 1.9 on the next page. The figure shows the re-scheduling of another power generating unit for a two-hour time horizon. Here, reserve power is deactivated in overlapping blocks. The ramp rate of the unit is 125.2 MW. These two examples confirm that the ability to ramp within the hour is extensively used, and is therefore highly important in short-term scheduling.

Finally, Fig. 1.10 on the facing page shows how a unit is rescheduled over a longer time horizon. In particular, the figure displays planned production and how reserves are activated and deactivated while the total production of the unit remains inside the minimum and maximum bounds for production.
Figure 1.9: Two-hour schedule for deactivation of reserves on a selected unit, showing five separate reserve deactivations. Each column represents the amount of energy deactivated in a five minute interval, e.g. 21 MWh/12, corresponding to producing at 21 MW below the planned production level in that interval.

Figure 1.10: Rescheduling of a selected conventional unit. The dark grey area is the planned production, the white area is activation of reserves, and the light grey area is deactivation of reserves. The two dashed lines show the minimum and maximum production levels of the unit.
1.8 Conclusion

In this paper, we formulated an intra-hour model that proactively re-dispatches the power generating units scheduled by a unit commitment model to account for imbalances in the system within the hour. Contrary to existing models from the literature, our model includes complex market rules for activation of reserves. Imbalances are caused by wind power forecast errors, ramping on power units and transmission lines, as well as intra-hour variations in demand. We generate a representative forecast of wind power production that will serve as input to the model. When investigating the influence of wind power forecast errors and ramping, we find that the ability to ramp within the hour is used extensively by the units in order to reduce imbalances. In spite of the cost of doing so, our results show that the benefits of growth in inexpensive wind power production is not outweighed by increased balancing cost. Moreover, we find that the proactive approach to re-dispatch the units is superior to the reactive approach, especially for the 2020 high wind cases. Thus, the approach is increasingly interesting for future power systems with the expected growth in wind penetrations.

Future work includes an extension of the model to consider stochastic wind power forecasts and to investigate the extent to which the results differ when comparing to our deterministic model with forecasts updated as frequently as once an hour. Furthermore, different solution methods may be explored to efficiently solve stochastic intra-hour models with high time resolution.

Acknowledgements

The authors gratefully appreciate many valuable comments and suggestions from Pierre Pinson from the Technical University of Denmark and Peter Meibom from the Danish Energy Agency, as well as discussions of the problem with Energinet.dk. Jeanne Andersen acknowledges support through the CFEM project and Trine Krogh Boomsma through the ENSYMORA project, both funded by the Danish Council of Strategic Research. Ditte Mølgård Heide-Jørgensen acknowledges the support through the iPover project.
Appendix

1.A Wind power forecasts: Parameters and correlations

Parameters of the Beta distributions

Consider the Beta distribution $\text{Be}(\alpha_t, \beta_t)$ with $\alpha_t > 0, \beta_t > 0$. To determine its parameters, we use the mean-variance model of Pinson et al. (2009).

The intra-hour mean is given by the linear interpolation between hourly values $\bar{w}_h$ and $\bar{w}_{h+1}$, and so

$$
\mu_t = \left( \frac{(60h - \tau t)}{(60 - \tau)} \right) \bar{w}_h + \left( \frac{(\tau (t-1) - 60(h-1))}{(60 - \tau)} \right) \bar{w}_{h+1}, \ t \in T_h.
$$

The variance is determined from the mean such that

$$
\sigma_t^2 = \theta_1 + \theta_2 \nu_t \mu_t (1 - \mu_t),
$$

where, for a constant mean, the function

$$
\nu_t = \sigma_0^2 \left( \frac{t_0 + t}{t_0 + T/\tau} \right)^\lambda,
$$

with $\lambda \in [0, 1]$ makes the variance an increasing and concave function of time. Here, $\sigma_0^2$ is a reference variance of the forecast at time $T/\tau$ (the end of the scheduling horizon), and $t_0$ is the time between generating and starting the forecast.

We use the parameter values $\sigma_0^2 = 0.1$ and $\lambda \in [0.4, 0.6]$ as in Pinson (2006), and $\theta_1 = 0.02$ and $\theta_2 = 4$, since the maximum variance (i.e. for $\mu_t = 0.5$ and $t = T/\tau$) should be close to $\sigma_0^2$. 

33
Finally, the parameters of the Beta distribution are computed using the identities\(^5\)

\[\mu_t = \frac{\alpha_t}{\alpha_t + \beta_t}, \quad \sigma^2_t = \frac{\alpha_t \beta_t}{(\alpha_t + \beta_t)^2(\alpha_t + \beta_t + 1)}.\]

**Correlation matrix**

Denote the \((T/\tau) \times (T/\tau)\)-correlation matrix by \(\Sigma = (\rho_{tt'}); t', t \in T\). With inspiration from Morales et al. (2009); Pinson and Girard (2012), we assume exponential decay of the correlations over time such that

\[\rho_{tt'} = \exp(-\phi(t' - t)(\tau/60)).\]

We use the parameter value \(\phi = 1/7\) as in Pinson and Girard (2012).

**1.B The effect of the cost of the automatic reserves**

**Cost of automatic reserves**

We investigate how the cost of balancing the system depends on the cost of automatic reserves. This is illustrated in Fig. 1.B.1. It is seen how total cost increase when activation of reserves

---

\(^5\text{For some } t, \text{ we occasionally obtain } \alpha_t \leq 0 \text{ or } \beta_t \leq 0, \text{ which means that the Beta distribution is not an accurate model for the data. In these cases we slightly modify } \mu_t \text{ or } \sigma^2_t.\)
becomes more expensive and deactivation becomes less expensive. By estimating the linear trend, we find the average total balancing cost to increase by 3.9%, when the reserve cost are increased by 10% (i.e. cost of activation of reserves increase by 10% and the cost of deactivation of reserves decrease by 10%).
Chapter 2

Stochastic model for short-term balancing of supply and consumption of electricity

Jeanne Andersen · Michal Kaut · Asgeir Tomasgard

J. Andersen: CORAL, Department of Economics and Business, Aarhus University, Fuglesangs Alle 4, 8210 Aarhus V, Denmark

M. Kaut: SINTEF Technology and Society, Applied Economics, S.P. Andersens vei 5, 7465 Trondheim, Norway

A. Tomasgard: Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology, 7491 Trondheim, Norway
Abstract
In this paper, we present a two-stage stochastic mixed integer model for the intra-hour balancing problem faced by system operators in electricity systems with large penetration of wind power production. Since wind power is non-controllable and intermittent, it is difficult to predict wind power production. Wind power prediction errors directly impact the balance between supply and demand of electricity. Therefore, it is of utmost importance for the system operators to understand and investigate these errors and plan accordingly when balancing the transmission system. In this model, we capture the uncertainty in wind power production forecasts by generating scenarios for the prediction errors. We apply the model on realistic Danish system data. We compare the stochastic solution and the deterministic solution to the solution of perfect foresight, and we find that wind power prediction errors entail huge balancing costs. Furthermore, we see that the stochastic solution incorporate a buffer when activating manual reserves compared to the deterministic solution. The buffer results in higher expected cost, but the actual cost incurred is lower compared to the deterministic solution in most of the cases.

Keywords: OR in energy, Scenario generation, Wind power prediction error, Power system balance

2.1 Introduction

With existing technology, electricity cannot be stored large-scale in any feasible way. Since it cannot be stored, the electricity has to be produced in the same second as it is consumed. It is the responsibility of the system operator (SO) to balance the system such that supply always equals consumption. With an increasing penetration of fluctuating renewable energy sources such as solar and wind power, the supply of electricity becomes highly uncertain. However, fluctuations can be met by planned conventional electricity production from thermal plants but only to the extent that the power from the fluctuating energy sources are forecasted accurately. When there is a prediction error, the excess or shortage of electricity must be handled.
through reserved power capacity on plants dedicated to this purpose. Thereby, the security of the electricity system depends on accessibility of reserved power capacity for increased or reduced electricity production.

The purpose of this paper is to develop a two-stage stochastic mixed integer model that can be used to analyse how to balance supply and demand of electricity within the hour seen from an SO’s perspective. The model is supposed to be used for analysing purposes, e.g. how different ramping patterns between flow levels from transmission lines effect the system balance or how different balancing policies result in different balancing costs.

Wind power production is one of the most promising large-scale renewables in Northern Europe to replace some of the conventional supply. Therefore, in this paper, we look at an electricity system with a large amount of installed wind capacity, and the stochasticity of the model lies in the uncertainty in the wind power production. In order to capture the uncertainty, we make scenarios for the wind power prediction error. The model is an extension of the deterministic Nordic balancing model presented in Andersen et al. (2014). A stochastic model can capture the uncertainty of wind and thereby give a more precise picture of intra-hour challenges related to wind power production.

When balancing an electricity system in the case of updated new information (e.g. new wind power forecasts), the SO can use two different strategies. Either they can be proactive and re-dispatch generating units by activating or deactivating reserved capacity before real-time operation, or they can wait until the imbalances occur and then re-dispatch the generating units by activating or deactivating reserved capacity at the time the imbalances occur. If the SO is proactive and able to forecast expected imbalances with high accuracy, smaller imbalances will occur real-time. This way balancing costs can be reduced since activation of reserved capacity is often cheaper the longer activation time, the generating unit has. We will refer to re-dispatching generating units before real-time operation as activating manual reserves and real-time activation or deactivation of reserved power capacity as activating automatic reserves.

In this paper, we assume that the SO is proactive and that the already committed units in the day-ahead market make their capacity for either additional or reduced electricity supply available to the SO. We consider balancing close to real-time operation when commitment sched-
ules, production plans for generating units, and forecasts for consumption and wind power production have been made and converted to an intra-hour time resolution. An example on how to convert hourly schedules and forecasts into an intra-hour time resolution can be seen in Andersen et al. (2014). On the basis of the intra-hour schedules and forecasts, the presented model re-dispatches the generating units before real-time operation in order to minimise balancing cost. It is assumed that the unit commitment schedule is fixed, and therefore our model only considers re-dispatch of already committed units.

Intra-hour balancing is closely related to the unit commitment (UC) problem, where start-ups, shut-downs, and production levels for the generating units are decided upon, usually with an hourly resolution. Stochastic UC models with scenarios for wind power production have been studied intensively in literature. Where some have made scenarios for the wind power production (e.g. Pappala et al. (2009)), others have made scenarios for the wind speed and afterwards converted wind speed to power (e.g. Papavasiliou and Oren (2013)). However, UC models with an hourly resolution do not capture fluctuations of wind power production within the hour. To get the detailed information about the system within the hour, intra-hour models can be used. Even though some intra-hour models have focused on uncertainty in wind power production (e.g. Lindgren and Söder (2008); Ela and O’Malley (2012)), only few have looked at intra-hour stochastic models for the balancing problem faced by SOs (e.g. Delikaraoglou et al. (2014)). However, to the best of our knowledge, no study has a stochastic balancing model as detailed as presented in this paper where technical ramping restrictions are included.

When making scenarios for wind power production in general, the emphasis has mainly been on wind power forecasts and not so much on the study of wind power prediction errors. Even though the wind power prediction error is a result of the forecast, understanding the nature of the error is extremely important for the SO since the error is one of the major sources for imbalances in an electricity system with high penetration of wind power. Studies (e.g. Doherty and O’Malley (2005); Menemenlis et al. (2012)) have been carried out on how the wind power prediction error affects the need for reserves and Menemenlis et al. (2012) show the importance of the wind power prediction error classification when determining the level of reserves.
It has often been assumed that the wind power prediction error can be described by a Gaussian distribution (e.g. see Doherty and O'Malley (2005); Bouffard and Galiana (2008)). However, as Bludszuweit et al. (2008) and Hodge and Milligan (2011) point out, the error distribution depends on the forecasting horizon and method. The prediction error also depends on the forecast being for a single wind turbine or for aggregated wind farms (Tewari et al. (2011)). The prediction errors from forecasts with a long forecast horizon and aggregated wind farms may be described by the Gaussian or Normal distribution where the distribution of prediction errors from forecasts with a shorter time horizon is heavy-tailed with variable kurtosis and hence cannot be described by the Gaussian distribution (Bludszuweit et al. (2008); Hodge and Milligan (2011)). To capture the heavy-tailed nature of the distribution of the short-term error, Bludszuweit et al. (2008) suggest the Beta distribution. Other distributions such as a Lévy alpha-stable distribution (Bruninx and Delarue (2014)), the Cauchy distribution (Hodge and Milligan (2011)), and the Gamma distribution (Menemenlis et al. (2012)) have also been suggested. Lately, Wu et al. (2014) proposed a mixed distribution based on the Laplace and Normal distribution to approximate the wind power prediction error. Lange (2005) takes another approach; assuming that the wind speed error to be Gaussian, they transform the wind speed error into wind power errors by Taylor-expansions.

In all of the before mentioned studies; the description of the correlation of error in time is absent. Pinson et al. (2009) address this problem by converting uniformly distributed series of wind power prediction errors to a multivariate Gaussian random variable, where the interdependence structure can be described by a unique covariance matrix. This matrix is recursively estimated in order to account for the variations in the characteristics of the errors. Ma et al. (2013) expand the work done in Pinson et al. (2009) by including empirical distributions of the prediction errors. In Litong-Palima et al. (2012), they account for the correlation of prediction errors in time by generating day-ahead and hour-ahead wind power prediction errors on a five minute resolution by an ARMA(1,1) process, where the independent random variable is assumed to follow a Gaussian distribution. The detailed resolution is obtained by using linear interpolation between the hourly wind power prediction errors. Söder (2004) and Weber et al. (2009), on the other hand, assume that the wind speed prediction error can be described
with an ARMA(1,1) process, where the independent random variable is assumed to follow a Gaussian distribution. They use a multidimensional ARMA(1,1) model to capture the positive correlation of the wind speed prediction error between different sites. In Matevosyan and Söder (2006), they use the approach described in Söder (2004) to generate different outcomes for the wind speed error and convert it to scenarios for the wind power production. The scenarios are used in a stochastic model which generates optimal wind power production bids from the wind farm owners to the power market.

In this paper, we use a different approach. We generate the outcomes of the error term using a copula-based heuristic from Kaut (2014). In particular, we use a nonparametric version of the method that attempts to replicate the distribution of the historical data, so we avoid extra assumptions about the underlying distributions.

To summarise, our contribution to the existing literature is twofold:

- We develop a stochastic two-stage model for the intra-hour balancing problem. The model can be used for analysing electricity systems operations within the hour, such as intra-hour variations, the effect of ramping patterns, or the minimum level of automatic reserves. With minor modifications, the model can be used in operational planning.

- We show how to make scenarios for the short-term wind power production by generating scenarios for the wind power prediction error by a copula-based heuristic.

The rest of the paper is divided into sections as follows. In the following section, we describe the stochastic balancing problem and present our model. Then, we describe our scenario-generation procedure in Section 2.3. Finally, we apply our model to a Danish case in Section 2.4.

### 2.2 The stochastic balancing model

The balancing problem described in this paper involves two sets of decisions. First, the manual reserves are activated based on expected imbalances, and afterwards, when the uncertain wind power prediction errors are revealed, the automatic reserves are activated. This can be seen as a two-stage stochastic model, where the SO decides on the level of manual reserves that have
to be activated in the first stage, and in the second stage, the automatic reserves are activated based on the actual imbalances caused by the wind power prediction errors.

In addition to the deterministic model presented in Andersen et al. (2014), we have in this model included uncertainty in the wind power forecast and made a simplification in the activation of manual reserves. However, for the readability and completeness of the model, we include all the notation and explanation of the first stage constraints, even though some of it is as presented in Andersen et al. (2014).

2.2.1 Notation

We start out by discretising the time horizon, \([0, T]\), of the stochastic model into \(\tau\)-minute intervals \([t-1]\tau, t\tau]\), \(t = 1, \ldots, T/\tau\). Let \(T = \{0, \ldots, T/\tau\}\). The time periods are further grouped into \(\Lambda\) groups, such that each group \(\lambda\), where \(1 \leq \lambda \leq \Lambda\), has \(|T_\lambda|\) time periods. Since we describe a two-stage model, we let \(\Lambda = 2\), a group for each stage. \(T\) then consists of the two subsets \(T_1\) and \(T_2\), where \(T_1 = \{0\}\) and \(T_2 = \{1, \ldots, T/\tau\}\). The first stage is assumed to be deterministic, so we have one stochastic stage \(T_2\) with \(T/\tau\) stochastic periods.

The transmission grid is modelled as a network \((\mathcal{N}, \mathcal{A})\) with nodes \(\mathcal{N} = \{1, \ldots, N\}\) and arcs \(\mathcal{A} = \{a : a = (n, n'), n, n' \in \mathcal{N} : n < n'\}\) representing transmission lines. We denote by \(\delta^{\text{out}}(n) = \{a : a = (n, n'), n \in \mathcal{N}\}\) and \(\delta^{\text{in}}(n) = \{a = (n', n), n' \in \mathcal{N}\}\) the sets of arcs either originating from or terminating in node \(n\), respectively. For \(a \in \mathcal{A}\), we let the capacity of transmission line \(a\) be \(L_{a,\max}\). Moreover, we let the flow allocated day-ahead to line \(a\) in time interval \(t\) be \(L_{a,t}\). If \(a = (n, n')\) and \(L_{a,t} > 0\) there is a net import from \(n\) to \(n'\). If \(L_{a,t} < 0\) there is a net export. We represent the intra-hour scheduled import determined by our model on the transmission line in the same time interval by the variable \(\Delta l_{a,t}\) using the same conventions regarding its sign. The allowed change in the flow on a transmission line between two time periods will be denoted by \(R_a\).

The set of all conventional units is denoted by \(\mathcal{I}\), the set of units online in time interval \(t\) is denoted \(\mathcal{I}_t\), and the units located in node \(n\) by \(\mathcal{I}_n\). For \(i \in \mathcal{I}\), we denote the planned production day-ahead in time interval \(t\) by the parameter \(P_{i,t}\), and the plant’s minimum and maximum production limits by \(P_{i,\min}\) and \(P_{i,\max}\), respectively. The model will minimise expected imbalances.
by doing short-term production planning, and it decides, based on updated weather forecasts, to produce $P_{it} + \Delta p_{it}^+ - \Delta p_{it}^-$, instead of $P_{it}$. The first-stage decision variables $\Delta p_{it}^+ \geq 0$ and $\Delta p_{it}^- \geq 0$ represent the balancing power provided by intra-hour activation of manual reserves on the unit in the time interval. When doing short-term production planning, we must for each unit $i$ and time interval $t$ obey

$$P_{i}^{p_{\min}} \leq P_{it} + \Delta p_{it}^+ - \Delta p_{it}^- \leq P_{i}^{p_{\max}}.$$  

However, in real time, there will still be imbalances that have to be dealt with using the automatic reserves. The extra produced or reduced power is illustrated by the second stage variables $q_{nts}^+ \geq 0$ and $q_{nts}^- \geq 0$, which represent the generation surplus or shortage in node $n$ in scenario $s \in S$ during time interval $t$. Note, that we do not have any limits on the production from automatic reserves. Firstly, there exist plants whose only purpose is to supply automatic reserves and hence are not included in the set $I$. Secondly, by not restricting the second stage variables, we are guaranteed to always have a feasible solution to the problem. Hence, the model can indicate which level of automatic reserves that are needed in order to maintain a secure system.

We denote by $C_i$ the variable generation cost of unit $i \in I$. The late scheduling of the units close to real time operation incurs extra variable cost $\gamma C_i$. Note, whereas activation of power generates a cost, deactivation of power results in cost savings from not producing. We therefore let $C_i^+ := (1 + \gamma) c_i$ and $C_i^- := (1 - \gamma) c_i$, with $\gamma \in [0, 1]$, be the costs of activated manual reserve power and savings from deactivated manual reserve power, respectively. The idea is to allow the costs and savings to reflect the additional stress imposed on the unit when using it for balancing purposes. Since $P_{it}$ is already decided, we do not include its costs in the model. The costs of activating automatic reserves in node $n$ are denoted by $C_n^+$ and $C_n^-$, where most likely $C_n^+ > \max_i C_i^+$ and $C_n^- < \min_i C_i^-$. In other words, it is likely that there is a higher cost for extra production and lower savings for production decrease in the case of automatic reserves compared to the manual reserves.

In order to keep track of the activated manual reserves, we let the variables $g_{it}^+$ and $g_{it}^-$ be the amount of recently activated and deactivated power on unit $i$ in time interval $t$, respectively. If new manual reserves are activated on a unit, it has to provide the committed level for at
least $\tau_{\text{res}}$ time periods. If we activate or deactivate manual reserve power, we let the associated binary variables $v^+_i$ or $v^-_i$ be one and zero otherwise. The amount of recently activated or deactivated manual reserve power has to be above the threshold value $G_{i\min}$. Furthermore, we introduce the variables $g^\text{ramp,}^+_i$ and $g^\text{ramp,}^-_i$ to record the ramping power generated by a unit providing manual reserves, where $g^\text{ramp,}^+_i$ is related to activating power and $g^\text{ramp,}^-_i$ to deactivating power. How much the production level is allowed to change from one time period to another will be denoted by $R^+_i$ when the unit is ramping up, and by $R^-_i$ when the unit is ramping down. A unit is only allowed to ramp $\tau_{\text{max}}$ time periods before a new activation level, and $\tau_{\text{max}}$ time periods after the activation period of $\tau_{\text{res}}$ time periods.

We assume that the demand is inflexible, meaning that the level of demand in each time period is maintained at the forecasted level. We denote by the parameter $D_{nt}$ its value in node $n$ in time interval $t$.

The uncertainty is modelled using scenarios $s \in \mathcal{S} = \{1, \ldots, S\}$ with probability $\Pi^s$ and $\sum_{s \in \mathcal{S}} \Pi^s = 1$. The only stochastic parameter in the model is the wind power production $\omega^s_{nt}$ for $t \in T_2$, with the first-stage deterministic values denoted by $\omega_{nt}$ for $t \in T_1$. This results in the scenario tree presented in Fig. 2.1, where we have omitted the node subscripts, i.e. all the values are vectors of size $N$.

### Figure 2.1: Scenario tree.

#### 2.2.2 Objective function

We schedule the activation of manual reserves $\Delta p^+_i$ and $\Delta p^-_i$, such as to cover any expected imbalances between supply and consumption. Occasionally, this may be technically infeasible,
or it may be feasible only at very high costs, in which case imbalances are left to automatic reserves $q_{nts}^+$ and $q_{nts}^-$. Since the uncertainty in wind power is revealed after activation of manual reserves, the amount of automatic reserves depends on the realisation of the wind power prediction error, and hence, we have an outcome with a probability $\Pi^s$ for each scenario. The optimal schedule is determined by a trade-off between the activation cost of manual and automatic reserves. The objective is

$$
\sum_{t \in T_2} \left( \sum_{i \in I_t} (C_i^+ \Delta p_{it}^+ - C_i^- \Delta p_{it}^-) + \sum_{s \in S} \Pi^s \left( \sum_{n \in N} (C_n^+ q_{nts}^+ - C_n^- q_{nts}^-) \right) \right),
$$

which is minimised subject to a number of constraints presented next.

### 2.2.3 Balancing constraint

The balancing constraint ensures system balance between supply and consumption. According to this constraint, if at any point in time scheduled production exceeds predicted consumption $D_{nt}$ or vice versa, we experience generation surplus or shortage, which will be left to the automatic reserves $q_{nts}^+$ and $q_{nts}^-$. We assume that it is always possible to provide the sufficient amount of automatic reserves. Production includes day-ahead planned generation on conventional units $P_{it}$, intra-hour activation of manual reserves $\Delta p_{it}^+$ and $\Delta p_{it}^-$, forecasted wind power production $\omega_{nts}$ and finally day-ahead and intra-hour net import/export on the transmission lines $L_{at}$ and $\Delta l_{at}$. Thus, we have that

$$
\sum_{i \in I_t \cap I_n} (P_{it} + \Delta p_{it}^+ - \Delta p_{it}^-) + \sum_{a \in \delta^+(n)} (L_{at} + \Delta l_{at}) - \sum_{a \in \delta^-(n)} (L_{at} + \Delta l_{at}) + q_{nts}^+ - q_{nts}^- = D_{nt} - \omega_{nts}, \quad n \in N, t \in T_2, s \in S. \quad (2.1)
$$

### 2.2.4 Limits on re-dispatch variables

Transmission flow is restricted by the available line capacity. In particular, intra-hour import $\Delta l_{at}$ on the transmission lines is bounded above by the line capacity $L_{at}^{\max}$ minus the capacity allocated day-ahead $L_{at}$. Thus, we have that

$$
-(L_{at}^{\max} - L_{at}) \leq \Delta l_{at} \leq L_{at}^{\max} - L_{at}, \quad a \in A, t \in T_2.
$$
Activation of manual reserve power is bounded above by the maximum capacity $P_{i \text{max}}^i$ that has not already been dispatched day-ahead $P_{i \text{it}}$, whereas deactivation is bounded by the dispatched capacity in excess of the minimum capacity $P_{i \text{min}}^i$. Formally,
\[
\Delta p_{i \text{it}}^+ \leq P_{i \text{max}}^i - P_{i \text{it}}, \quad i \in I, t \in T_2,
\]
\[
\Delta p_{i \text{it}}^- \leq P_{i \text{it}} - P_{i \text{min}}^i, \quad i \in I, t \in T_2.
\]

### 2.2.5 New generation level of a unit

We require the minimum threshold value $G_{i \text{min}}^i$ to be obtained when activating manual reserves. The threshold value is included in order to make it profitable for the unit to change the production level. Manual reserves are activated when $v_{i \text{it}}^+ (v_{i \text{it}}^-)$ is equal to one at the level $g_{i \text{it}}^+ (g_{i \text{it}}^-)$.

\[
G_{i \text{min}}^i v_{i \text{it}}^+ \leq g_{i \text{it}}^+ \leq M v_{i \text{it}}^+, \quad i \in I, t \in T_2, \tag{2.2}
\]
\[
G_{i \text{min}}^i v_{i \text{it}}^- \leq g_{i \text{it}}^- \leq M v_{i \text{it}}^-, \quad i \in I, t \in T_2, \tag{2.3}
\]

where $M$ is a sufficiently large number, e.g. $P_{i \text{max}}^i$.

It is only allowed to either activate or deactivate manual reserve power on the same unit in one time interval, hence,

\[
v_{i \text{it}}^+ + v_{i \text{it}}^- \leq 1, \quad i \in I, t \in T_2. \tag{2.4}
\]

Additional levels of reserves may be activated on the same unit at a given point in time. This we refer to as activating new reserves. The amount of manual reserves a unit provides at a given point in time can be calculated by

\[
\Delta p_{i \text{it}}^+ = \sum_{t' = \max\{1, t - \tau_{\text{res}} + 1\}}^{t} g_{i \text{it}}^+ + g_{i \text{it}}^{\text{ramp},+}, \quad i \in I, t \in T_2, \tag{2.5}
\]
\[
\Delta p_{i \text{it}}^- = \sum_{t' = \max\{1, t - \tau_{\text{res}} + 1\}}^{t} g_{i \text{it}}^- + g_{i \text{it}}^{\text{ramp},-}, \quad i \in I, t \in T_2, \tag{2.6}
\]

where $\tau_{\text{res}}$ is the fixed time from time $t$ and forward, for which the activated reserves have to provide the level $g_{i \text{it}}^+ (g_{i \text{it}}^-)$ of power. The actual manual reserves provided by a unit are the activation level plus the ramping power $g_{i \text{it}}^{\text{ramp},+} (g_{i \text{it}}^{\text{ramp},-})$ it provides in order to reach the agreed level.
2.2.6 Ramping

Ramping constraints on the transmission lines are restrictions on the change in allocated transmission flow from one time interval to another and apply to net import. In order to record the change from one time interval to another, we need to include both the day-ahead agreed flow amount $L_{at}$ and the variation decided by our model $\Delta l_{at}$ from the time intervals. Thus,

$$- R_a - L_{at(t+1)} + L_{at} \leq \Delta l_{at(t+1)} = R_a - L_{at(t+1)} + L_{at},$$

$$a \in A, t \in T_2 : t \leq |T_2| - 1.$$

It is allowed to change the flow on the transmission line by $R_a$ between two time intervals.

For the generating units, we include detailed ramping restrictions where the change in day-ahead planned production $P_{it}$ is taken into account.

$$-(R_i^- + P_{it(t+1)} - P_{it}) \leq \Delta p_{it(t+1)}^- \leq (R_i^+ - P_{it(t+1)} + P_{it}),$$

$$i \in I, t \in T_2 : t \leq |T_2| - 1,$$

$$-(R_i^+ - P_{it(t+1)} + P_{it}) \leq \Delta p_{it(t+1)}^+ \leq (R_i^- + P_{it(t+1)} - P_{it}),$$

$$i \in I, t \in T_2 : t \leq |T_2| - 1.$$

The allowed ramping speed between two time intervals is $R^+$ for activation of power, where it is $R^-$ for deactivation.

Generating units can only ramp towards a new level in $\tau^{\text{max}}$ time periods before the amount of activated reserves has to be fully activated and in the $\tau^{\text{max}}$ time periods after the activation period of $\tau^\text{res}$ time periods.

$$g_{it}^{-\text{ramp}} \leq M \min\{T_{t+\tau^{\text{max}}}, \tau^\text{res}+1\} v_{it}^+ \leq M \max\{1, t-\tau^{\text{res}}\} v_{it}^-,$$

$$i \in I, t \in T_2,$$

$$g_{it}^{+\text{ramp}} \leq M \min\{T_{t+\tau^{\text{max}}}, \tau^\text{res}+1\} v_{it}^- \leq M \max\{1, t-\tau^{\text{res}}\} v_{it}^+,$$

$$i \in I, t \in T_2,$$

where $M$ is a sufficiently large number, e.g. $P_{it}^{\text{max}}$. If $v_{it}^+$ ($v_{it}^-$) is equal to 1 in the constraint, it indicates that an activation (deactivation) of manual reserve power has occurred at time $t'$.
thereby it is possible for the unit to ramp by activating (deactivating) \( g_{it}^{ramp,+} \) \( g_{it}^{ramp,-} \) power in the \( \tau_{\text{max}} \) time periods before time \( t' \) and after time \( t' + \tau_{\text{res}} \).

If new reserves are activated, and thereby a new level of production is committed on a unit in one time period, ramping is not allowed in the same period on the same unit. Hence,

\[
\begin{align*}
    g_{it}^{ramp,+} &\leq M(1 - v_{it}^+), \quad i \in I_t, t \in T_2, \\
    g_{it}^{ramp,-} &\leq M(1 - v_{it}^-), \quad i \in I_t, t \in T_2,
\end{align*}
\]

(2.11) (2.12)

where \( M \) is a sufficiently large number, e.g. \( P_{i}^{\text{max}} \).

Since ramping in optimisation can adapt to the imbalances, we restrict the ramping to a linear pattern.

\[
\begin{align*}
    g_{it}^{ramp,+} &= \frac{\Delta p_{i(t-1)}^+ + \Delta p_{i(t+1)}^+}{2}, \quad i \in I_t, t \in T_2 : 1 < t \leq |T_2| - 1, \\
    g_{it}^{ramp,-} &= \frac{\Delta p_{i(t-1)}^- + \Delta p_{i(t+1)}^-}{2}, \quad i \in I_t, t \in T_2 : 1 < t \leq |T_2| - 1.
\end{align*}
\]

(2.13) (2.14)

Hence, the level of manual reserves \( \Delta p_{it}^+ \) and \( \Delta p_{it}^- \) in time period \( t \) depends on the previous and the subsequent levels.

### 2.3 Generating scenarios for the wind power production

In this section, we describe the process of generating values for the stochastic wind power production \( \omega_{nt}^s \) for \( n \in N, s \in S \), and \( t \in T_2 \). For this, we assume that we have, at time \( t = 0 \), access to wind power forecast \( \hat{\omega}_{nt} \) for all \( t \in T_2 \); this forecast is typically obtained from a third party. The scenario values for wind power production, \( \omega_{nt}^s \), are then computed as

\[
\omega_{nt}^s = \hat{\omega}_{nt} + \epsilon_{nt}^s,
\]

(2.15)

where \( \epsilon_{nt}^s \) is the prediction error in period \( t \) and scenario \( s \). This implies that we first create the scenario tree for the prediction errors and then combine these with the forecast to get the wind production values; this is illustrated graphically in Fig. 2.2 on the following page.

The prediction error is a result of the imprecision of the wind power production forecast. Its statistical properties depend on the length of the forecast and need to be estimated from
historical data. For this purpose, we assume that we have access to historical data for both the production and the forecasts for all the forecast lengths \( \Delta t \) used in the model: \( \Delta t = \tau t \) with \( t \in T_2 = \{1, \ldots, T/\tau\} \). We use a tilde to distinguish the historical data from the model parameters, so \( \tilde{\omega}_{n,t} \) denotes the historical wind production at some time \( t < 0 \), \( \tilde{\epsilon}_{n,t+\Delta t|t} \) stands for the forecast for time \( t + \Delta t \) made at time \( t \), and \( \tilde{\epsilon}_{n,t+\Delta t|t} \) is the error of this forecast, i.e.

\[
\tilde{\epsilon}_{n,t+\Delta t|t} = \tilde{\omega}_{n,t+\Delta t|t} - \tilde{\omega}_{n,t+\Delta t|t}.
\]

Once we have computed the historical errors \( \tilde{\epsilon}_{n,t+\Delta t|t} \), we can treat them as data series for each \( n \) and \( \Delta t \). Hence, we have a historical data set with dimension \( |\mathcal{N}| \times |T_2| \) for which we want to generate scenarios. Before we delve into the details about how we generate these scenarios, we need to explain how we use the generated values. Since each \( \Delta t \) is a separate data series, we only need to generate values one period ahead. To explain why, let us denote the generated values by \( \epsilon_{n,\Delta t} \). In the model, on the other hand, we need \( \epsilon_{nt} = \epsilon_{n,\tau t} \), used in (2.15) to compute the wind production values \( \hat{\omega}_{nt} \). This way, we transform scenarios for \( |\mathcal{N}| \times |T_2| \)
variables one period ahead into scenarios for $|\mathcal{N}|$ variables $|\mathcal{T}_2|$ periods ahead. This implies, that if $\bar{\tilde{\varepsilon}}_{n,\tau \Delta t}$ correctly captures the $|\mathcal{N}| \times |\mathcal{T}_2|$-dimensional distribution, then $\tilde{\varepsilon}_{n,t}$ will correctly describe not only the dependencies between nodes within each period, but also inter-temporal dependencies such as autocorrelations.

To achieve this, we generate the scenarios using an algorithm from Kaut (2014). This method generates scenarios that try to replicate a specified multivariate distribution; in our case provided by the historical data for forecast errors $\tilde{\varepsilon}_{n,t+\Delta t|t}$. It does so by matching all the univariate distribution functions plus the bivariate copulas of all the variable pairs. For the purpose of this paper, a copula can be thought of as a generalisation of the linear correlation, which can fully describe the dependence between stochastic variables – unlike correlations that only capture linear dependencies. See Nelsen (1998) for more information about copulas, and Kaut and Wallace (2011) for a discussion about their use in scenario generation.

With $|\mathcal{N}| \times |\mathcal{T}_2|$ variables, matching copulas for all the $|\mathcal{N}| \times |\mathcal{T}_2| (|\mathcal{N}| \times |\mathcal{T}_2| - 1) / 2$ variable pairs becomes impractical. By trying to match copulas for so many pairs, the approximation error can be expected to be significant, especially since we can only solve the model with a couple of hundreds scenarios for realistic instances. For this reason, we concentrate only on the most important pairs of variables: for each forecast length $\Delta t$, we match the copulas between all the locations/nodes $n$, and for each node $n$, we match the copulas between forecasts of similar length, specifically for $0 < \Delta t_2 - \Delta t_1 \leq U \tau$. There, $U \geq 1$ is a case-dependent constant, whose value has to be adjusted to the data and also the generated number of scenarios. This reduces the number of bivariate copulas to

$$|\mathcal{T}_2| \times \frac{|\mathcal{N}| \times (|\mathcal{N}| - 1)}{2} + |\mathcal{N}| \times \sum_{l=1}^{U} (|\mathcal{T}_2| - l),$$

(2.16)
a reduction by the factor of more than $T/(1 + 2U/N)$. Naturally, there is a price to pay for this reduction: when we do not specify the dependence between a given pair of variables, it does not mean that the two variables become independent. It simply means that the algorithm does not have any control over the dependence, except for the constraints implied by the other pairs. For this reason, we need to check the quality of the scenario-generation procedure and its suitability for our optimisation model prior to using it on real data. In our case, these
checks indicate that the scenarios-generation procedure performs satisfactorily, as shown in Section 2.4.2.

2.4 Real-world problem

To investigate the usefulness of the model, we apply it on realistic data from Denmark. We look at two areas, Denmark West and Denmark East, and optimise the use of reserves during four weeks in 2012: a week in January, a week in April, a week in July, and finally, a week in October. We apply the model in a rolling horizon manner where we solve the model with a two hour time horizon every hour. The two hour time horizon is chosen to avoid undesired end effects. We will use the result for the first hour and disregard the result from the second hour, since the second hour will be re-optimised in the subsequent run of the model. The initialisation of each run is based on the state of the last optimised hour that is not disregarded. By utilising rolling horizon, we can take updated information about the system into consideration, and in our case, we get updated wind power production forecasts every hour. Our test instances each cover a week, and thereby they each consist of 168 hours. However, we will only report of 167 hours, since we use the first hour of each case for initialisation. The time resolution of the model, meaning each time interval $|\tau(t-1), \tau t|, t \in T_2$, will represent five minutes.

In each of the four cases, we will compare the cost of balancing the system when having scenarios for the wind power prediction error to the cost of balancing the system when assuming the prediction error to be zero, i.e. to the deterministic case. We investigate how the solution differs in the two cases. Furthermore, in order to show that it makes sense to predict the imbalances in the system and be proactive, we calculate the cost of balancing the system in each of the four weeks by only using automatic reserves. Finally, we investigate what the cost in the four weeks would have been if we were able to predict the actual wind power production, i.e. in the case of perfect foresight. The solution of perfect foresight is a lower bound on the actual cost of balancing the system, which the stochastic solution will never be able to match, but it indicates how far we are from the perfect solution.

We run the model on a computer with an Intel Core 2.30GHz processor and 4 GB RAM. The model is implemented in the 64 bit GAMS framework version 24.2.3 for Windows using
Table 2.1: Aggregated information about the system for four different weeks in year 2012 used as input to our model. The numbers represent the mean (standard deviation) of the 2016 time periods in the optimisation period, and hence is calculated in MWh/12. Note, capacity is showed in MW and day-ahead forecasts in MWh.

<table>
<thead>
<tr>
<th></th>
<th>January</th>
<th>April</th>
<th>July</th>
<th>October</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demand (MWh/12)</td>
<td>4520(863)</td>
<td>3940(644)</td>
<td>3537 (625)</td>
<td>3976(731)</td>
</tr>
<tr>
<td>Conv. prod. (MWh/12)</td>
<td>3363(527)</td>
<td>2571(890)</td>
<td>940 (524)</td>
<td>2342(763)</td>
</tr>
<tr>
<td>Online capacity (MW)</td>
<td>5634(249)</td>
<td>5299(337)</td>
<td>3188(1289)</td>
<td>5205(331)</td>
</tr>
<tr>
<td>Wind power prod. forecast DA (MWh)</td>
<td>2054(836)</td>
<td>1005(668)</td>
<td>949 (488)</td>
<td>894(791)</td>
</tr>
<tr>
<td>Wind power prod. forecast HA (MWh/12)</td>
<td>1958(812)</td>
<td>999(662)</td>
<td>927 (529)</td>
<td>927(785)</td>
</tr>
<tr>
<td>Wind power prod. (MWh/12)</td>
<td>1914(811)</td>
<td>996(660)</td>
<td>921 (526)</td>
<td>934(786)</td>
</tr>
<tr>
<td>Installed wind capacity (MW)</td>
<td>3926</td>
<td>3969</td>
<td>4025</td>
<td>4038</td>
</tr>
<tr>
<td>Import (MWh/12)</td>
<td>931(637)</td>
<td>1811(637)</td>
<td>2322 (443)</td>
<td>1750(801)</td>
</tr>
<tr>
<td>Export (MWh/12)</td>
<td>1698(584)</td>
<td>1478(512)</td>
<td>608 (250)</td>
<td>943(748)</td>
</tr>
</tbody>
</table>

the CPLEX 12.6 solver.

2.4.1 Data

We get unit commitment data from the unit commitment model, Sivael, used by the Danish transmission system operator, Energinet.dk. Real data for the Danish system in 2012 are given as input to the unit commitment model. Afterwards, we use an intra-hour model, SimBa from Energinet.dk, to convert the hourly output from the unit commitment model into a time resolution of five minutes. A description of SimBa can be found in Hansen et al. (2011).

Aggregated overall information about demand, conventional production, conventional capacity, wind power production forecasts, actual wind power production, installed capacity, import, and export can be found in Table 2.1. **Online capacity** is conventional capacity disposal for the market, meaning started up capacity where some of it is already committed by the unit commitment model, the rest of the capacity can be used as reserves for additional electricity production in our model. **Wind power prod. forecast DA** is the forecast made to clear the day-ahead market used in the unit commitment model, where the **wind power prod. forecast HA** is the updated wind power production forecast used in our model. The updated wind power forecasts used are made just before the hour of operation.

In order to resemble real life balancing in the Nordic countries, we follow present market rules and let the ramping period, $\tau^{\text{max}}$, for import, export, and conventional production units
be 15 minutes. Having a time resolution of five minutes implies $\tau^{\text{max}} = 3$. Once a generating conventional unit has been re-dispatched, the unit commits to the change in 30 minutes, i.e $\tau^{\text{res}} = 6$. Finally, the unit is guaranteed to have a re-dispatch amount greater than 10 MWh/12, i.e. $G_i^\text{min} = 10$ for all $i \in I$.

The costs of manual reserves are the individual marginal costs of the generating units altered with the parameter $\gamma$. The level of marginal costs of the generating units are within the range €2.27-205.73, where $\gamma$ is set to 0.1 to represent the additional stress imposed on the unit. Recall, that for up-regulating cost the marginal cost of a unit is multiplied by $1 + \gamma$, where for down-regulating cost the marginal cost of a unit is multiplied by $1 - \gamma$.

In practice, the costs for automatic reserves are not known when optimising and activating manual reserves for the next hour. To run our model, we need to estimate a fixed number for the cost of the automatic reserves, and hence we base the level on historical data from 2012. We let the 95% fractile for the historical cost of manually activating additional electricity in the balancing market be the cost for the automatic reserves when additional electricity is activated. Likewise, we let the 5% fractile for the historical cost in the balancing market of manually deactivating power be the cost for the automatic reserves in the case of deactivation of power. Hence, in our cases, we will assume the cost of activating automatic reserve power to be €75 and the cost of deactivating automatic reserve power to be €10.

### 2.4.2 Scenarios for the wind power production

The wind power production scenarios are generated as described in Section 2.3. Data input to the procedure comes from historical data for the wind power production in Denmark during 2012. The data consists of information about the installed capacity, the actual wind power production, the forecasts made day-ahead and the updated forecasts made just before the hour of operation where the system had to be balanced. Subsequently, we have calculated the historical prediction errors normalised with installed capacity, which is the actual input to the scenario generation procedure. The time resolution of all these time series is five minutes.

Our data analysis of the historical wind data shows a significant difference in the prediction errors between onshore and offshore wind power production, as documented in Fig. 2.A.1 and
Fig. 2.3: The graph to the left shows 50 generated scenarios for onshore Denmark West from the scenario generation procedure for four hours where the graph to the right shows 50 generated scenarios for two hours during another time period. The gray line shows the wind production forecast, and the thick black line shows the actual wind power production.

Fig. 2.A.2 in Appendix 2.A on page 65. For this reason, we separate the onshore and offshore wind power production in both the studied regions, so we end up with four locations ($N = 4$). Since we solve the model for two hours ahead with five minute steps, we have $T = 120$, $\tau = 5$, and $|T_2| = T/\tau = 24$. As a result, the model has $N \times |T_2| = 96$ random variables.

It follows that there are 4560 bivariate copulas to match – too many to get a good match, especially since we can solve the model with a couple of hundreds of scenarios at most. We therefore only match a subset of the copulas as discussed in Section 2.3. After studying the data, we have decided to use $U = 2$, i.e. to model dependencies between forecast errors with forecasts lengths that differ by two periods at most. Using (2.16), we see that this reduces the number of copulas to 324; a far more manageable amount.

In order to construct the scenarios, we need exogenous given wind power forecasts, see Fig. 2.2 on page 50. For this, we use the historical updated wind power forecasts from 2012.

Fig. 2.3 shows two graphs. The graph on the left shows 50 scenarios generated by the described scenario generation method. As expected for a short-time forecast, most of the sce-
narios lie around the forecast (the gray line) with some extreme scenarios farther apart. In most cases, also the actual wind production lies within the range of the scenarios. In rare occasions, however, the forecast is wrong from the start, so the actual wind production will be outside of the range covered by the scenarios, especially at the start of the horizon. This is illustrated by the graph to the right.

![Scenarios for one week](image)

**Figure 2.4:** The graph shows 50 generated scenarios for onshore Denmark West for a whole week.

The graph in Fig. 2.4 shows 50 scenarios for a whole week. Since the model is re-optimised every hour when new scenarios are generated, the scenarios will be close to the new updated forecast every hour, whereafter they will spread out and gather again in the beginning of the next hour.

**In-sample and out-of-sample stability tests**

Stability and high accuracy of the scenario generation procedure are important to ensure consistency and high quality of the solutions coming from the intra-hour model. In-sample and out-of-sample stability tests, as described in Kaut and Wallace (2007), can be used to measure
the quality of the scenario generation procedure in relation to the model. According to King and Wallace (2012), there are at least two ways to perform these tests. Which tests to use depends on the scenario generation procedure itself which will be explained below.

For in-sample stability, it is important to ensure that the objective function value of the model is approximately the same each time it is run on the same data. There are at least two ways to perform tests for in-sample stability. First, if the scenario generation procedure generates different scenario trees, it is important to test if different trees of the same size generate the same objective function value. Second, if the scenario generation procedure generates the same tree each time, then trees of different sizes have to be investigated. In order to have in-sample stability in the latter of the two cases, the objective function value should not change when changing the tree size with a small amount.

For out-of-sample stability, the value of implementing the first-stage variables and optimise the second-stage variables with respect to the true distribution of the random variables should not change between different runs. Again, when investigating the stability, the stochasticity of the scenario generation procedure has to be taken into account as described under in-sample stability.

Another important factor to investigate is the bias of the model. If the objective function value of the model when it is optimised over the entire distribution of the random variables is different from the objective function values given by the model when optimised over the scenario trees, then the scenarios do not represent the underlying distribution well enough.

The copula-based method we use for generating the scenarios for the error terms is somewhere between the two cases described above: since it is probabilistic repeated runs with the same input parameters, it typically results in very few distinct trees. For this reason, we combine the two approaches to stability and run the intra-hour model 10 times for the following number of scenarios: 1, 5, 10, 20, 30, 50, 75, 100, 125, 150, 175, and 200. However, if stability is achieved before we have investigated all the tree sizes, we stop. The case with 1 scenario is the deterministic variant of the model.

For the stability tests, we report the following values: (V1) is the in-sample objective value, i.e. the objective value of the scenario-based model. The out-of-sample value (V2) is computed
Figure 2.5: In this figure, we see two graphs showing the objective function value for our in- and out-of-sample stability tests for the two studies months. • represents values from (V1), ○ represents values from (V2), and the black line represents values from (V3).

by fixing the first-stage decisions and re-optimising using all the historical data as our scenarios. In order to measure the bias, we also calculate the value (V3) for the objective value of the model optimised over all the historical data. Note, that we are able to calculate (V3) only because we investigate these two small cases. It will be too time consuming to apply the model on the full cases with all the historical data for the prediction errors. In the stability tests, we should optimally see both (V1) and (V2) stabilising and approaching (V3) with the increasing number of scenarios.

We run the model with a two hour time horizon. Due to the complexity of the model, we have chosen two hours in January and two hours in July as representative cases for the full instances. The January two-hour case is difficult for the model to solve, while the July two-hour case is solved easily. Since the model is normally run with rolling horizon where the last hour of a model run is re-optimised in the subsequent run of the model, the output function of the model is only implemented for the first hour. Hence, we report the costs in these test only for the first hour.

Fig. 2.5 shows the results of the tests. The figure shows that, for a given size of the scenario tree, the scenario generation procedure is indeed stable: for most of the tree sizes, we only
see one dot in the figure for each of the measurements (V1) and (V2). However, running the model with different tree sizes shows some instability. Looking at the graph to the left in Fig. 2.5, we can see that the values improve fast when including up until 50 scenarios. After 50 scenarios, the values are approaching each other very slowly and the values of the three different measurements lie relatively close. However, with 200 scenarios, the values of the three measurements are still not exactly the same. Looking at the graph on the right, we see that the three measurements give approximately the same value already when including 30 scenarios.

We choose to run our full test instances with 50 scenarios. The tests show that the objective function value is reasonably stable around 50 scenarios, while the solution time is still manageable: it would be too time consuming to run the full tests with the 200 scenarios that are required for significantly better stability.

2.4.3 Results

In this section, we will present the obtained results. We will present results from four different versions of the model presented in Section 2.2. The first version is the stochastic model, where we apply the scenario generation procedure described in Section 2.3 to make 50 scenarios for the wind power production for each of the individual hours in the test instances. The second version is the deterministic model, where we only have one scenario for the wind power production and that is the expected value, i.e. where the expectation of the error-term is zero. The third version is the case of perfect foresight, where we also only have one scenario for the wind power production, but this scenario is the actual realised wind power production. The fourth and final version is also the stochastic model, but with an extra constraint forcing the manual reserves to zero, such that only automatic reserves are utilised.

Table 2.2 on the next page displays the average objective function value (which we will call average expected cost) and its standard deviation over the individual hours in the investigated weeks for each of three models: the stochastic, the deterministic, and the case of perfect foresight. The high standard deviation indicates significant variability in the costs between individual hours. This is further supported by Fig. 2.6 on page 61, which shows the expected
Table 2.2: Average objective function values and actual values in € and CPU time of the cases. The total CPU time is displayed hours first and then additional minutes (h:m).

<table>
<thead>
<tr>
<th>Case</th>
<th>Model</th>
<th>Objective value</th>
<th>Actual cost</th>
<th>CPU time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Average</td>
<td>Std.</td>
<td>Average</td>
</tr>
<tr>
<td>January</td>
<td>Stochastic</td>
<td>2751</td>
<td>16 862</td>
<td>4765</td>
</tr>
<tr>
<td></td>
<td>Deterministic</td>
<td>1465</td>
<td>16 892</td>
<td>5269</td>
</tr>
<tr>
<td></td>
<td>Perfect foresight</td>
<td>2123</td>
<td>17 225</td>
<td>2123</td>
</tr>
<tr>
<td>April</td>
<td>Stochastic</td>
<td>6844</td>
<td>14 069</td>
<td>7338</td>
</tr>
<tr>
<td></td>
<td>Deterministic</td>
<td>5598</td>
<td>14 144</td>
<td>7639</td>
</tr>
<tr>
<td></td>
<td>Perfect foresight</td>
<td>5643</td>
<td>13 774</td>
<td>5643</td>
</tr>
<tr>
<td>July</td>
<td>Stochastic</td>
<td>−82</td>
<td>7274</td>
<td>405</td>
</tr>
<tr>
<td></td>
<td>Deterministic</td>
<td>−1392</td>
<td>7052</td>
<td>727</td>
</tr>
<tr>
<td></td>
<td>Perfect foresight</td>
<td>−1383</td>
<td>7087</td>
<td>−1383</td>
</tr>
<tr>
<td>October</td>
<td>Stochastic</td>
<td>−3442</td>
<td>8102</td>
<td>−3276</td>
</tr>
<tr>
<td></td>
<td>Deterministic</td>
<td>−4883</td>
<td>8089</td>
<td>−2893</td>
</tr>
<tr>
<td></td>
<td>Perfect foresight</td>
<td>−5072</td>
<td>8151</td>
<td>−5072</td>
</tr>
</tbody>
</table>

Comparing the objective function values of the stochastic and deterministic models, we can see that the stochastic solutions have higher expected costs; this is to be expected, since the stochastic model needs to hedge against uncertainty. Table 2.2 also displays the actual (real) cost of implementing the solution value of the first-stage decision variables. Here, the difference is reversed: the deterministic solutions cannot cope with the uncertainty and end up being more expensive than the stochastic ones on average. We can thus conclude that for the investigated problem, it is beneficial (on average) to use a stochastic model instead of a deterministic model.

Looking at the running times of the model (also displayed in Table 2.2), we see that for most of the hours, the model is quite fast: between 68.8% and 99.4% of the individual hours in each case are solved within a minute. However, a few of the instances take considerably more time: between 0% and 9.6% of the instances take more than ten minutes (and often even much more) to solve. For the numbers displayed in Table 2.2 the optimisation of an individual hour was stopped after 1000 sec., which happened in 1.8% of the cases. For those hours not solved to optimality, the gap was 0.6% at most.

Analysing the difficult instances, we note that these represent hours with large changes in the supply of power over a very short time, or hours, where the overall production level is
close to the overall minimum or maximum capacity level. Looking at the overall running time for each of the different cases for the four weeks, we see that the stochastic model is solved faster than the deterministic model which is not what we expected to see. We also see that there are more of the individual hours that require a long running time in order to find the optimal solution in the deterministic model compared to the stochastic model. In order to explain this, we will first point out that the second-stages of the stochastic model are solved very fast. Second, in Table 2.3 on the following page, we see that the stochastic solution on average activates more additional power and deactivate less than the deterministic solution. In all four cases, the gap between the aggregated planned production level and the overall minimum capacity level is smaller than the gap between the maximum production level and the aggregated planned production level. This means that on the short-term horizon, we can
activate more power than we can deactivate. Since cases where the production level is close to the capacity limits are difficult to solve, an explanation for the longer running time of the deterministic cases could be that the production level comes close to the minimum capacity level.

Table 2.4 on the next page shows the aggregated total cost for the whole week given by the stochastic solution, the deterministic solution, and in case of perfect foresight for each of the four test instances. The results in Table 2.4 support the results in Table 2.2 and show that in all cases, the total cost given by the objective function of the stochastic solution is higher than the total cost given by the objective function of the deterministic solution. However, implementing the deterministic solution compared to implementing the stochastic solution entails a higher total actual cost in all of the four weeks. Looking at Table 2.5 on the facing page, we see that the actual cost of the stochastic solution is actually smaller than the actual cost given by the deterministic solution in 62% - 77% of the individual hours in the four weeks.

As pointed out earlier, in the stochastic solution we manually activate more additional power and deactivate less than in the deterministic solution. It can be seen in Table 2.4 that the stochastic solution also deactivates more automatic reserve power than in the deterministic case, and it activates less automatic reserve power. Since it is cheaper to activate manual reserve power than automatic reserve power, the stochastic solution builds in a buffer by assuring a higher level of produced power before the hour of operation. If there is no need for the additional power, it will be deactivated by the automatic reserves.

Now, one could think that incorporating this buffer could be rather expensive and that it could be beneficial not to be proactive and just leave all the imbalances to the automatic reserves. However, when looking at Table 2.6 on page 64, we see that the actual cost of such an
Table 2.4: Weekly costs of operating the system. All numbers listed are in €.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td>Stochastic</td>
<td>1 232 805</td>
<td>−1 372 805</td>
<td>709 147</td>
<td>−109 734</td>
<td>459 413</td>
</tr>
<tr>
<td></td>
<td>Deterministic</td>
<td>1 222 213</td>
<td>−1 460 396</td>
<td>518 206</td>
<td>−35 329</td>
<td>244 694</td>
</tr>
<tr>
<td></td>
<td>Perfect foresight</td>
<td>1 269 633</td>
<td>−1 392 438</td>
<td>510 245</td>
<td>−32 868</td>
<td>354 572</td>
</tr>
<tr>
<td>April</td>
<td>Stochastic</td>
<td>2 431 796</td>
<td>−2 180 726</td>
<td>1 013 361</td>
<td>−121 524</td>
<td>1 142 907</td>
</tr>
<tr>
<td></td>
<td>Deterministic</td>
<td>2 425 102</td>
<td>−2 236 951</td>
<td>796 819</td>
<td>−50 079</td>
<td>934 891</td>
</tr>
<tr>
<td></td>
<td>Perfect foresight</td>
<td>2 397 138</td>
<td>−2 228 160</td>
<td>823 970</td>
<td>−50 538</td>
<td>942 410</td>
</tr>
<tr>
<td>July</td>
<td>Stochastic</td>
<td>627 105</td>
<td>−671 029</td>
<td>199 857</td>
<td>−169 679</td>
<td>−13 746</td>
</tr>
<tr>
<td></td>
<td>Deterministic</td>
<td>590 006</td>
<td>−716 442</td>
<td>432 8</td>
<td>−110 351</td>
<td>−232 459</td>
</tr>
<tr>
<td></td>
<td>Perfect foresight</td>
<td>576 937</td>
<td>−707 088</td>
<td>404 581</td>
<td>−156 676</td>
<td>121 469</td>
</tr>
<tr>
<td>October</td>
<td>Stochastic</td>
<td>758 591</td>
<td>−1 419 053</td>
<td>226 467</td>
<td>−140 882</td>
<td>−574 877</td>
</tr>
<tr>
<td></td>
<td>Deterministic</td>
<td>693 241</td>
<td>−1 504 818</td>
<td>23 525</td>
<td>−27 475</td>
<td>−815 527</td>
</tr>
<tr>
<td></td>
<td>Perfect foresight</td>
<td>669 776</td>
<td>−1 509 467</td>
<td>21 028</td>
<td>−28 356</td>
<td>−847 019</td>
</tr>
</tbody>
</table>

Table 2.5: Comparison of the solutions given by the stochastic model and the deterministic model. For each of the two models, it is shown in how many of the individual hours the model found a solution with lowest actual cost.

<table>
<thead>
<tr>
<th>Case</th>
<th>Deterministic best</th>
<th>Stochastic best</th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td>23%</td>
<td>77%</td>
</tr>
<tr>
<td>April</td>
<td>37%</td>
<td>63%</td>
</tr>
<tr>
<td>July</td>
<td>34%</td>
<td>66%</td>
</tr>
<tr>
<td>October</td>
<td>38%</td>
<td>62%</td>
</tr>
</tbody>
</table>

The approach is very expensive. Hence, it is better to be proactive even though it can occasionally result in activation of manual reserves which is deactivated by automatic reserves when the uncertainty is revealed.

If there had been no uncertainty in the wind power production, we would have had the cost given by the solution of perfect foresight. If we could predict the wind power production precisely, it would result in huge savings. For July, it is more than four times the cost of the stochastic solution, and else it is between 23% and 55%.
Table 2.6: Weekly actual costs of operating the system with only automatic reserves. All numbers listed are in €.

<table>
<thead>
<tr>
<th>Case</th>
<th>Total cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td>2 368 144</td>
</tr>
<tr>
<td>April</td>
<td>3 838 005</td>
</tr>
<tr>
<td>July</td>
<td>809 076</td>
</tr>
<tr>
<td>October</td>
<td>462 443</td>
</tr>
</tbody>
</table>

2.5 Conclusion

In this paper, we have presented a stochastic model for the short-term balancing problem between demand and consumption of electricity. We have made scenarios for the wind power prediction error by a copula-based heuristic that captures the dependency between all the wind variables but also the dependency of the errors of the individual wind variables through time.

The results show that the stochastic model is superior to the deterministic model when looking at the actual cost of the solutions. The stochastic model builds in a buffer of additional electricity by activating more manual reserve power than the deterministic model. It does so since it is cheaper to activate the manual reserves than the automatic reserves, and if there is no need for the additional power, it is deactivated by the automatic reserves.

The results also show that it is a good idea to be proactive and activate manual reserves before the actual imbalances occur. It will be much more expensive to only let the automatic reserves handle the imbalances.

Acknowledgments

The authors gratefully appreciate all the support they have received from Energinet.dk in getting data and lending the models Sivael and Simba throughout this project. Furthermore, the authors appreciate the work Stephan Wöllner at Energinet.dk has made to make Simba work together with our model OPTIBA. Jeanne Andersen acknowledges support through the CFEM project.
Appendix

2.A Graphs for historical wind data

Supplementary Fig. 2.A.1: The average of the normalised wind power prediction errors in Denmark recorded in 2012 for each of the four months we investigate. The thin line represents values for January, the thick line represents values for April, the scattered line represents values for July, and the dotted line represents values for October.
**Supplementary Fig. 2.A.2:** Standard deviation of the normalised wind power prediction errors in Denmark recorded in 2012 for each of the four months we investigate. The thin line represents values for January, the thick line represents values for April, the scattered line represents values for July, and the dotted line represents values for October.
Chapter 3

Supply chain network and route design for second generation bioethanol

Jeanne Andersen · Claus Aage Grøn Sørensen
Andreas Klose · Kim Allan Andersen

J. Andersen and K. A. Andersen: CORAL, Department of Economics and Business, Aarhus University, Fuglesangs Alle 4, 8210 Aarhus V, Denmark

C. A. G. Sørensen: Department of Engineering - Operations Management, Aarhus University, Inge Lehmanns Gade 10, 8000 Aarhus C, Denmark

A. Klose: Department of Mathematics, Aarhus University, Ny Munkegade 118, 8000 Aarhus C, Denmark

Preliminary version
Supply chain network and route design for second generation bioethanol

Abstract
We present a mixed integer linear model, which can be used to design the supply chain network of biomass for second generation of bioethanol. In our case the particular biomass is straw that may be converted into compressed briquettes. In addition to other models for biomass supply chain design, we optimise the network taking routes for trucks into account. We propose a Lagrangian relaxation by variable splitting as the solution method. On small test instances, this method provides reasonable gaps between upper and lower bounds on the objective function value.

Keywords: OR in energy, Supply chain network design, Routing

3.1 Introduction
Due to the environmental drawbacks and diminishing availability of fossil fuels, alternative and renewable energy sources are needed to achieve sustaining and secure energy supply. However, a problem is that some renewable energy sources are non-controllable and intermittent, e.g. energy from wind and solar. As intermittent energy production is growing, a need for more controllable energy sources emerge. Biomass is such a controllable energy source as it can be stored and converted to energy on demand, and studies indicate that in the future, biomass can be a significant resource for production of electricity, heat, and transport fuels (Meyer et al. (2014)), e.g. see Intergovernmental Panel on Climate Change (2011); Energinet.dk (2010); IEA (2012).

However, the future role of biomass depends on the biomass-to-energy supply chain to overcome barriers that may hinder this development, see Bravo et al. (2012). When designing supply chains for biomass, an important key driver is to minimise the total costs (cost-efficiency). At present, the production of bioenergy is far more expensive than exploiting traditional sources of energy (Haas et al. (2006); IEA (2012)). One major reason is the high logistics
costs arising in bioenergy provision as energy crops are grown over large and isolated geographical areas, see Allen et al. (1998) and Caputo et al. (2005). Improving the logistics of the biomass-to-energy supply chain is therefore a requirement for enabling bioenergy supply to become an economical and environmentally sound additional long-term source of energy. For achieving efficiency in logistics operations, the design of the supply chain network is of utmost importance. It involves long-term decisions regarding transport flows, capacities as well as numbers and types of facilities used in the network for providing biomass and transforming it into energy.

In this paper, we will look at how to optimise the design of the supply chain network of biomass. We are inspired by a real problem faced by Maabjerg BioEnergy. The firm is located in Holstebro, Denmark, and is in the process of building a large second generation bioethanol plant that will be able to produce bioethanol from straw. The project, which in total will cost approximately €290 million, is supported through the EU NER 300 program by €39.3 million, and it is expected, that the plant should be operational in year 2017. This new plant will interact with the combined heat and power plant, as well as the biogas plant already installed in the area. The by-products, molasses and lignin, from the production of bioethanol will be used as fuel in the biogas plant and the combined heat and power plant, respectively. Furthermore, the by-product lignin is expected to substitute the considerable amount of waste imported from London, that today constitute the fuel at the combined heat and power plant. Once the new plant is operational it will be able to produce approximately 80 million litres of bioethanol. In order to produce that amount of bioethanol, it is estimated, that about 300,000 tons straw evenly distributed over a year are needed to run the plant efficiently. Therefore, Maabjerg BioEnergy faces a huge challenge at the moment in setting up their supply chain of locally supplied straw. They estimate that there is enough supply at local farmers to cover their demand within a range of 80 km from the plant, but in case of increased competition in the demand market and/or varying seasonal production outputs, it may be necessary to collect straw from further away (Kjærgaard et al. (2014)). Using Maabjerg BioEnergy as a case study, we investigate the supply chain network design for straw that may be converted into compressed briquettes. The decisions will include where to locate storage and converting fa-
ilities, the capacity of these facilities, and the transportation routes of straw from their origins (farms or fields) to the destinations (the bioethanol plant). It will also be decided how many of a heterogeneous fleet of trucks that is needed to carry out the transportation of straw. See Fig. 3.1 for an overview of the underlying network structure of our problem.

![Network structure of the supply chain.](image)

Clearly, our problem relates to the set of location-routing problems, as well as to the design of supply chains. Since a location-routing problem is a combination of both the classical location and the classical vehicle routing problem, which both are NP-hard, location-routing problems are NP-hard as well. One possible way to approach the problem may be to separate the location decisions from the routing decisions. However, it was rather early discovered by Maranzana (1964) and Webb (1968) that this approach may lead to suboptimal decisions. This insight has later led to an increased attention about the combined problem, and in the 1980s, a more formal classification of location-routing problems were given. One of the first papers on location-routing problems was by Perl and Daskin (1985). They described the warehouse location-routing problem and solved it using a heuristic. Since that classical paper appeared, the location-routing problem has been an active research area and attracted a lot of attention. Due to the hardness of the problem, most of the proposed solution methods involves using heuristics.

However, we would like to stress that there are a few papers on location-routing problems which describe exact solution methods, among them Belenguer et al. (2011). In their paper, they propose a branch-and-cut method for the capacitated location-routing problem and solve
instances with up to 50 customers and 5-10 possible locations. But not all the instances were solved to optimality. So, in general, real-life instances, which tend to be much larger, will be very hard to solve to optimality. To include a complete survey on location-routing models in this paper, is out of the scope. Instead, we refer to the survey papers by Nagy and Salhi (2007) and by Prodhon and Prins (2014).

Regarding the design of supply chains for biomass, there exists a variety of models depending on which type of biomass the supply chain needs to cover. A few examples of biomass, for which models have been developed, are wood (Gunnarsson et al. (2004), Gronalt and Rauch (2007), Huang et al. (2014)), corn (Čuček et al. (2010), Huang et al. (2014)), and straw (van Dyken et al. (2010)). In the past few years, this area has generated substantial interest. Some recent survey papers on biomass supply chain design are Iakovou et al. (2010) and Sharma et al. (2013). Below, we mention a few recent works on the design of supply chains for biomass. All developed models in these works have a size that could be solved using GAMS and CPLEX.

Gunnarsson et al. (2004) published some work on biomass supply chain planning. Both papers consider specialised supply chain models of forest fuel. The first paper studies the problem of deciding when and where forest residues should be converted into forest fuel, and how the residues should be transported and stored in order to satisfy demand at heating plants. The problem is formulated as a large mixed integer linear program, and a heuristic solution approach for minimising the total costs is proposed. The model also involves locational decisions regarding the locations of terminals at which the forest residues are chipped. Also Gronalt and Rauch (2007) consider the problem of locating chipping terminals in a forest fuel supply network. For given demands and locations of heating and energy plants, different configurations of chipping terminals are compared regarding the estimated resulting system cost.

Čuček et al. (2010) present a mixed integer linear programming model for regional biomass supply chain with profit maximisation as the optimisation criterion, where the environmental impact is evaluated by the carbon footprint as a primary performance indicator. The types of biomass considered are corn stover and wood. The problem determines the optimal locations of collection points and of processing plants. Transportation costs between the facilities are also taken into account. Several transportation modes may be used, but the distinct transportation
routes used are not determined in the model. The mathematical model is demonstrated using a case study. In Čuček et al. (2012), the work is extended to also look at how to obtain sustainable solutions where both economic, environmental, and social aspects are present, using multi-objective optimisation. Furthermore, they look at both the direct and indirect effects from a life-cycle perspective.

In van Dyken et al. (2010), a mixed integer programming model for biomass supply chains is presented. It includes transport as well as storage and processing. The types of biomass considered are spruce, chips, and pellets. The main objective of the paper is to present a modelling framework that can be applied to a variety of biomass supply chains. Therefore, not much effort was spent on obtaining realistic data. The most important part has been to represent the relationship between moisture and energy content of different kinds of biomass and to handle long-term processes in the optimisation.

Huang et al. (2014) develop a two-stage mixed integer stochastic model for designing a biofuel supply chain network at minimum total cost. The available feedstock is uncertain and represented by a probability function. The type of biomass considered is biowaste resources such as corn stover and forest residues. The decisions include the locations and sizes of refineries and storage facilities. Furthermore, the operational decisions include procurement and storage of feedstock, as well as of seasonal ethanol productions, storage, and distributions. A progressive hedging method is developed to solve the resulting model. While the model can be solved using CPLEX, the developed method works much faster.

Kim et al. (2011) develop a multi-period mixed integer stochastic programming model for designing a biomass supply chain network for biofuels under uncertainty, with the objective of maximising total profit. The types of biomass considered are logging residuals, thinnings, prunings, grasses, and chips/shavings. There are several uncertain parameters in the model, some of them being the transportation costs and the sales prices for the products. The model determines the sizes and the locations of different types of plant facilities.

Another important key driver for promoting biomass as replacement for fossil fuel is the reduction of Green House Gas (GHG) emissions throughout its life cycle (IEA (2012)). Therefore, GHG emission reduction potentials have to be considered when designing the supply chain.
network of biomass-to-energy. A few papers have considered both the minimisation of cost (or maximising profit) and minimisation of GHG emissions. This is accomplished by using multi-objective optimisation.

Mele et al. (2011) present a detailed bi-objective (net present value vs. GHG emissions) model for the optimal design of supply chains for the combined production of sugar and bioethanol. They illustrate their model through a case study based on the Argentinean sugar cane industry.

In Giarola et al. (2011) and Zamboni et al. (2009), they present a detailed bi-objective mixed integer linear model for the supply chain of bioethanol, which they apply on a case study for the future Northern Italy biomass-based ethanol production. The biomass considered in both cases is corn grain. The two objectives are to maximise net present value and to minimise GHG emission. The model has been further developed in Bernardi et al. (2013) to take water footprints into account.

In Pérez-Fortes et al. (2012), they develop a multi-criteria mixed integer linear programming model for the design of biomass supply chain for electricity generation. The biomass considered is cassava waste. Sustainability is in focus and three objectives are taken into account, namely economic, environmental, and social criteria. They account for transportation in their environmental performance, but it does not depend on the mode.

In You et al. (2012), a detailed multi-criteria mixed integer linear programming model for the design of cellulosic biofuel supply chain is presented. It takes economic, environmental, and social objectives into account. It also considers different modes of transportation, and their influence on the environmental performance.

While all the models in the papers presented above could be solved using GAMS and CPLEX, this is not the case for the problem considered in this paper. The main difference is that we consider both the design of the supply chain for biomass, as well as the construction of the optimal routes. To the best of our knowledge, this combined problem has not been considered before. Given the size of our problem, we need to design a tailor-made algorithm to solve it.

The rest of the paper is organised as follows: in Section 3.2, we introduce the mathematical
model, and in Section 3.3, we present a solution procedure for the problem. In Section 3.4, we present the data, and in Section 3.5, we solve a prototype example.

3.2 The mathematical formulation

The biomass supply chain will be represented by a network. In the network, we have a set of possible locations where storage and converting facilities may be opened. Also, a number of different technologies are available for being installed at a facility. One technology may be to convert straw into compressed briquettes, and another technology may be to keep the straw in its baled form. If it is decided to open a facility at a particular location, it is also necessary to decide which technology to install.

Whenever a new facility or a plant is opened, it is decided which technology it should be equipped with along with its capacity. For the facilities, the capacity level must not exceed a given maximum. Opening facilities and managing storage imposes a fixed cost that depends on the capacity level and the technology installed. It is assumed that the relationship between the capacity level and the fixed cost is linear.

Each arc in the network is associated with a transportation time and a transportation cost. The transportation cost depends on the weight of each load. Therefore, the transportation cost is divided into two parts. The first part is the transportation cost for a truck with no load, and the second part is the added transportation cost for a truckload of straw. The straw is in a form determined by the in-field treatment, e.g. bales, or in a form determined by some conversion technology, e.g. briquettes. This means that if a truck carries a full load of straw treated with a particular technology, then the total transportation cost is equal to the sum of these two parts.

In each time period, the model has to decide how much straw that has to be transported between the nodes in the network. For the transport a heterogeneous fleet of vehicles, located at the same depot, is available. Straw treated with a particular technology can be carried only on trucks designed to carry that kind of straw. When designing the routes, it is required that the total duration of a route must not exceed a certain threshold value.

The network can be represented by a graph $G = (\mathcal{N}, \mathcal{A})$, where $\mathcal{N} = \{0, 1, \ldots, N\}$ is a set of nodes and $\mathcal{A} = \{(i,j) : i \neq j, i, j \in \mathcal{N}\}$ is a set of arcs. The set of nodes is divided
Sets

\[ \mathcal{T} \quad \text{Set of time periods} = \{1, 2, \cdots, T\}. \]

\[ \mathcal{M} \quad \text{Set of different technologies which can be implemented in the storage and converting facilities} = \{1, 2, \cdots, M\}. \]

\[ \mathcal{M}_m \quad \text{Subset of } \mathcal{M} \text{ where the elements can be treated with technology } m \in \mathcal{M}. \]

\[ \mathcal{K}_m^t \quad \text{Set of trucks available to carry straw treated with technology } m \in \mathcal{M} \text{ in period } t \in \mathcal{T} = \{ M(t-1)K + (m-1)K + 1, \ldots, M(t-1)K + mK \}. \]

\[ K \quad \text{is an upper limit on the number of trucks per technology in any period.} \]

\[ \mathcal{K} \quad \text{Set of trucks, i.e. } \mathcal{K} = \bigcup_{m \in \mathcal{M}} \bigcup_{t \in \mathcal{T}} \mathcal{K}_m^t. \]

Parameters

\[ c^k_{ij} \quad \text{Transportation cost on arc } (i, j) \in \mathcal{A} \text{ for an empty truck } k \in \mathcal{K}. \]

\[ c^m_{ij} \quad \text{Added transportation cost on arc } (i, j) \in \mathcal{A} \text{ for a truckload of straw treated with technology } m \in \mathcal{M}. \]

\[ c_{it}^m \quad \text{Inventory carrying cost in storage and converting facility established at location } i \in \mathcal{N}_1 \cup \mathcal{N}_2 \text{ in time period } t \in \mathcal{T}. \]

\[ f_{Km}^t \quad \text{Fixed cost for truck } k \in \mathcal{K}_m^t. \]

\[ f_j^m \quad \text{Capacity cost per unit for opening facilities equipped with technology } m \in \mathcal{M} \text{ at node } j \in \mathcal{N}_2 \cup \mathcal{N}_3. \]

\[ \tau_{ij} \quad \text{Transportation time by traversing arc } (i, j) \in \mathcal{A} \text{ for an empty truck.} \]

\[ \tau_{\text{MAX}} \quad \text{The maximum duration time of a route for a truck.} \]

\[ \beta_{im} \quad \text{The weight of a truckload of straw treated with technology } m \in \mathcal{M} \text{ when transported from node } i \in \mathcal{N}_1 \cup \mathcal{N}_2. \]

\[ \pi^m \quad \text{Cost for transforming one unit of straw into one unit of converted straw using technology } m. \]

\[ \text{CAP}_{jmt} \quad \text{Maximum capacity level of a storage and converting facility established at location } j \in \mathcal{N}_1 \cup \mathcal{N}_2 \text{ and equipped with technology } m \in \mathcal{M}. \]

\[ \text{CAP}_{jmt} \quad \text{Amount of straw available in node } i \in \mathcal{N}_1, \text{ in time period } t \in \mathcal{T}, \text{ for technology } m \in \mathcal{M}. \]

\[ D_{jt} \quad \text{Demand at node } j \in \mathcal{N}_3 \text{ in time period } t \in \mathcal{T}. \]

Decision variables

\[ x_{ijt}^k \quad \text{Number of times truck } k \in \mathcal{K} \text{ uses the arc } (i, j) \in \mathcal{A} \text{ in time period } t \in \mathcal{T}. \]

\[ y_{jm} \quad \text{Binary variable that indicates if a facility equipped with technology } m \in \mathcal{M} \text{ is opened at node } j \in \mathcal{N}_2. \]

\[ I_{jt}^m \quad \text{Ultimo inventory level at time period } t \in \mathcal{T} \text{ of straw treated with technology } m \in \mathcal{M} \text{ at node } j \in \mathcal{N}_2 \cup \mathcal{N}_3. \]

\[ v_{ijt}^m \quad \text{Number of loads of straw treated with technology } m \in \mathcal{M} \text{ that have to be shipped between nodes } i \text{ and } j. \]

\[ L_{Km}^t \quad \text{The number of trucks needed to carry straw treated with technology } m \in \mathcal{M} \text{ in period } t \in \mathcal{T}. \]

\[ C^m_j \quad \text{Capacity level of a facility equipped with technology } m \in \mathcal{M} \text{ in node } j \in \mathcal{N}_2 \cup \mathcal{N}_3. \]

\[ \omega_{ijt}^k \quad \text{Amount of a fictitious flow on arc } (i, j) \in \mathcal{A} \text{ in time period } t \in \mathcal{T} \text{ for the truck } k \in \mathcal{K}. \]

\[ D_{jt}^m \quad \text{Nonnegative help variables, such that } \sum_{m \in \mathcal{M}} D_{jt}^m = D_{jt}, j \in \mathcal{N}_3, t \in \mathcal{T}. \]
into smaller disjunctive subsets $\mathcal{N} = \{0\} \cup \mathcal{N}_1 \cup \mathcal{N}_2 \cup \mathcal{N}_3$, where $\{0\}$ represents the depot for the trucks, $\mathcal{N}_1$ consists of all the nodes representing the location of the farmers, $\mathcal{N}_2$ consists of all the nodes representing the possible locations for the storage and converting facilities, and $\mathcal{N}_3$ consists of the nodes representing the location of bioethanol plants. In our case, we have $|\mathcal{N}_3| = 1$. If a possible location of a storage and converting facility is at a farmers place, then the farming place and the possible location of the facility are modelled as two separate nodes in the mathematical formulation.

### 3.2.1 Objective function

In the objective function, we include the yearly fixed cost for operating facilities, the inventory cost, and the cost for the trucks; both the fixed cost and the variable transportation cost.

$$
\min z = \sum_{t \in T} \sum_{m \in \mathcal{M}} f^m_{k, m, t} + \sum_{t \in T} \sum_{m \in \mathcal{M}} \sum_{(i, j) \in A} \left( \sum_{k \in \mathcal{K}_{m, t}} c^k_{i, j, k, m, t} + c^m_{i, j, m, t} \right) + \sum_{m \in \mathcal{M}} \sum_{j \in \mathcal{N}_2 \cup \mathcal{N}_3} \left( f^m_{j, m} \mathcal{C}^m_{j, m} + \sum_{t \in T} c^m_{j, m} \mathcal{I}^m_{j, m, t} \right) + \sum_{m \in \mathcal{M}} \sum_{t \in T} \sum_{m \in \mathcal{M}} \sum_{i \in \mathcal{N}_1 \cup \mathcal{N}_2} \sum_{j \in \mathcal{N}_3} \pi^m_{i, j, m, t}.
$$

The objective function is a sum of six costs. The first one is the fixed cost associated with the trucks. The second part is the transportation cost associated with transportation with trucks that are empty, and the third part is the extra cost imposed when the trucks carry straw that has been treated with some technology. The fourth part is the cost associated with the capacity of the new facilities, and the fifth part is the inventory cost. The last part is the total cost for transforming straw into other physical forms.

### 3.2.2 Constraints

**Capacities and location of facilities**

The level of capacity for storage at the plant and at the converting and storage facilities has to be decided upon:

$$
I^m_{j, m} \leq \mathcal{C}^m_{j, m}, \quad j \in \mathcal{N}_2 \cup \mathcal{N}_3, m \in \mathcal{M}, t \in T, \quad (3.1)
$$

$$
\mathcal{C}^m_{j, m} \leq \mathcal{C}\mathcal{A}\mathcal{P}_{\text{jan}} \mathcal{Y}_{j, m}, \quad j \in \mathcal{N}_2, m \in \mathcal{M}. \quad (3.2)
$$
In equation (3.2), we also make sure to open a facility that is used. Storage at the bioethanol plant has to be implemented no matter what, which is why equation (3.2) does not apply to $N_3$. If an upper bound on the capacity for the plant is needed, it can be implemented by including the upper bound in equation (3.1). Only one facility can be open at one location with one kind of technology. Hence,

$$\sum_{m \in M} y_{jm} \leq 1, \ j \in N_2. \quad (3.3)$$

Furthermore, the capacity of the farming sites has to be obeyed.

$$\sum_{t' \leq t} \sum_{j \in N_2 \cup N_3} \beta_{im} v_{ijt}^m \leq \sum_{t' \leq t} \text{CAP}_{im'}, \ i \in N_1, m \in M, t \in T. \quad (3.4)$$

The amount of straw leaving facility $i$ in time period $t$ has to be less or equal to the amount available at the facility. Here, we assume one period for handling the straw transported to the facility, and therefore, the available amount of straw is the inventory from last period.

$$\sum_{j \in N_2 \cup N_3} \beta_{im} v_{ijt}^m \leq I_{mj}(t-1), \ i \in N_2, m \in M, t \in T. \quad (3.5)$$

**Inventory**

We have to keep track of the inventory at the plant in each period.

$$\sum_{i \in N_1 \cup N_2} \beta_{im} v_{ijt}^m + I_{jt}^m - D_{jt}^m = I_{jt}^m, \ j \in N_3, m \in M, t \in T, \quad (3.6)$$

where,

$$\sum_{m \in M} D_{jt}^m = D_{jt}, \ j \in N_3, t \in T. \quad (3.7)$$

Furthermore, we also have to keep track of the inventory at the facilities in each period.

$$\sum_{m' \in M^m \ i \in N_1 \cup N_2} \beta_{im'} v_{ijt}^{m'} - \sum_{i \in N_2 \cup N_3} \beta_{jm} v_{jit}^m + I_{jt}^m = I_{jt}^m, \ j \in N_2, m \in M, t \in T. \quad (3.8)$$

Here, we have subsets $M^m$ of the original set $M$ since bales of straw can be converted into briquettes but briquettes cannot be converted into bales of straw.
Truck routes

We need to transport all the required amount of truckloads between node $i$ and $j$.

$$\sum_{k \in K^{m,t}} x_{ijt}^k \geq v_{ijt}^m, \quad (i, j) \in A, m \in M, t \in T. \quad (3.9)$$

If a truck is visiting a node $j$, it also has to leave it again.

$$\sum_{i \in N} x_{ijt}^k - \sum_{i \in N} x_{jit}^k = 0, \quad j \in N, k \in K^{m,t}, m \in M, t \in T. \quad (3.10)$$

The number of trucks leaving the depot in each time period is recorded, such that we know exactly how many trucks will be needed to perform the transportation of straw.

$$\sum_{j \in N \setminus \{0\}} \sum_{k \in K^{m,t}} x_{0jt}^k \leq L_{K^{m,t}}, \quad m \in M, t \in T. \quad (3.11)$$

Furthermore, we assume that the number of trucks available to carry straw treated with technology $m \in M$ in time period $t \in T$ is limited by a constant $K$.

$$L_{K^{m,t}} \leq K, \quad m \in M, t \in T. \quad (3.12)$$

We make sure that the time duration of a tour does not exceed the limit, by

$$\sum_{(i,j) \in A} \tau_{ij} x_{ijt}^k \leq \tau_{\text{MAX}}, \quad k \in K^{m,t}, m \in M, t \in T. \quad (3.13)$$

Subtour elimination constraints

To ensure that the routes of the trucks do not have any subtours we have the following multi-commodity network flow constraints.

First, we make sure that the flow on an arc is at least as large as the number of times the arc is used by the truck.

$$x_{ijt}^k \leq \omega_{ijt}^k, \quad (i, j) \in A, k \in K^{m,t}, m \in M, t \in T, \quad (3.14)$$

$$\omega_{ijt}^k \leq |A| x_{ijt}^k, \quad i \in N, j \in N \setminus \{0\}, k \in K^{m,t}, m \in M, t \in T. \quad (3.15)$$

Then, for each node the truck visits, we send flow back to the depot.

$$x_{ijt}^k \leq \omega_{0jt}^k, \quad i \in N \setminus \{0\}, j \in N, k \in K^{m,t}, m \in M, t \in T. \quad (3.16)$$
We ensure that no node accumulates flow in the network,
\[ \sum_{i \in N} \omega_{ijt}^k = \sum_{j \in N} \omega_{jlt}^k, \quad j \in N, k \in K^{m,l}, m \in M, t \in T. \] (3.17)
Finally no flow is allowed to flow from the depot to itself.
\[ \omega_{00t}^k = 0, \quad k \in K^{m,l}, m \in M, t \in T. \] (3.18)

### 3.3 Lagrangian relaxation by variable splitting

Inspired by Fisher et al. (1997), we propose to use Lagrangian relaxation by variable splitting to solve the model presented in Section 3.2. Variable splitting or Lagrangian decomposition was originally introduced by Guignard and Kim (1987) and Glover and Klingman (1988). By looking at the problem in Section 3.2, it can be seen, that the first part of the model consists of the design of the supply chain network, and the second part of the model consists of transporting the straw between facilities. There is only one constraint binding the two parts together, namely constraint (3.9). We add the following constraints to our problem:
\[ v_{ijt}^m = \tilde{v}_{ijt}^m, \quad (i, j) \in A, m \in M, t \in T. \] (3.19)
We also change constraints (3.9) into:
\[ \sum_{k \in K^{m,l}} x_{ijt}^k \geq \tilde{v}_{ijt}^m, \quad (i, j) \in A, m \in M, t \in T. \] (3.20)

By relaxing constraint (3.19) with multipliers \( \lambda = (\lambda_{ijt}^m) \), we obtain the Lagrangian relaxation
\[
z(\lambda) = \min \left\{ \sum_{t \in T} \sum_{m \in M} f_{K^{m,l}} L_{K^{m,l}} + \sum_{t \in T} \sum_{m \in M} \sum_{(i, j) \in A} \left( \sum_{k \in K^{m,l}} c_{ijt}^k x_{ijt}^k + v_{ijt}^m \right) \right\}
+ \sum_{m \in M} \sum_{j \in N_2 \cup N_3} \left( f_j^m C_j^m + \sum_{t \in T} c_{jt}^m I_{jt} \right) + \sum_{t \in T} \sum_{m \in M} \sum_{i \in N_1 \cup N_2} \sum_{j \in N_3} \pi_i^m v_{ijt}^m
+ \sum_{t \in T} \sum_{m \in M} \sum_{(i, j) \in A} \lambda_{ijt}^m (v_{ijt}^m - \tilde{v}_{ijt}^m) \right\}
\] (3.21)

s.t.
Constraints (3.1) – (3.8), (3.10) – (3.18), (3.20),
and domain constraints on variables.
Solving problem (3.21) gives a lower bound on the optimal value of the model presented in Section 3.2. The best lower bound is found by solving the Lagrangian dual

\[ z_{LD} = \max_\lambda z(\lambda). \]  

(3.22)

Problem (3.21) decomposes into

\begin{align*}
  z^1(\lambda) = \min & \sum_{t \in T} \sum_{m \in M} \sum_{(i,j) \in A} (c_{ij}^m + \lambda_{ij}^m) \nu_{ijt}^m + \sum_{m \in M} \sum_{j \in N_3 \cup N_5} \left( f_j^m C_j^m + \sum_{t \in T} c_{jt}^m I_{jt}^m \right) \\
  \text{s.t.} & \sum_{t \in T} \sum_{m \in M} \sum_{i \in N_1 \cup N_2} \sum_{j \in N_5} \pi_{ijt}^m \nu_{ijt}^m \\
\end{align*}

(3.23)

and

\begin{align*}
  \min & \sum_{t \in T} \sum_{m \in M} f_{K^m,t}^L K_{K^m,t}^L + \sum_{t \in T} \sum_{m \in M} \sum_{(i,j) \in A} \sum_{k \in K_{K^m,t}} c_{ij}^m x_{ijt}^k - \sum_{t \in T} \sum_{m \in M} \sum_{(i,j) \in A} \lambda_{ijt}^m \bar{v}_{ijt}^m \\
  \text{s.t.} & \sum_{k \in K_{K^m,t}} x_{ijt}^k \geq \bar{\nu}_{ijt}^m, (i,j) \in A \\
  & \sum_{i \in N} x_{ijt}^k - \sum_{i \in N} x_{ijt}^k = 0, j \in N, k \in K_{K^m,t} \\
  & \sum_{j \in N \setminus \{0\}} \sum_{k \in K_{K^m,t}} x_{ijt}^k \leq L_{K^m,t} \\
  & L_{K^m,t} \leq K \\
  & \sum_{(i,j) \in A} \tau_{ij} x_{ijt}^k \leq \tau_{MAX}, k \in K_{K^m,t} \\
  & x_{ijt}^k \leq \omega_{ijt}^k \leq |A|x_{ijt}^k, (i,j) \in A, k \in K_{K^m,t} \\
  & x_{ijt}^k \leq \omega_{ijt}^k, i,j \in N \setminus \{0\}, k \in K_{K^m,t} \\
  & \sum_{i \in N} \omega_{ijt}^k = \sum_{i \in N} \omega_{ijt}^k, j \in N, k \in K_{K^m,t} \\
\end{align*}

(3.24)

Constraints (3.1) – (3.8), and domain constraints on variables.

Problem (3.24) can in turn be partitioned into \(|T| \times |M|\) independent problems:

\begin{align*}
  \min z^2_{lm}(\lambda) = f_{K_{K^m,t}} L_{K_{K^m,t}} + \sum_{(i,j) \in A, k \in K_{K^m,t}} c_{ij}^m x_{ijt}^k - \sum_{(i,j) \in A} \lambda_{ijt}^m \bar{v}_{ijt}^m \\
  \text{s.t.} & \sum_{k \in K_{K^m,t}} x_{ijt}^k \geq \bar{\nu}_{ijt}^m, (i,j) \in A \\
  & \sum_{i \in N} x_{ijt}^k - \sum_{i \in N} x_{ijt}^k = 0, j \in N, k \in K_{K^m,t} \\
  & \sum_{j \in N \setminus \{0\}} \sum_{k \in K_{K^m,t}} x_{ijt}^k \leq L_{K^m,t} \\
  & L_{K^m,t} \leq K \\
  & \sum_{(i,j) \in A} \tau_{ij} x_{ijt}^k \leq \tau_{MAX}, k \in K_{K^m,t} \\
  & x_{ijt}^k \leq \omega_{ijt}^k \leq |A|x_{ijt}^k, (i,j) \in A, k \in K_{K^m,t} \\
  & x_{ijt}^k \leq \omega_{ijt}^k, i,j \in N \setminus \{0\}, k \in K_{K^m,t} \\
  & \sum_{i \in N} \omega_{ijt}^k = \sum_{i \in N} \omega_{ijt}^k, j \in N, k \in K_{K^m,t} \\
\end{align*}

(3.25)

Domain constraints on variables.
We see that the value of the Lagrangian dual is: 
\[ z_{LD} = \max_\lambda \left( z^1(\lambda) + \sum_{t \in T} \sum_{m \in M} z^2_{im}(\lambda) \right). \]

**Lemma 1** If \( c_{ij} \) obey the triangle inequality, then there is an optimal solution to problem (3.25) that can be found by only including node \( i \in N_1 \cup N_2 \) in the problem if either \( \lambda^t_{ijm} > 0 \) or \( \lambda^t_{jim} > 0 \) for some \( j \in N_1 \cup N_2 \).

**Proof** For a fixed \( t \) and \( m \) let \( z^2_{im}(\lambda) \) be the optimal solution value to problem (3.25). Assume that there is a node \( i \in N_1 \cup N_2 \) for which \( \bar{v}^m_{ijt} > 0 \) and where both \( \lambda^m_{ijt} \leq 0 \) and \( \lambda^m_{jim} \leq 0 \), \( \forall j \in N_1 \cup N_2 \). We fix the value of \( \bar{v}^m_{ijt} \) to 0. We hereby obtain a new feasible solution with objective value \( \bar{z}^2_{im}(\lambda) \), where \( \bar{z}^2_{im}(\lambda) \geq z^2_{im}(\lambda) \) due to the negative sign of the last term in the objective function. Due to the triangle inequality, the cost of traveling cannot increase if \( \sum_{k \in K^m, t} x^k_{ijt} = 0 \).

Hence, if a node \( i \) only has \( \lambda^m_{ijt} \leq 0 \) and \( \lambda^m_{jim} \leq 0 \), \( \forall j \in N_1 \cup N_2 \), all \( \bar{v}^m_{ijt} = 0 \) and \( x^k_{ijt} = 0 \), \( \forall j \in N_1 \cup N_2 \), \( k \in K^m \), we can exclude the node from the problem. \( \square \)

**Lemma 2** A lower bound on the optimal solution value \( z^* \) is given by
\[
z^* \geq \min \sum_{t \in T} \sum_{m \in M} \sum_{(i,j) \in A} \frac{\tau_{ij} v^m_{ijt}}{\tau_{MAX}} + \sum_{t \in T} \sum_{m \in M} \sum_{(i,j) \in A} \left( \bar{c}_{ij} + c^m_{ij} \right) v^m_{ijt} \\
+ \sum_{j \in N_2 \cup N_3} \sum_{i \in N_1} \left( f^m_j \sigma^m_j + \sum_{t \in T} c^m_{jt} l^m_{jt} \right) \]
\[ \text{s.t. Constraints (3.1) – (3.18), and domain constraints on variables}, \]
where
\[ \bar{c}_{ij} = \min_{k \in K^m} \{ c^k_{ij} \}. \]

**Proof** The total transportation time used is: \( \sum_{(i,j) \in A} \sum_{m \in M} \tau_{ij} v^m_{ijt} \). Because the maximum duration of a route for a truck is \( \tau_{MAX} \), a lower bound on the number of trucks needed is given by:
\[
L_{K^m} \geq \frac{\sum_{(i,j) \in A} \sum_{m \in M} \tau_{ij} v^m_{ijt}}{\tau_{MAX}}. \]
This implies that
\[
\sum_{t \in T} \sum_{m \in M} f^m_{K^m, t} L_{K^m} \geq \sum_{t \in T} \sum_{m \in M} f^m_{K^m, t} \frac{\sum_{(i,j) \in A} \sum_{m \in M} \tau_{ij} v^m_{ijt}}{\tau_{MAX}}. \]
Constraint (3.9) implies that
\[
\sum_{t \in T} \sum_{m \in M} \sum_{(i,j) \in A} \sum_{k \in K} \chi^k_{ijt} x^k_{ijt} \geq \sum_{t \in T} \sum_{m \in M} \sum_{(i,j) \in A} \bar{c}_{ijt} v^m_{ijt}.
\]

Using these two observations, the result follows immediately. \(\square\)

### 3.3.1 Solution procedure

The Lagrangian dual is solved using a subgradient procedure. A pseudocode is presented in Algorithm 1 on the facing page. In the first three lines, the algorithm is initialised. The main loop runs from line 4 to line 16. During each iteration, the problems (3.23) and (3.25) are solved in lines 5-7, each time using an updated set of Lagrangian multipliers. In line 8, it is checked if the solution is optimal, in which case the algorithm stops. During the first iteration, we calculate an upper bound in lines 11-12. If an optimal solution is not found, the Lagrangian multipliers are updated in lines 13-15. If the stop criterion is not satisfied, the algorithm proceeds with the next iteration. Otherwise, it stops and returns a lower bound and the upper bound to problem (3.22) in line 17. The stop criterion used is a limit on the number of iterations in the loop between lines 4-16.

The final result of the procedure will provide us with a lower and upper bound on the optimal value of problem (3.21). Additional upper bounds on \(z^*\) can be found as follows. Choose any of the solutions determined by solving problem (3.23) in line 5. Then fix the values of \(\bar{v}^m_{ijt}, (i,j) \in A, m \in M, t \in T\) in problem (3.25) to the values of \(v^m_{ijt}, (i,j) \in A, m \in M, t \in T\) obtained by solving problem (3.23) and solve problem (3.25). The solution will provide an upper bound on \(z^*\).

### 3.4 Description of data

In this section, we describe the data for our case. Several people have helped us in obtaining the necessary data. The firm Maabjerg BioEnergy has helped with detailed information about the bioethanol project and the numbers related to the plant, Knowledge Centre for Agriculture in Denmark has helped with overall aggregated numbers for the farmers, the briquette manufacturer C.F. Nielsen has kindly supplied us with detailed information about briquette stations,
Algorithm 1: Procedure Lagrangian dual

\textbf{Input}: Problem (3.23) and problem (3.25)

\textbf{Output}: Lower and upper bound to problem (3.22)

1 Initialization:
2 Iteration counter: \( p = 0 \)
3 \((\lambda_{ijm}^p)^p = 0, \ t \in \mathcal{T}, m \in \mathcal{M}, (i,j) \in \mathcal{A}.\)

4 while Stop criterion not fulfilled do
5 \quad Solve problem (3.23) with \( \lambda_{ijm}^p = (\lambda_{ijm}^p)^p, \ t \in \mathcal{T}, m \in \mathcal{M}, (i,j) \in \mathcal{A}. \)
6 \quad for \( t \in \mathcal{T} \) and \( m \in \mathcal{M} \) do
7 \quad \quad Solve problem (3.25), with \( \lambda_{ijm}^p = (\lambda_{ijm}^p)^p, \ (i,j) \in \mathcal{A}. \)
8 \quad \quad /* The present value of the Lagrangian relaxation is \( z_{LR}(\lambda^p) = z^1(\lambda^p) + \sum_{t \in \mathcal{T}} \sum_{m \in \mathcal{M}} z^2_{tm}(\lambda^p). \) This provides a lower bound on \( z. \) */
9 \quad if \( \bar{v}_{ijt}^m = \bar{v}_{ijt}^m, \ \forall t \in \mathcal{T}, m \in \mathcal{M}, (i,j) \in \mathcal{A} \) then
10 \quad \quad STOP. An optimal solution has been found.
11 \quad \quad for \( t \in \mathcal{T} \) and \( m \in \mathcal{M} \) do
12 \quad \quad \quad Solve problem (3.25), with \( \bar{v}_{ijt}^m = \bar{v}_{ijt}^m \) and \( \lambda_{ijm}^p = 0, \ (i,j) \in \mathcal{A}. \)
13 \quad \quad \quad /* The upper bound is \( z^{up} = z^1(0) + \sum_{t \in \mathcal{T}} \sum_{m \in \mathcal{M}} z^2_{tm}(0) \) */
14 \quad Use the subgradient procedure to update \( (\lambda_{ijm}^p)^p, \ t \in \mathcal{T}, m \in \mathcal{M}, (i,j) \in \mathcal{A}: \)
15 \quad Steplength: \( s_p = r_p(z^{up} - z_{LR}(\lambda))/\sum_{ijm} (\bar{v}_{ijt}^m - \bar{v}_{ijt}^m)^2, \) where \( r_p \in [0,2] \) is a scalar and \( z^{up} \) is the best known upper value on \( z. \)
16 \quad \quad \quad \quad (\lambda_{ijm}^p)^{p+1} = (\lambda_{ijm}^p)^p + s_p(\bar{v}_{ijt}^m - \bar{v}_{ijt}^m), \ t \in \mathcal{T}, m \in \mathcal{M}, (i,j) \in \mathcal{A}. \)
17 \quad \quad \quad \quad Let \( p = p + 1 \) and continue until stop criterion fulfilled.
18 \quad \quad \quad /* Continue with step 1 (solve problem (3.23) with the new values of \( \lambda \)) until stop criterion is reached. */
19 Return lower and upper bound to problem (3.22)
and the Department of Engineering - Operations Management, at Aarhus University, has contributed with knowledge sharing. In the following subsections, we will explain the foundation for the data input to the model.

3.4.1 Optimisation horizon

Since the life cycle of biomass production in the form of straw, with replanting, growth, and harvesting, is one year, the optimisation horizon for this real world problem will also be one year.

3.4.2 Demand and the plant

Maabjerg BioEnergy plans to produce around 80 million litre bioethanol for which they will need approximately 300,000 tons of straw evenly distributed over the year. At the plant, there will be enough capacity for storing straw in order to operate the plant for 65 hours.

3.4.3 Farming sites

The location of the farming sites and how much of straw that is available to Maabjerg BioEnergy is estimated by the Knowledge Centre for Agriculture in Denmark. The procedure for estimating straw availability involves three calculations. First, the total available amount of straw has to be estimated based on crop production by the farmers. Second, the technical available amount of straw is 80% of the produced amount. Third, due to local conditions such as price relations and competition not all technical available straw will be available to Maabjerg BioEnergy. Hence, the technical available straw has to be reduced by a certain percentage (10-50%) depending on the distance to Maabjerg (the longer the distance the higher the reduction). The Knowledge Centre for Agriculture estimates that there will be enough available straw to feed the ethanol plant for one year within a range of up to 80 kilometres from Maabjerg.

Since straw is a homogeneous good, we assume the price of straw paid to the farmers to be equal no matter where the straw is purchased, and hence, the price can be treated like a constant. Therefore, we do not include this price in the calculation. The handling cost of the
straw at the farmers is included in the price paid for the straw and is therefore also excluded from the model.

### 3.4.4 Converting and storage facilities

We will assume that the farming sites are possible locations for the converting and storage facilities. Often there are old barns or similar that can easily be rebuilt at a low cost. Furthermore, if there is a need for either storage or the possibility to convert the straw, it is likely that the farming sites will be a fine place to install a facility, since the farmer next to the facility will not have any transportation cost of significance.

The cost of building the facilities depends on whether the facilities need to be able to manage briquettes or only straw. In case of briquettes, the barn needs to have a solid floor and the cost of installing a briquette machine also has to be included in the fixed cost for the facility. We have linearised the cost of building such facilities even though it is not linear. Since we do not know in which size to build, and since the buildings are not standardised but can be customised into many sizes, we have made this simplification.

<table>
<thead>
<tr>
<th>Facility</th>
<th>Fixed cost</th>
<th>Variable cost</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Building cost</td>
<td>Installation cost</td>
</tr>
<tr>
<td>Straw</td>
<td>230</td>
<td>-</td>
</tr>
<tr>
<td>Briquettes</td>
<td>150</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 3.1 displays the average fixed cost per ton for building a facility and the average variable cost per ton processed at the facility. The fixed cost is an estimation based on a depreciation of 10% and an interest rate of 5% per year. We have taken the 15% of the total cost of building a facility and divided it with the capacity. Since the cubic measure of straw is higher than the cubic measure of briquettes, the cost of a facility is higher for straw than for briquettes. We have made the same exercise with the fixed cost for installing a briquette machine.

The variable cost in Table 3.1 is an average cost per ton for the indirect cost of salary, electricity, and water used to fabricate the briquettes. The variable cost depends on the sizes of the briquette machines. Its usage of electricity and water is not necessarily linear nor is the number
of employees needed to operate the machines. However, due to the complexity of the model, we have made a linear estimate. We have chosen three different machine sizes and for each of them divided the variable cost with the throughput of the machine. Afterwards, we have taken the average between the three variable unit costs and used it as the estimate for the variable cost in our case.

3.4.5 Inventory carrying cost

The inventory carrying cost in the model is the cost of having one unit in inventory for one time period. The cost is usually expressed on a yearly basis and measured as a percentage of the inventory items’ value. The percentages differ among products and depend on the durability of the product. Straw is not as perishable as other biomass products if it is kept dry, and therefore we set the percentage at 20% which is a common percentage for carrying cost.

A ton of straw can be sold for approximately €80, which gives us a carrying cost of €16 per year and €0.044 per day.

3.4.6 Transportation cost

Transportation cost covers distance travelled, driving time, and time for loading and unloading the trucks. The cost of the distance travelled is calculated by number of tons transported times the number of kilometres times the cost of diesel per kilometre per ton of load. When transporting the straw, it is important to distinguish between the truck and the load on the truck, since the cost of driving the truck depends on its weight. Table 3.2 shows the weight of the truck and the loads separately. Hence, in case of transporting bales of straw, the truck will weigh 27.7 tons, and in the case of briquettes, it will weigh 47.5 tons. The weight of the truck is then multiplied with the diesel usage per ton per kilometre in order to get the diesel usage of the truck per kilometre of the distance.

<table>
<thead>
<tr>
<th>Empty truck</th>
<th>Straw load</th>
<th>Briquette load</th>
<th>Diesel usage</th>
<th>Cost of diesel</th>
<th>Fixed cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>14.5 ton</td>
<td>13.2 ton</td>
<td>33 ton</td>
<td>0.0131 ton/km</td>
<td>€1.3</td>
<td>€260</td>
</tr>
</tbody>
</table>
Distance travelled, $d_{pq}$, can be obtained by either of two methods; either by using road network distances from Google Maps or by calculating the Euclidian distances

$$d_{pq} = G\left((x^1_p - x^1_q)^2 + (x^2_p - x^2_q)^2\right)^{\frac{1}{2}}, \quad (3.26)$$

where $(x^1_p, x^2_p)$ and $(x^1_q, x^2_q)$ are coordinates of site $p$ and $q$, respectively. $G$ is a constant used for correcting the Euclidian distances for any abnormality in the landscape (Gelders et al. (1987)).

Fixed cost for the trucks is included in the optimisation model. We have assumed that the number of trucks can change from day to day, meaning that the trucks are assumed to be leased and therefore the fixed cost represents the leasing fee.

The transportation time can be estimated by means of Google Maps or calculated using an average transport speed which is estimated to be 51.5 km/hour. In the transportation time, time for unloading and loading needs to be included. The transportation time can be converted to cost, using the salary of the driver.

### 3.5 Solution of the prototype example

![Figure 3.2: Locations used for the farming sites in the test instances as well as the address for the depot and plant: Source Google Maps.](image)

In this section, we solve the model presented in Section 3.2 on small test instances. We apply Algorithm 1 as solution method and compare its output to the solution obtained by
letting CPLEX solve the whole model of Section 3.2. Furthermore, we investigate how good a lower bound Lemma 2 provides when applied on the model of Section 3.2.

We have created four test instances. We will refer to them as \( z_1 \times z_2 \times z_3 \), where \( z_1 \) is the number of farmers, \( z_2 \) is the number of possible facility locations, and \( z_3 \) is the number of time periods. All four test instances are smaller versions of the real case, where the main difference is the number of nodes and time periods. The first two test instances have nine farmers and nine possible locations for the facilities. These two test instances are differentiated by the number of time periods; the first test set has a single time period and the second has two time periods. The two other test instances have both thirteen farmers and thirteen possible locations for the facilities. These two test instances are also differentiated by the number of time periods; they have one and two time periods, respectively.

In the test instances \( 9 \times 9 \times z_3 \), we have lowered the demand at the plant to 500 tons per time period. The demand is lowered in order to differentiate the instances \( 9 \times 9 \times z_3 \) and \( 13 \times 13 \times z_3 \) even more, and to make the smaller instances easier to solve. Furthermore, we have not included the cost of the drivers and loading and unloading times at the nodes in any of the test instances. Fig. 3.2 on the preceding page shows a map of the locations of the depot, the plant, and all the farmers for our examples. For the \( 9 \times 9 \times z_3 \) instances, we have excluded four of the farmers shown in Fig. 3.2. The depot is in the middle of the map and is placed in the city center of Holstebro. The plant is the node just above the depot.

We run the model on a computer with an Intel Core 2.30GHz processor and 4 GB RAM. The model is implemented in the 64 bit GAMS framework version 24.2.3 for Windows using the CPLEX 12.6 solver.

First, we solve each of the four test instances, where we apply CPLEX to the whole model of Section 3.2. If CPLEX is not able to find an optimal solution within an allowed time limit, CPLEX returns the best known solution (BS) and a gap between the best known solution and the best known lower bound.

Then we solve the four instances with the proposed solution method of Algorithm 1, where we apply Lemma 1 to reduce the size of the routing problems. The algorithm is stopped in case an optimal solution is found or if a maximum of 100 iterations is reached. The solution
method finds a lower bound (LB). Furthermore, for \((i, j) \in A, m \in M, t \in T\), we find an upper bound (UB) by fixing the values \(\bar{v}_{ijt}^m\) in problem (3.25) to the values of \(v_{ijt}^m\) obtained by solving problem (3.23), where in both problems \(\lambda_{ijt}^m = 0\). This approach can, in principle, be used to find a new upper bound each time a new lower bound is found, but due to high computation time, we only compute an upper bound once. The upper bound is used as \(z_{up}\) in Algorithm 1.

Finally, we find a lower bound (LB) by applying Lemma 2 to the full model in Section 3.2. Again, we find an upper bound (UB) by repairing the solution of the lower bound into a feasible solution. This is done by fixing the values \(v_{ijt}^m\) and \(\bar{v}_{ijt}^m\), \((i, j) \in A, m \in M, t \in T\), to the values \(v_{ijt}^m\), \((i, j) \in A, m \in M, t \in T\), given by the solution of Lemma 2, and then solving problem (3.23) and problem (3.25).

The results from each of the three methods are displayed in Table 3.3. The model presented in Section 3.2 is a very difficult problem to solve as mentioned in the introduction. The real complexity of the model is indicated by the results given by CPLEX, columns 2-4 in Table 3.3. CPLEX is not able to find a proven optimal solution in any of the four test instances within the allowed time limits. For the first three test instances, we stop CPLEX after four hours and for the last instance after 16 hours. CPLEX is able to find feasible solutions in the first three instances. If only CPLEX is used to solve the problems, one may be in doubt about the quality of the feasible solutions obtained, because the gaps between the upper bound and the best lower bound found are rather high. However, looking at the results obtained by Algorithm 1, columns 5-8 in Table 3.3, it seems that CPLEX in particular fails to determine a good lower bound. In the last instance, it does not succeed in finding a feasible solution.

Algorithm 1 is not able to find a proven optimal solution in any of the four test instances either, see columns 5-8 in Table 3.3. However, the algorithm provides us with an upper and
a lower bound for each instance where the gap is of a reasonable quality. The CPU time and
the gaps are significantly better than those provided by CPLEX. However, the upper bound
provided by Algorithm 1 is slightly worse than or equal to the upper bounds provided by
CPLEX, except in the case where CPLEX did not find a feasible solution, i.e., test instance
13 × 13 × 2. If Lemma 1 is not applied in connection with Algorithm 1, then the algorithm uses
too much time to solve the test instances. We tried to apply Algorithm 1 to test instance 9×9×1
without applying Lemma 1. After 48 hours, the algorithm was still running, so we stopped
it. This indicates that it may be worthwhile to apply preprocessing rules each time problems
(3.23) and (3.25) are solved.

Computing the lower bound of Lemma 2 takes only a negligible amount of time for the test
instances considered here. Therefore, the CPU times displayed in column 9 of Table 3.3 are the
times used to find the upper bound. This method performs in general worse than Algorithm 1
with respect to the lower bounds and the gaps, see columns 11-12 in Table 3.3. However, for
the last instance, the method finds the best upper bound and is much faster than Algorithm 1
and CPLEX, and it has the best found upper bound.

Fig. 3.3 on the facing page shows the routes constructed by the upper bound solution given
by Algorithm 1 for the instance 9 × 9 × 1. No new facility is opened, and a total of 38 truckloads
of straw are transported to the plant from a total of three farmers.

In general, the structure of the routes given by the upper bound solutions of Algorithm 1 to
each of the four instances is similar to those shown in Fig. 3.3. The trucks drives in loops from
the closest farmers to the plant. According to the upper bound solutions, it is not beneficial in
any of the instances to open new facilities.

3.6 Future work

We have not yet received all data for the farmers, and hence we cannot apply our solution
method to the full case at the moment. But once obtained we need to investigate how well our
solution method works on the full size problem.

The routing problems are the difficult part to solve in our solution method. The CPU time
has been improved by Lemma 1. But in order to solve the real problem, we may need to
find ways to improve the CPU time even further while still obtaining optimal solutions to the routing problems so that the Lagrangian relaxation gives a valid lower bound.

The routing problem used to find the upper bound is a Full Truckload Transportation Problem (FTTP), and it can be rewritten as a multiple Traveling Salesman Problem (mTSP) as shown by Desrosiers et al. (1988). The mTSP is also a NP-hard problem, but for smaller cases, there exist good exact solution algorithms which may speed up the solution time. A drawback, though, is that transforming the routing problem into the mTSP requires an introduction of a node for each truckload of a given technology that has to be transported during one time period. If the problem finds that no briquette facilities have to be installed, then 300,000 tons of straw have to be transported with trucks that can carry 13.2 tons per load. This will result in \((300,000/13.2)/365 \approx 62\) truckloads to be transported during one time period. Which in turn results in 62 nodes in the mTSP and then this problem will also be difficult to solve. However, the general routing problem in (3.25) cannot be viewed as a traditional FTTP, since the truckloads are variables, and hence the problem in (3.25) cannot be transformed into the mTSP.

In Arunapuram et al. (2003), they formulate the FTTP in a different way than the one presented in this paper. Their formulation can be applied to the routing problem in (3.25), and it would be interesting to see if their solution method can improve the CPU time of our model.

Another rather important aspect to consider is how to incorporate multiple objective functions. Supply chain design of biomass naturally have different objectives which are important to incorporate. It would be interesting to see how the two key drivers, cost and environment,
influence each other. To do that, we may define two objective functions; one that minimises total costs and one that minimises GHG emissions. To find a solution we need to determine all nondominated points. The nondominated points will reveal the trade-off between all Pareto optimal solutions to the bi-objective mixed integer linear program. Unfortunately, at this time, it is very hard to solve bi-objective mixed integer linear programs as we suggest here. However, a method that may be possible to use has recently been published, see Stidsen et al. (2014). That method is being developed further to take advantage of parallel processing.
Bibliography


Energinet.dk (September 2010). *Energi 2050 - Udviklingsspørg for energisystemet*. Energinet.dk, Denmark.


PhD Theses since 1 July 2011

2011-4  Anders Bredahl Kock: Forecasting and Oracle Efficient Econometrics
2011-5  Christian Bach: The Game of Risk
2011-6  Stefan Holst Bache: Quantile Regression: Three Econometric Studies
2011:12 Bisheng Du: Essays on Advance Demand Information, Prioritization and Real Options in Inventory Management
2011:13 Christian Gormsen Schmidt: Exploring the Barriers to Globalization
2011:16 Dewi Fitriasari: Analyses of Social and Environmental Reporting as a Practice of Accountability to Stakeholders
2011:22 Sanne Hiller: Essays on International Trade and Migration: Firm Behavior, Networks and Barriers to Trade
2012-1 Johannes Tang Kristensen: From Determinants of Low Birthweight to Factor-Based Macroeconomic Forecasting
2012-2 Karina Hjortshøj Kjeldsen: Routing and Scheduling in Liner Shipping
2012-3 Soheil Abginehchi: Essays on Inventory Control in Presence of Multiple Sourcing
2012-4 Zhenjiang Qin: Essays on Heterogeneous Beliefs, Public Information, and Asset Pricing
2012-5 Lasse Frisgaard Gunnersen: Income Redistribution Policies
2012-6 Miriam Wüst: Essays on early investments in child health
2012-7 Yukai Yang: Modelling Nonlinear Vector Economic Time Series
2012-9 Henrik Nørholm: Structured Retail Products and Return Predictability
2012-10 Signe Frederiksen: Empirical Essays on Placements in Outside Home Care
2012-11 Mateusz P. Dziubinski: Essays on Financial Econometrics and Derivatives Pricing
2012-12 Jens Riis Andersen: Option Games under Incomplete Information
2012-13 Margit Malmmose: The Role of Management Accounting in New Public Management Reforms: Implications in a Socio-Political Health Care Context
2012-14 Laurent Callot: Large Panels and High-dimensional VAR
2012-15 Christian Rix-Nielsen: Strategic Investment
2013-1 Kenneth Lykke Sørensen: Essays on Wage Determination
2013-2 Tue Rauff Lind Christensen: Network Design Problems with Piecewise Linear Cost Functions
2013-4 Rune Bysted: Essays on Innovative Work Behavior
2013-5 Mikkel Nørlem Hermansen: Longer Human Lifespan and the Retirement Decision
2013-7 Mark Strøm Kristoffersen: Essays on Economic Policies over the Business Cycle
2013-8 Philipp Meinen: Essays on Firms in International Trade
2013-9 Cédric Gorinas: Essays on Marginalization and Integration of Immigrants and Young Criminals – A Labour Economics Perspective
2013-12 Paola Andrea Barrientos Quiroga: Essays on Development Economics
2013-13 Peter Bodnar: Essays on Warehouse Operations
2013-14 Rune Vammen Lesner: Essays on Determinants of Inequality
2013-15 Peter Arendorf Bache: Firms and International Trade
2013-16 Anders Laugesen: On Complementarities, Heterogeneous Firms, and International Trade
2013-17  Anders Bruun Jonassen: Regression Discontinuity Analyses of the Disincentive Effects of Increasing Social Assistance

2014-1  David Sloth Pedersen: A Journey into the Dark Arts of Quantitative Finance


2014-3  Lukas Bach: Routing and Scheduling Problems - Optimization using Exact and Heuristic Methods

2014-4  Tanja Groth: Regulatory impacts in relation to a renewable fuel CHP technology: A financial and socioeconomic analysis

2014-5  Niels Strange Hansen: Forecasting Based on Unobserved Variables

2014-6  Ritwik Banerjee: Economics of Misbehavior

2014-7  Christina Annette Gravert: Giving and Taking – Essays in Experimental Economics

2014-8  Astrid Hanghøj: Papers in purchasing and supply management: A capability-based perspective

2014-9  Nima Nonejad: Essays in Applied Bayesian Particle and Markov Chain Monte Carlo Techniques in Time Series Econometrics

2014-10  Tine L. Mundbjerg Eriksen: Essays on Bullying: an Economist’s Perspective

2014-11  Sashka Dimova: Essays on Job Search Assistance

2014-12  Rasmus Tangsgaard Varneskov: Econometric Analysis of Volatility in Financial Additive Noise Models

2015-1  Anne Floor Brix: Estimation of Continuous Time Models Driven by Lévy Processes

2015-2  Kasper Vinther Olesen: Realizing Conditional Distributions and Coherence Across Financial Asset Classes

2015-3  Manuel Sebastian Lukas: Estimation and Model Specification for Econometric Forecasting

2015-4  Sofie Theilade Nyland Brodersen: Essays on Job Search Assistance and Labor Market Outcomes

<table>
<thead>
<tr>
<th>Year</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015-6</td>
<td>Sanni Nørgaard Breining: The Sibling Relationship Dynamics and Spillovers</td>
</tr>
<tr>
<td>2015-7</td>
<td>Marie Herly: Empirical Studies of Earnings Quality</td>
</tr>
<tr>
<td>2015-8</td>
<td>Stine Ludvig Bech: The Relationship between Caseworkers and Unemployed Workers</td>
</tr>
<tr>
<td>2015-9</td>
<td>Kaleb Girma Abreha: Empirical Essays on Heterogeneous Firms and International Trade</td>
</tr>
<tr>
<td>2015-10</td>
<td>Jeanne Andersen: Modelling and Optimisation of Renewable Energy Systems</td>
</tr>
</tbody>
</table>