Essays on stochastic modeling in electricity markets

With applications to wind power and market interconnectedness

Pircalabu, Anca

DOI (link to publication from Publisher):
10.5278/vbn.phd.eng.00027

Publication date:
2017

Document Version
Publisher's PDF, also known as Version of record

ESSAYS ON STOCHASTIC MODELING IN ELECTRICITY MARKETS

WITH APPLICATIONS TO WIND POWER AND MARKET INTERCONNECTEDNESS

BY
ANCA PIRCALABU

DISSERTATION SUBMITTED 2018

AALBORG UNIVERSITY
DENMARK

ESSAYS ON STOCHASTIC MODELING IN ELECTRICITY MARKETS

WITH APPLICATIONS TO WIND POWER AND MARKET INTERCONNECTEDNESS

PhD Dissertation

Anca Pircalabu

Dissertation submitted January 2, 2018

Dissertation submitted: January 2, 2018

PhD supervisors: Assoc. Prof. Esben Høg
Department of Mathematical Sciences
Aalborg University

Jesper Jung, PhD
Neas Energy

Thomas A. Fredholm, PhD
Neas Energy

PhD committee: Professor Rasmus Waagepetersen (chairman)
Aalborg University

Professor Asger Lunde
Aarhus University

Professor Mark Shackleton
Lancaster University

PhD Series: Faculty of Engineering and Science, Aalborg University

Department: Department of Mathematical Sciences

ISSN (online): 2446-1636
ISBN (online): 978-87-7210-123-1

Published by:
Aalborg University Press
Langagervej 2
DK – 9220 Aalborg Ø
Phone: +45 99407140
aauf@forlag.aau.dk
forlag.aau.dk

Printed in Denmark by Rosendahls, 2018

Preface

This thesis unifies the research carried out during my years as an Industrial PhD student at the Department of Mathematical Sciences, Aalborg University, and Neas Energy. My studies have been funded jointly by Neas Energy and Innovation Fund Denmark through the Industrial PhD program.

The thesis consists of a collection of research papers written in the period from February 2015 to January 2018. Each paper is self-contained, with a separate bibliography. While the overall theme of the thesis is modeling in electricity markets, particular focus is on wind power and the price coupling of day-ahead electricity markets. Empirical analyses on actual market data constitute a significant part of each paper, and the developed models are applied for purposes such as the pricing of financial instruments, hedging and risk management.

There are a number of people that have influenced and helped shape this thesis, and who I wish to acknowledge. First and foremost, I wish to thank my university supervisor Esben Høg, company supervisor Jesper Jung and company co-supervisor Thomas Aalund Fredholm for their continued support and excellent guidance over the years. A special thanks to Esben for establishing the contact to Neas Energy and for encouraging me to take on this project, and to Jesper and Thomas for allowing me the opportunity to pursue the PhD degree. Without your help and support, I would not be where I am today.

Thanks are also extended to my colleagues at Neas Energy and the Department of Mathematical Sciences, Aalborg University, for providing an inspiring work environment and for all the help and assistance in various matters. Christian Sønderup, Jakob Vive Munk and my colleagues in Quantitative Analytics, Sebastian Christensen and Rune Elgaard Mikkelsen, deserve special recognition for their valuable comments, feedback on my papers, and all the conversations that served as a great source of inspiration. Further, thanks to my fellow PhD students for all the fun hours spent together during PhD courses, coffee and lunch breaks, and in particular Rune Hjorth Nielsen for all the interesting discussions on the German electricity market, and Thomas Hvolby for our collaboration and friendship. Fellow industrial

PhD student, friend and co-author Troels Sønderby Christensen deserves a special thanks for the great work and effort he put in our two joint papers.

From August to December 2016, I had the privilege of visiting Fred Espen Benth at the Department of Mathematics, Oslo University. I am truly grateful for this opportunity, and I wish to thank Fred for our collaboration and for being an excellent host. Moreover, I thank Fred for his continued support and the many stimulating conversations during as well as after my stay in Oslo. To everyone I had the pleasure of meeting at the Department of Mathematics, Oslo University, thank you for your hospitality.

Last but definitely not least, a special thank you is reserved for my mom Carmen and my boyfriend Søren for their tremendous patience, unconditional love and support.

Anca Pircalabu
Aalborg University, January 2, 2018

Thesis details

Thesis title:	Essays on stochastic modeling in electricity markets - with applications to wind power and market interconnectedness
PhD student:	Anca Pircalabu
PhD supervisors:	Assoc. Prof. Esben Høg, Department of Mathematical Sciences, Aalborg University Jesper Jung, PhD, Neas Energy Thomas A. Fredholm, PhD, Neas Energy

This thesis is based on the six papers enumerated below:

Paper I. A. Pircalabu, T. Hvolby, J. Jung and E. Høg. Joint price and volumetric risk in wind power trading: A copula approach. *Energy Economics*, 62:139–154, 2017.

Paper II. A. Pircalabu and J. Jung. A mixed C-vine copula model for hedging price and volumetric risk in wind power trading. *Quantitative Finance*, 17(10):1583–1600, 2017.

Paper III. A. Pircalabu and F. E. Benth. A regime-switching copula approach to modeling day-ahead prices in coupled electricity markets. *Energy Economics*, 68:283–302, 2017.

Paper IV. F. E. Benth and A. Pircalabu. A non-Gaussian Ornstein-Uhlenbeck model for pricing wind power futures. Submitted to *Applied Mathematical Finance*.

Paper V. T. S. Christensen, A. Pircalabu and E. Høg. A seasonal copula mixture for hedging the clean spark spread with wind power futures. Submitted to *Energy Economics*.

Paper VI. T. S. Christensen and A. Pircalabu. On the spatial hedging effectiveness of German wind power futures for wind power generators. Submitted to *Journal of Energy Markets*.

Summary

This thesis consists of an introductory section and six self-contained papers on different electricity-related topics. Particular focus is on the volumetric risk associated with wind power generation, how wind power interacts with the market, and the price coupling of electricity markets.

Papers I and II study the problem of joint price and volumetric risk in the context of energy trading companies entering into long-term commitments with wind power generators, where the fluctuating wind power production is bought at a pre-determined fixed price. Multiple aspects concerning this type of commitment are studied, such as determining its fair value, quantifying its risk, and identifying a suitable hedging strategy. To address these issues, a model for the joint behavior of prices and wind power production is established and fitted to the Danish power market primarily. The performed empirical investigations not only highlight the importance of capturing the negative dependence between day-ahead electricity prices and wind power production for many applications, but also allow for quantifying its effect. As far as hedging is concerned, an implementable strategy that outperforms the industry standard is proposed.

Paper III deals with the structural change in the joint behavior of spot electricity prices that came with the price coupling of European electricity markets. In addition to proposing a model that accounts for this change, we focus on capturing the non-linear dependence between price pairs. One particularly interesting finding is that the probability of simultaneous extreme events in price coupled areas is, in many cases, high. Through practically relevant applications such as the pricing of financial transmission rights and the forecasting of tail risk, we illustrate the benefits of our proposed model.

Papers IV, V and VI are all centered around a particular weather derivative that has only recently been introduced to the market. With the focus on green growth that many European countries have experienced, a financial instrument that can mitigate the volume risk associated with the generation of wind electricity has long been missing; as a result, German-based wind power futures were launched, with plans of extending the concept to other markets. While Paper IV studies the pricing of such instruments and answers

questions related to the risk premium in this newly established market, Papers V and VI relate to the detailed study of the actual benefits of wind power futures for the buyer and seller group, respectively.

Resumé

(Danish Summary)

Denne afhandling består af en introduktion og seks selvstændige artikler omhandlende forskellige energirelaterede emner. Der er særligt fokus på volumenrisiko i vindkraftproduktion, hvordan vindkraft interagerer med markedet, og priskoblingen mellem europæiske elmarkeder.

Artikel I og II omhandler pris- og volumenrisiko i forbindelse med energivirksomheder, der indgår langsigtede aftaler med vindkraftproducenter, hvor den varierende vindkraftproduktion købes til en forudbestemt fast pris. Flere aspekter vedrørende en sådan type aftale undersøges, såsom beregning af dens fair værdi, kvantificering af dens risiko, samt fastlæggelse af hedging strategier. For at løse disse problemer etableres en simultan model for elpriser og vindkraftproduktion, som primært anvendes på det danske marked. Dernæst fremhæves det via empiriske undersøgelser at den negative afhængighed mellem elpriser og vindkraftproduktion er vigtig at fange. Hvad angår hedging er der foreslået en implementerbar strategi, der overgår industristandarden.

Artikel III omhandler den strukturelle ændring der er opstået som følge af indføreselen af en simultan prisformation i de europæiske elmarkeder. Vi foreslår en model der inkluderer denne ændring, og som fokuserer på at indfange den ikke-lineære afhængighed mellem prispar. Et særligt interessant resultat er, at sandsynligheden for samtidige ekstreme begivenheder i priskoblede områder i mange tilfælde er høj. Via praktiske anvendelser som prisfastsættlese af finansielle transmissionsrettigheder og forecasting af halerisiko, illustrerer vi fordelene ved den foreslåede model.

Artikel IV, V og VI er alle centreret omkring et bestemt vejrderivat, som for nylig er blevet introduceret på børsen. Med det store fokus på grøn vækst som mange europæiske lande har oplevet, har der længe manglet et finansielt instrument, der kan nedsætte volumenrisikoen forbundet med vindkraftproduktion. Som følge heraf blev tyskbaserede wind power futures lanceret, med planer om at udvide konceptet til andre markeder. Mens Artikel IV behandler prisfastsættelsen af sådanne instrumenter og besvarer spørgsmål

vedrørende risikopræmien i dette nyoprettede marked, undersøger Artikel V og VI de faktiske fordele ved wind power futures for henholdsvis køber- og sælgergruppen.

Contents

Part A

Background

Introduction

This section provides a brief introduction to each of the six papers included in this thesis, together with the motivation behind each paper. Why and how the papers are linked to each other is commented on, and wherever applicable, remarks on the latest related research are included.

1 Joint price and volumetric risk in wind power trading

Owing to the rapid rise of wind power that many countries have experienced over the past decades, the interaction between wind power generation and the market has become important in a wide range of applications and for different types of market players. The first problem studied in this thesis is concerned with the joint price and volumetric risk in wind power trading. The perspective considered is that of energy trading companies entering into long-term commitments that involve buying the fluctuating wind power generation at a fixed price.

Unlike typical commodities like e.g. oil, metals, and gas, electricity cannot be stored directly and is consumed instantaneously and continuously. As a results, the link between electricity prices and the underlying price drivers can be very strong. One situation where this link is clearly reflected relates to the mechanism of day-ahead electricity price formation. In deregulated electricity markets, day-ahead prices are set by matching supply and demand, where supply curves are constructed according to the merit order principle. This price-setting mechanism prioritizes the cheap renewable sources, and implies that a large share of wind power in the electrical grid will pull the clearing price downwards. Similarly, a small share of wind power generation will set an upward pressure on the clearing price. For energy trading companies with the above-named fixed price obligations, the relation between day-ahead electricity prices and wind power production is paramount for all applications that involve pricing, hedging and risk management. The first two papers in this thesis are concerned with different aspects of this problem.

Paper I, coauthored with Thomas Hvolby, Jesper Jung and Esben Høg, proposes a model for the joint behavior of day-ahead electricity prices and wind power generation, and provides a detailed study of the dependence between the two in Denmark. Since Denmark is a top wind power producing country, conducting such analysis on the Danish power market was an obvious choice. To model the joint behavior of prices and wind power production, the flexible copula models are employed, allowing us to separate the treatment of marginals from that of the dependence structure. The flexibility of this modeling approach is reflected in the fact that we can easily capture stylized facts of the individual time series, such as seasonality, serial dependence and non-normality; concurrently, an in-depth study of the dependence structure is facilitated. The appropriateness of the proposed copula model is assessed through statistical testing, and Monte Carlo simulations enable us to address many practical concerns related to pricing and risk quantification. In particular, we show how the negative dependence between prices and wind power generation affects the profit distribution of energy trading companies with long-term commitments that involve buying the fluctuating wind power production at a fixed price. Moreover, we quantify the price of correlation risk that must be accounted for when pricing such long-term commitments.

Aside from the aspects related to pricing and risk quantification addressed in Paper I, a major concern relates to how energy trading companies can hedge their exposure to joint price and volumetric risk. While Paper I mentions this aspect, the considered hedging portfolio is simplistic in that it consists of a single product, namely the standard power forward contract.

Paper II, coauthored with Jesper Jung, extends the work carried out in the first paper by studying the hedging of joint price and volumetric risk in much more detail. Given that the aforementioned fixed-price-for-fluctuating-volume contracts give rise to a non-linear payoff, a hedging portfolio would ideally include options. However, not only is the range of available derivative products that can be used to address such risks very limited in the Danish power market, we are also faced with illiquidity, even with respect to standard power forwards. Owing to the efforts made to achieve increased market integration across electricity markets in Europe, the Danish day-ahead electricity prices are strongly related to the prices in the much more liquid German market. In Paper II, we extend the copula model in Paper I by including the German electricity price as a third variable in the model. This enables us to study the potential benefit from employ German financial instruments to hedge the Danish joint price and volumetric risk. Hence, also from a modeling perspective, Paper II extends the work in Paper I.

In Paper II, the instruments included in the hedging portfolio are standard power forwards on the Danish and German prices, wind power futures, financial transmission rights, and put options. Wind power futures are a

recent financial instrument, and can be used to mitigate the volumetric risk associated with wind power production. Financial transmission rights belong to the class of real options, and give the right to transfer electricity across the border. In more technical terms, they are zero strike spread options written on the spot price differential in two interconnected electricity markets. The put options we include in the hedging portfolio are also real options: Specifically, they relate to combined heat and power plants equipped with electric boilers, which can benefit from converting electricity to heat whenever the day-ahead electricity price is low enough. The study of the hedging effectiveness of the proposed instruments is facilitated by Monte Carlo simulations from the proposed three-dimensional copula model. We find that our proposed hedge contributes with significant additional risk-reducing benefits compared to the market standard, which entails having solely forwards in the hedging portfolio.

While the hedging application in Paper II was to some extent motivated by Oum et al. (2006), Oum and Oren (2010) and Coulon et al. (2013), the hedging of joint price and volumetric risk has attracted much attention recently, and more related studies have become available. Since these studies present perspectives and applications that are (or can become) very relevant in the context of Paper II, they should be mentioned here. Considering the topic on a more general level are the studies of Brik and Roncoroni (2016), Roncoroni and Brik (2017) and Lange (2017), where interesting alternative strategies are proposed. The studies of Ernstsen et al. (2017) and Tegnér et al. (2017) consider the hedging of price and volumetric risk on the buyer side, i.e., distribution companies that cover the fluctuating electricity demand of customers at a fixed price. Empirical analyses are performed on the Danish power market, where optimal hedging strategies implied by different risk measures are compared. Since many of the ideas are directly applicable to the problem in Paper II, it would be interesting to investigate whether the conclusions in Ernstsen et al. (2017) and Tegnér et al. (2017) are transferable to the seller side of price and volumetric risk in wind power trading.

In the hedging exercise of Paper II, two of the employed hedging instruments stood out, namely the *financial transmission rights* and the *wind power futures*. The special features and the many unexplored aspects regarding these instruments called for a separate in-depth treatment, which motivated greatly the subsequent work in this thesis. In the following two sections, we elaborate on how these instruments are included in our studies.

2 Financial transmission rights and the price coupling of European electricity markets

Financial transmission rights have long been a topic of interest in the energy-related literature. Recently however, a structural reorganization of many electricity markets in Europe has affected all pricing and decision-making that relies on the joint behavior of day-ahead electricity prices. This structural change relates to the price coupling of areas in order to achieve increased integration. Under price coupling, the supply and demand curves of individual areas are considered simultaneously, according to the overall merit order and the available transmission capacity. As a result, the joint behavior of day-ahead electricity prices exhibits a particular pattern compared to other economical or financial data: If the available transmission capacity between two countries is not fully utilized (non-congestion) for a given hour of the day, exact price convergence is achieved; otherwise, we could keep on transferring electricity from the cheap to the expensive area until the difference in prices is eliminated. Prices will thus only start deviating from each other once the available transmission capacity is fully utilized (congestion).

In **Paper III**, coauthored with Fred Espen Benth, we consider the problem underlying the pricing of financial transmission rights under price coupling: The modeling of pairs of electricity prices in interconnected areas. Since the price coupling of most electricity markets dates only a few years back, models for the joint behavior of spot prices that include the implications of this change are very scarce [some examples are Füss et al. (2017) and Kiesel and Kustermann (2016)]. Inspired by ideas in Haldrup and Nielsen (2006) and Haldrup et al. (2010), we propose a regime-switching AR–GARCH copula model for pairs of day-ahead electricity prices. Since equality of prices implies non-congestion, and non-equality of prices implies congestion, the regimes are easily identifiable, and most importantly, they are observable. While the AR–GARCH filter describes the conditional serial dependence and heteroscedasticity, the copula in included to model the non-linear dependence in the state of non-equal prices.

In a study of various European price pairs, we find the congestion state to have significant tail dependence. As applications of the proposed model we consider the pricing of financial transmission rights (in more detail compared to Paper II) and the out-of-sample forecasting of tail quantiles. We highlight through comparison studies the importance of capturing effects such as heavy tails, asymmetric dependence and tail dependence. The study in Paper III differs from the rest in that it does not concern wind power; instead, it focuses on the relation between pairs of day-ahead prices in interconnected markets, and the challenges that followed the launch of certain initiatives to further market integration.

3 Wind power futures and hedging weather risk

The last three papers in this thesis are concerned with the recently introduced German wind power futures. As a result of the installed wind power capacity growing in many electricity markets, so has the demand for financial instruments that can address directly the volume risk associated with wind power generation. To meet this demand, wind power futures were recently introduced in Germany, with plans of later extending to e.g. Denmark and the UK. Clearly, the wind power producer's profit is directly affected by how the wind blows; and since he wishes to lower his exposure to low wind scenarios, which (with everything else being equal) lower his profit, he is a natural seller of wind power futures. The conventional generator, who would have his competitiveness on the day-ahead market weakened by a large share of cheap renewable generation in the grid, is on the other hand a natural buyer of wind power futures. While the literature on weather derivatives generally speaking is rich [Davis (2001), Cao and Wei (2004), Campbell and Diebold (2005), Brody et al. (2002), Benth and Benth (2011) and Gersema and Wozabal (2017) to name a few], work specifically directed at wind derivatives is very limited, and our analyses are among the first attempts to study the new market for German wind power futures.

Paper IV, coauthored with Fred Espen Benth, addresses the pricing of wind power futures. Unlike the other studies in this thesis, time series models and copulas are not employed here. Instead, we propose describing the German wind power production index, which is the underlying of the wind power futures, as a product of two components: The first component is a deterministic function that takes into account the seasonal pattern clearly present in the historical wind index data; the second component consists of the exponential of an Ornstein-Uhlenbeck that is driven by an appropriate Lévy process. We discuss the properties of the model, estimation of the model parameters, and derive a closed-form pricing formula for wind power futures while also dealing with measure change. To strengthen further the advantages of the model, we also include an outline concerning the pricing of European options written on wind power futures contracts, which can be achieved through Fourier analysis. The model is then applied to German data, where the objective is to investigate the sign and magnitude of the market price of risk in this newly established market. Our empirical analysis reveals a negative risk premium, translating to the fact that wind power producers are willing to sell wind power futures at a discounted price. This finding motivated the two subsequent studies, which represent a pursue to investigate further the benefits of wind power futures for the buyer and the seller groups.

In **Paper V**, coauthored with Troels Sønderby Christensen and Esben Høg,

we return to the joint modeling of wind power generation and prices as in Papers I and II. This time however, we bring into focus the conventional generator. By considering the clean spark spread as a measure of the conventional generator's profitability, Paper V studies the benefits of wind power futures for gas-fired power plants that operate in the day-ahead electricity market during peak hours. First, a detailed study of the dependence between the clean spark spread and the wind power generation is carried out using copulas. Second, the hedging effectiveness of wind power futures is illustrated through a simplified version of the hedging decision that conventional generators are faced with every day. Third, we highlight the effects of asymmetric dependence, tail dependence, and seasonal dependence on the optimal hedge quantities and the conventional producer's hedged profit distribution.

Paper VI, coauthored with Troels Sønderby Christensen, is very similar to Paper V in terms of both modeling approach and application. In contrast however, we consider here the seller rather than the buyer side, and illustrate through simple analyses the spatial hedging benefits of wind power futures for wind power generators with different geographical locations in Germany. Although the analysis in Paper VI only differs in terms of context, it complements and completes the wind power futures theme. Also, this study highlights that while geographical conditions are unimportant for conventional generators, they can have a substantial impact for wind power generators, as far as the benefits of wind power futures are concerned.

References

F. E. Benth and J. Šaltytė Benth. Weather derivatives and stochastic modelling of temperature. *International Journal of Stochastic Analysis*, Vol. 2011:Article ID 576791, 21 pages, 2011.

R. Id Brik and A. Roncoroni. Static mitigation of volumetric risk. *Journal of Energy Markets*, 9(2):111–150, 2016.

D. Brody, J. Syroka, and M. Zervos. Dynamical pricing of weather derivatives. *Quantitative Finance*, 2(3):189 – 198, 2002.

S. D. Campbell and F. X. Diebold. Weather forecasting for weather derivatives. *Journal of the American Statistical Association*, 100(469):6 – 16, 2005.

M. Cao and J. Wei. Weather derivatives valuation and market price of weather risk. *Journal of Futures Markets*, 24(11):1065 – 1089, 2004.

M. Coulon, W. B. Powell, and R. Sircar. A model for hedging load and price risk in the Texas electricity market. *Energy Economics*, 40:976 – 988, 2013.

M. Davis. Pricing weather derivatives by marginal value. *Quantitative Finance*, 1(3):305 – 308, 2001.

R. R. Ernstsen, T. K. Boomsma, M. Jönsson, and A. Skajaa. Hedging local volume risk using forward markets: Nordic case. *Energy Economics*, 68: 490–514, 2017.

R. Füss, S. Mahringer, F. Paraschiv, and M. Prokopczuk. Electricity spot and derivatives pricing under market coupling. *University of St. Gallen, School of Finance Research Paper No. 1323*, 2017.

G. Gersema and D. Wozabal. An equilibrium pricing model for wind power futures. *Energy Economics*, 65:64 – 74, 2017.

N. Haldrup and M. Ø. Nielsen. A regime switching long memory model for electricity prices. *Journal of Econometrics*, 135(1–2):349 – 376, 2006.

N. Haldrup, F. S. Nielsen, and M. Ø. Nielsen. A vector autoregressive model for electricity prices subject to long memory and regime switching. *Energy Economics*, 32(5):1044 – 1058, 2010.

R. Kiesel and M. Kustermann. Structural models for coupled electricity markets. *Journal of Commodity Markets*, 3(1):16 – 38, 2016.

N. Lange. Correlation in energy markets. *Ph.D. thesis, Copenhagen Business School*, PhD Series, no. 08.2017, 2017.

Y. Oum and S. S. Oren. Optimal static hedging of volumetric risk in a competitive wholesale electricity market. *Decision Analysis*, 7(1):107 – 122, 2010.

Y. Oum, S. S. Oren, and S. Deng. Hedging quantity risks with standard power options in a competitive wholesale electricity market. *Naval Research Logistics*, 53(7):697 – 712, 2006.

A. Roncoroni and R. Id Brik. Hedging size risk: Theory and application to the US gas market. *Energy Economics*, 64:415–437, 2017.

M. Tegnér, R. R. Ernstsen, A. Skajaa, and R. Poulsen. Risk-minimisation in electricity markets: Fixed price, unknown consumption. *Energy Economics*, 68:423–439, 2017.

References

Part B

Papers

Paper I

Joint price and volumetric risk in wind power trading: A copula approach

List of authors: Anca Pircalabu[*,†], Thomas Hvolby[*], Jesper Jung[†] and Esben Høg[*]
[*] Department of Mathematical Sciences, Aalborg University
[†] Quantitative Analytics, Neas Energy

Published in
Energy Economics Vol. 62, pp. 139–154, 2017.
DOI: http://dx.doi.org/10.1016/j.eneco.2016.11.023

The layout has been revised.

ABSTRACT

This paper examines the dependence between wind power production and electricity prices and discusses its implications for the pricing and the risk distributions associated with contracts that are exposed to joint price and volumetric risk. We propose a copula model for the joint behavior of prices and wind power production, which is estimated to data from the Danish power market. We find that the marginal behavior of the individual variables is best described by ARMA–GARCH models with non-Gaussian error distributions, and the preferred copula model is a time-varying Gaussian copula. As an application of our joint model, we consider the case of an energy trading company entering into longer-term agreements with wind power producers, where the fluctuating future wind power production is bought at a predetermined fixed price. We find that assuming independence between prices and wind power production leads to an underestimation of risk, as the profit distribution becomes left-skewed when the negative dependence that we find in the data is accounted for. By performing a simple static hedge in the forward market, we show that the risk can be significantly reduced. Furthermore, an out-of-sample study shows that the choice of copula influences the price of correlation risk, and that time-varying copulas are superior to the constant ones when comparing actual profits generated with different models.

1 Introduction

Since the European electricity market reforms in the late 1990's, the electricity markets have undergone considerable structural changes. Liberalization has led to extremely volatile electricity prices, and the prioritization of renewable energy sources in order to reduce CO_2 emissions has introduced further challenges in terms of financial risk management. One particular challenge that we study in this paper is related to the production uncertainty associated with wind power generation. Wind power is highly non-dispatchable and therefore fundamentally different from the more traditional thermal power sources in the sense that the production cannot be planned and controlled to the same extent. The dependency on weather variations (wind speed and air density among others) makes the exact future production of a wind turbine or wind-farm very hard to predict. Thus, in addition to facing price volatility, wind power generators are exposed to production uncertainty, often referred to as volumetric risk.

The joint exposure to price and volumetric risk can be further amplified by a high penetration ratio of wind power in the grid. This is due to the mechanism of day-ahead price formation, which is based on finding the equilibrium between supply and demand bids made to the exchange, where the

supply curve is built according to merit order stack[1]. Because wind power has a very low marginal cost, a high production for a given hour will, other things being equal, pull the market clearing price downwards. Similarly, if wind power production is low for a given hour, demand will have to be met by either import or turning on more costly generating plants. The latter (and possibly the former) will, again other things being equal, pull the prices upward. This leads to prices and wind power production being negatively correlated, which depending on the strength of this correlation, enhances the joint price and volumetric risk significantly. Empirical evidence regarding this relation between spot electricity prices and wind power production has been demonstrated in the literature, e.g. Jónnson et al. (2010) for the Danish power market, Gelabert et al. (2011) for Spain, and Elberg and Hagspiel (2015) and Paraschiv et al. (2014) for Germany.

In practice, it is usually energy trading companies that act on the exchange on behalf of the producers. Due to increasing wind power production in some power markets, some trading companies offer, in addition to the management of production, a predetermined fixed price in exchange for the fluctuating production. Companies offering such insurances against price movements will naturally attempt to cover their exposure, and a typical solution that will eliminate some of the risk is to sell energy on the forward market corresponding to the expected wind power production. The remaining exposure will inevitably cause the energy trading companies to purchase energy on the spot market when being short, and dispose of excess energy on the spot market when expecting less than the realized production. Furthermore, the negative relationship between prices and wind power production adds an additional correlation risk: If being short, chances are that the missing energy will have to be bought at a higher price; similarly, if having to dispose of excess electricity, chances are that this will be sold at times of a lower price. As a result, the negative dependence between price and production introduces a "double" risk that is not straight forward to address or diminish without having a well-specified model for the dependence structure.

The problem of joint price and volumetric risk stems back some decades, and was first discussed in McKinnon (1967) in relation to the classical farmer's problem – who faces both price and production uncertainty at the time of harvest. In McKinnon (1967), the author considered futures as hedging instruments, and presented an explicit formula for the optimal position in futures contracts (from a minimum-variance perspective); this formula pointed out that the correlation between the two sources of uncertainty is an essential feature of the problem. Later, the work of Moschini and Lapan (1995) included options in the hedging portfolio due to the non-linearity of profit.

[1]Supply bids from different power stations are ranked according to their production costs, and the market clearing price corresponds to the highest bid needed to match demand.

More recently, energy related work on the subject became available, and some interesting discussions on the hedging of volumetric risk associated with consumers' load (demand-side risk) were presented in Oum and Oren (2009, 2010) and Coulon et al. (2013). In Oum and Oren (2009, 2010), the authors assumed bivariate lognormality for electricity prices and consumers' demand of electricity with a constant correlation, and focused on hedging strategies that 1) maximize the expected utility of the hedged profit and 2) maximize the expected profit subject to a Value-at-Risk constraint. In Coulon et al. (2013), the authors propose a structural model that captures the complex dependence structure of electricity price and load dynamics as a base for hedging. While many of the ideas in the existing literature regarding the hedging of volumetric risk can be used in our application, there are some major distinctions between supply-side volumetric risk (associated with wind power production) and demand-side volumetric risk (consumers' load) that pose some challenges when having to specify a joint model for day-ahead electricity prices and wind power production.

One issue of concern when considering a joint model for electricity prices and wind power production is that the price dynamics are very different from the production dynamics, causing us to expect the benchmark bivariate (log) normality assumption to be too restrictive;[2] in fact, the two variables might have univariate marginal distributions from different families, making it very challenging to decide upon a suitable bivariate density. The assumption of constant correlation might also prove too restrictive, and many studies have shown evidence of time-varying dependence between economic time series, see e.g. Avdulaj and Barunikl (2015), Salvatierra and Patton (2013), Dias and Embrechts (2004), Patton (2006), and Wen et al. (2012). Thus, before addressing issues such as the valuation of correlation risk in the context of fixed price obligations with fluctuating wind power production or the hedging of portfolios containing such obligations, a large part of this paper is concerned with developing a joint model that correctly characterizes the marginal behavior of electricity prices and wind power production and also their dependence structure. For this purpose, we propose the use of copula models.

Copulas are flexible tools that can be used to completely describe the dependence structure between random variables while allowing for arbitrary marginal distributions. They were introduced in the literature by Sklar (1959), and have found various applications in economics and finance over the past decades: See Cherubini and Luciano (2002) for the use of copulas in pricing different types of bivariate options, Embrechts et al. (1999) for an application to risk management, and Patton (2013) for a thorough review on copula-based models, including methods for estimation, inference and

[2] Alone the fact that electricity prices can go negative rules out the lognormality assumption in some marketplaces.

model-selection. Applications of copula models in energy markets are less common, but some examples are Alexander (2004), Benth and Kettler (2011), Elberg and Hagspiel (2015), and González-Pedraz et al. (2015).

Specifically, we offer two contributions: Firstly, we propose a flexible joint model that relaxes the assumption of bivariate normality and that accounts for the time variation we observe in the dependence structure. Our empirical study is based on data from the Danish power market; nonetheless, we expect our results to be generally applicable in all liberalized energy markets with a high penetration of wind power in the grid. By performing statistical tests and Monte Carlo simulation studies, we demonstrate that our proposed empirical model captures the joint distribution accurately, and also its time-varying behavior.

Secondly, we provide applications of our model that are of interest to e.g. an energy trading company managing a large share of wind turbines. We estimate the risk distribution and the price of correlation risk associated with a specific contract exposed to joint price and volumetric risk, i.e. a contract implying that an energy trading company offers wind power producers an insurance against price movements, by purchasing their fluctuating production at a predetermined fixed price. We show that the negative relation between prices and wind power production plays an important role both in relation to the pricing and the risk distribution of such contracts. We find that the price of correlation risk amounts to a significant percentage of the price of a regular fixed price agreement with no volumetric risk (a standard forward contract). Also, the risk distribution becomes left-skewed under the assumption of negative dependence compared to the case of independence. Lastly, we compare the out-of-sample performance of competing models, and show that time-varying copula models outperform the constant copula models.

This paper is organized as follows: Section 2 briefly introduces the notion of copula and the methodology used in building a joint model for electricity prices and wind power production. In Section 3, we apply the theory to data from the Danish power market. In Section 4, we present a simulation study and investigate how different wind scenarios affect the conditional distribution of spot electricity prices. Section 5 presents an application to pricing and risk management, and in Section 6 we conclude.

2 Modeling dependence with copula models

Formally, a d-dimensional copula is a distribution function $C(u_1, \ldots, u_d)$ defined on the unit cube $[0,1]^d$ with uniform margins. Since our application is a bivariate one, we shall consider the case where $d = 2$, however copula theory holds for the general multivariate case. The central result when working with copula models is Sklar's theorem, which shows how to decompose a

joint distribution function into its univariate marginal distribution functions and a copula.

In our application, we wish to condition on the information generated by past observations of our variables, denoted by $\mathcal{F}_{t-1}$. Thus, we shall consider an extension to Sklar's theorem proposed in Patton (2006), which holds for conditional joint distributions. The theorem states that if we let $F(\cdot \mid \mathcal{F}_{t-1})$ be the bivariate conditional distribution function of the random vector $\boldsymbol{Y}_t \equiv (Y_{1,t}, Y_{2,t})'$, with conditional marginal distribution functions $F_1(\cdot \mid \mathcal{F}_{t-1})$ and $F_2(\cdot \mid \mathcal{F}_{t-1})$, then there exists a two dimensional conditional copula $C(\cdot \mid \mathcal{F}_{t-1})$, such that

$$F((y_1, y_2) \mid \mathcal{F}_{t-1}) = C(F_1(y_1 \mid \mathcal{F}_{t-1}), F_2(y_2 \mid \mathcal{F}_{t-1}) \mid \mathcal{F}_{t-1}). \tag{1}$$

Furthermore, if the marginal distribution functions are continuous, the copula is unique. The converse also holds, such that given two conditional marginal distributions, we can use the conditional copula to link the variables to form a conditional joint distribution with the specified margins. It is especially this second part of the theorem that is useful here, since it allows us to isolate the description of the dependence structure from the marginal behavior of the individual variables.

Moreover, let us define the probability integral transform variables

$$U_{i,t} \equiv F_i(Y_{i,t} \mid \mathcal{F}_{t-1}), \quad \text{for } i = 1, 2, \tag{2}$$

and let $\boldsymbol{U}_t \equiv (U_{1,t}, U_{2,t})'$. Then $U_{i,t} \sim \text{Unif}(0, 1)$, and note furthermore that the conditional copula in Eq. (1) is simply the conditional distribution of $\boldsymbol{U}_t \mid \mathcal{F}_{t-1}$:

$$\boldsymbol{U}_t \mid \mathcal{F}_{t-1} \sim C(\cdot \mid \mathcal{F}_{t-1}). \tag{3}$$

In this paper, we consider different copulas from the elliptical and archimedean families, which are commonly used in the financial literature. For a detailed treatment of these copulas and their properties, we refer to the reference books by Joe (1997) and Nelsen (1999).

2.1 Marginal models

As a first step when working with copulas, we need to find proper marginal distribution models. Here, we restrict our attention to marginal models of the ARMA–GARCH type to model the conditional mean and the conditional variance of the individual variables.[3] For example, an ARMA(p, q)–

[3] A variety of other parametric specifications can be considered for the conditional mean, such as ARMAX models, long memory models, linear and nonlinear regression models, etc. The same holds for the conditional variance where, among others, different extensions to the ARCH model can be considered; see Bollerslev (2008) for a long list of such models.

GARCH$(1,1)$ model for the margins can be written as

$$Y_{i,t} = \sum_{j=1}^{p} \phi_{i,j} Y_{i,t-j} + \sum_{k=1}^{q} \theta_{i,k} \varepsilon_{i,t-k} + \varepsilon_{i,t}, \tag{4}$$

$$\varepsilon_{i,t} = \sigma_{i,t} \eta_{i,t}, \tag{5}$$

$$\sigma_{i,t}^2 = \omega_i + \alpha_i \varepsilon_{i,t-1}^2 + \beta_i \sigma_{i,t-1}^2, \tag{6}$$

for $i = 1, 2$, where ω_i, α_i, β_i follow the restrictions posed in e.g. Nelson and Cao (1992), and $\alpha_i + \beta_i < 1$. Furthermore,

$$\eta_{i,t} \,|\, \mathcal{F}_{t-1}^{(i)} \sim F_i(0,1), \quad \text{for } i = 1, 2 \text{ and all } t. \tag{7}$$

For the marginal distributions we consider the case where F_i does not vary with time and has a parametric form. Also, we relax the normality assumption, allowing for more general distributions. The ARMA–GARCH models function as filters that produce innovation processes $\eta_{1,t}$ and $\eta_{2,t}$ that are serially independent; it is the conditional distributions of $\eta_{1,t}$ and $\eta_{2,t}$ that are then coupled using the conditional copula.

One note of caution has to be made regarding the conditioning set $\mathcal{F}_{t-1}$ emphasizing that this set is generated by $(\boldsymbol{Y}_{t-1}, \boldsymbol{Y}_{t-2}, \dots)$. In our specification for the marginal models however, we do not condition on $\mathcal{F}_{t-1}$, but only a subset $\mathcal{F}_{t-1}^{(i)} \subset \mathcal{F}_{t-1}$. When using such models, the copula is, according to Fermanian and Wegkamp (2012), a true copula if and only if

$$Y_{i,t} \,|\, \mathcal{F}_{t-1} \overset{d}{=} Y_{i,t} \,|\, \mathcal{F}_{t-1}^{(i)}, \tag{8}$$

for $i = 1, 2$ and all t. If the equality in Eq. (8) is not satisfied, then the joint conditional distribution of $\boldsymbol{Y}_t \,|\, \mathcal{F}_{t-1}$ does not have the specified conditional marginal distributions. To study if the equality in Eq. (8) holds, we test for cross-equation effects by including lags of one variable in the conditional mean equation of the other variable and vice versa, and perform a standard Wald test for the joint significance of the added explanatory variables, as proposed in Patton (2013).

2.2 Estimation procedure for the joint model

To estimate the joint model, we perform maximum likelihood estimation. The joint conditional density is obtained by differentiating the joint conditional distribution function in Eq. (1). Thus, the log-likelihood function takes the

form

$$\log \mathcal{L} = \sum_{t=1}^{T} \log f((y_{1,t}, y_{2,t}) \mid \mathcal{F}_{t-1}; \Theta) = \sum_{t=1}^{T} \log f_1(y_{1,t} \mid \mathcal{F}_{t-1}; \Theta_1) + \sum_{t=1}^{T} \log f_2(y_{2,t} \mid \mathcal{F}_{t-1}; \Theta_2) + \sum_{t=1}^{T} \log c((u_{1,t}, u_{2,t}) \mid \mathcal{F}_{t-1}; \gamma), \tag{9}$$

where f_1 and f_2 are the conditional marginal densities, c is the conditional copula density defined as

$$c((u_{1,t}, u_{2,t}) \mid \mathcal{F}_{t-1}) = \frac{\partial^2}{\partial u_1 \partial u_2} C((u_{1,t}, u_{2,t}) \mid \mathcal{F}_{t-1}), \tag{10}$$

and

$$u_{i,t} = F_i(y_{i,t} \mid \mathcal{F}_{t-1}; \Theta_i), \quad \text{for } i = 1, 2. \tag{11}$$

In Eq. (9), Θ denotes the set of parameters for the entire model, and Θ_1, Θ_2 and γ denote the parameters for the two marginal models and the copula, respectively, and have no common elements. For simplicity, we assume that the copula is completely described by one single parameter γ. We perform multistage maximum likelihood estimation, where we consider the two marginal models and the copula model separately. For details on the validity of this procedure, consult Patton (2013).

2.3 Time-varying copula models

Since the dependency between electricity prices and wind power production might change through time, extending copula models to allow for time-varying dependence is relevant. Before specifying a parametric model for the copula dependence parameter, it is useful to investigate what type of time variation (if any) we can detect in the data. Here, we employ two tests proposed in Patton (2013): One that tests for the presence of a break in the rank correlation by performing the classical "sup" test, and another that tests for the presence of autocorrelation in a measure of dependence. For a comprehensive description of the two tests the reader is referred to Patton (2013).

The Generalized Autoregressive Score model

To model time-varying dependence, we employ the *Generalized Autoregressive Score* (GAS) model of Creal et al. (2013). In order to ease the presentation, we consider the case where the copula has one dependence parameter. For

the GAS(1,1) model, a possible updating equation for the transformed copula dependence parameter g_{t+1} is:

$$g_{t+1} = \omega + \alpha g_t + \beta I_t^{-\frac{1}{2}} s_t, \tag{12}$$

where

$$\begin{aligned} g_t &= h(\gamma_t), \\ s_t &= \frac{\partial}{\partial \gamma} \log c((u_{1,t}, u_{2,t}); \gamma_t), \\ I_t &= \mathbb{E}_{t-1}\left[s_t^2\right]. \end{aligned}$$

In Eq. (12), s_t denotes the score of the copula log-likelihood and I_t is the Fisher information. Moreover, γ_t denotes the time-varying copula dependence parameter, which is usually constrained to lie in a particular range; see Table 9 in B for details regarding the range of different copula dependence parameters. For estimation purposes, we apply a transformation $h(\cdot)$ to γ_t, to obtain g_t which takes values on the entire real axis. We note that the updating mechanism given in Eq. (12) is one of many possible specifications: The GAS model can be extended to include e.g. more lags or exogenous variables. Moreover, the scaling quantity $I_t^{-1/2}$ is simply one convenient choice. GAS models can be generalized to allow for asymmetries or long memory, and to include regime-switching, however such extensions are not considered in the present work.

The parameter estimates from the GAS model can be obtained by maximum likelihood estimation, as proposed by Creal et al. (2013). The only challenge can be finding a closed-form expression for the Fisher information, and thus deriving the updating mechanism in Eq. (12). To overcome this issue, the Fisher information is evaluated numerically for most copula specifications by performing the following steps:

1. Given a copula specification, construct a grid of values for the dependence parameter, $[\gamma^{(1)} < \gamma^{(2)} < \cdots < \gamma^{(n)}]$.
2. For each dependence parameter in the grid,
   (a) perform a large number of simulations from the chosen copula model,
   (b) evaluate the score function at each simulation,
   (c) compute the Fisher information, by taking the mean over the evaluated scores squared.
3. Finally, use linear interpolation to get the Fisher information at intermediate points.

2.4 Quantile dependence

As a preliminary study before specifying copula models, one can examine the dependence in the data by considering quantile dependence. For the case of negatively dependent variables, the quantile dependence is defined as:

$$\lambda^q = \begin{cases} \mathbb{P}(U_{1,t} \leq q \,|\, U_{2,t} \geq 1-q), & 0 < q \leq 1/2, \\ \mathbb{P}(U_{1,t} > q \,|\, U_{2,t} < 1-q), & 1/2 < q < 1. \end{cases} \tag{13}$$

By computing quantile dependence coefficients at different quantiles q, we obtain a richer description of the dependence structure. This can help narrow down the set of possible parametric copulas to a collection of models that are able to capture some of the characteristics we observe in the data. To obtain standard errors for the quantile dependence coefficients, we use boostrapping; specifically, we follow the procedure proposed in Patton (2013), which is based on the stationary block-bootstrap of Politis and Romano (1994), where the optimal block-length is chosen according to Politis and White (2004) and Politis et al. (2009).[4]

2.5 Selection of copula models

To test for whether or not a copula is well specified, we perform two widely used goodness-of-fit tests (GoF): The Kolmogorov-Smirnov (KS) and the Cramer von-Mises (CvM) tests. Under the null that the conditional copula is well specified, we should find that the empirical copula provides a good nonparametric estimate of the null conditional copula. Suppose we have the random sample $\{\boldsymbol{u}_t\} = \{(u_{1,t}, u_{2,t})\}_{t=1}^T$ from $\boldsymbol{U}_t$. Then the test statistics can be written as

$$KS^{(C)} = \max_t \left| C(\boldsymbol{u}_t; \hat{\gamma}) - \hat{C}(\boldsymbol{u}_t) \right|, \tag{14}$$

$$CvM^{(C)} = \sum_{t=1}^{T} \left\{ C(\boldsymbol{u}_t; \hat{\gamma}) - \hat{C}(\boldsymbol{u}_t) \right\}^2, \tag{15}$$

where $C(\boldsymbol{u}_t; \hat{\gamma})$ is an estimator of the null conditional copula. Moreover, $\hat{C}$ denotes the empirical copula defined as

$$\hat{C}(\boldsymbol{z}) \equiv \frac{1}{T+1} \sum_{t=1}^{T} \mathbf{1}\{u_{1,t} \leq z_1, u_{2,t} \leq z_2\}, \tag{16}$$

where $\mathbf{1}$ denotes the indicator function and $\boldsymbol{z} = (z_1, z_2) \in [0,1]^2$. The KS and CvM tests described above work solely for the testing of constant copula

[4]The same bootstrapping procedure can be used to perform inference on other measures of dependence, e.g. linear correlation, rank correlation.

models. A slight modification will however allow for the additional testing of time-varying copulas: The KS and CvM tests based on the Rosenblatt transform. In our case, the transformation is simply

$$V_{1,t} = U_{1,t} \tag{17}$$
$$V_{2,t} = C_{2|1,t}(U_{2,t}|U_{1,t}; \hat{\gamma}_t), \tag{18}$$

where $C_{2|1,t}$ denotes the conditional copula of the random variable $U_{2,t} \,|\, U_{1,t}$. Applying the Rosenblatt transform to the data will yield *iid* and Unif$(0,1)$ variables, and hence we can compare the empirical copula of a random sample $\{\boldsymbol{v}_t\} = \{(v_{1,t}, v_{2,t})\}_{t=1}^{T}$ from $\boldsymbol{V}_t$, against the independence copula, defined as

$$C^{indep}(\boldsymbol{v}_t; \hat{\gamma}_t) \equiv \prod_{i=1}^{2} v_{i,t}. \tag{19}$$

A simulation-based approach is used to obtain p-values for the GoF tests described above, since the test statistics in Eqs. (14) and (15) depend on estimated parameters. This approach is described in detail in Berg (2009), Genest et al. (2009) and Patton (2013), and will not be elaborated on here.

Another very important issue when dealing with copulas is choosing the best copula model among competing models. Here, we consider pairwise comparisons, where we follow Rivers and Vuong (2002) for most in-sample (IS) model comparisons and Diks et al. (2010) for out-of-sample (OOS) model comparisons. The IS comparison test can be performed when the models are non-nested; for the case where the models are nested, a likelihood ratio test can usually be used. The OOS model comparison test works for both nested and non-nested models. Also, both tests can be applied regardless of whether the copula is constant or time-varying.

For the IS case, the idea is to compare two models using their joint log-likelihood, and test the null

$$H_0 : \mathbb{E}\left[\boldsymbol{L}^{(1)} - \boldsymbol{L}^{(2)}\right] = 0, \tag{20}$$

against

$$H_1 : \mathbb{E}\left[\boldsymbol{L}^{(1)} - \boldsymbol{L}^{(2)}\right] > 0 \quad \text{and} \quad H_2 : \mathbb{E}\left[\boldsymbol{L}^{(1)} - \boldsymbol{L}^{(2)}\right] < 0, \tag{21}$$

where the superscripts (1) and (2) denote two competing models. The case of comparing joint log-likelihoods reduces in our case to comparing copula log-likelihoods, c.f. Eq. (9), since we use the same marginal distribution models. Hence, $\boldsymbol{L}^{(i)} = \log c^{(i)}(\boldsymbol{u}; \gamma^{(i)})$ or $\boldsymbol{L}^{(i)} = \log c^{(i)}(\boldsymbol{u}; \gamma_t^{(i)})$, $i = 1, 2$, depending on whether the copula is constant or time-varying. Rivers and Vuong (2002)

show that under the null,

$$\frac{\sqrt{T}\left(\bar{L}^{(1)} - \bar{L}^{(2)}\right)}{\sqrt{\hat{\sigma}^2}} \xrightarrow{d} \mathrm{N}(0,1) \tag{22}$$

where

$$\bar{L}^{(i)} = \frac{1}{T}\sum_{t=1}^{T} \log c^{(i)}\left(\hat{\boldsymbol{u}}_t; \hat{\gamma}_t^{(i)}\right), \quad \text{for } i = 1,2. \tag{23}$$

As an estimator for the asymptotic variance of $\sqrt{T}\left(\bar{L}^{(1)} - \bar{L}^{(2)}\right)$ we use the Newey-West heteroskedasticity and autocovariance consistent (HAC) estimator.

For OOS comparisons, we consider a fixed estimation window, where the model is estimated using the data from $[1, T]$. We then evaluate the conditional predictive ability of two competing copulas on the OOS period, i.e. on R observations, where $R = T^* - T$, $T^* > T$. The test for comparing the predictive ability of competing copula models conditional on the estimated parameters proposed by Diks et al. (2010) is in fact a special case of the more general framework presented in Giacomini and White (2006). The null hypothesis for the OOS case is the same as for the IS case, and a test statistic based on the difference between the sample averages of the copula log-likelihoods can again be used, and is shown by Giacomini and White (2006) to be asymptotically $\mathrm{N}(0,1)$ under the null. As an estimator for the asymptotic variance, we use the HAC estimator.

3 Empirical results

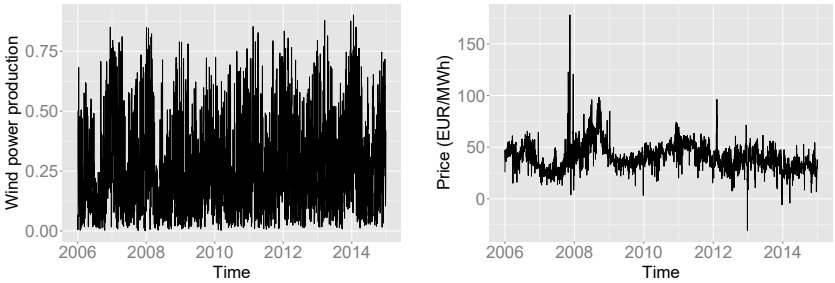
A joint model for electricity prices and wind power production is interesting to consider in an area with a high penetration ratio of wind power in the grid. Here, we analyze data from Denmark, which has long been among the top wind power producing countries. According to *Energinet.dk*, the Danish Transmission System Operator, more than a third of the Danish power consumption was covered by wind power in 2013, and in December that year, 57.4% of the consumption came from wind turbines. In 2014, wind turbines produced on average what corresponds to over 39% of the Danish power consumption. Also, in January 2014, 61.7% of the consumption was covered by wind power.

Specifically, we base our analysis on data from one of the two Danish price areas, DK1 (Western Denmark), and a sample period that spans from 1 January 2006 to 31 December 2014. The first time series, Fig. 1(a), consists of total daily wind power production in DK1 relative to the total installed

capacity, and is obtained by performing the normalization

$$\frac{\text{Total daily wind power production (MWh)}}{\text{Installed capacity (MW)} \cdot H} \tag{24}$$

for each day in the sample, where H denotes the total number of hours in the day. We note that we work in UTC time, so $H = 24$ always. The second time series, Fig. 1(b), represents the daily average of spot electricity prices.[5,6]



(a) Daily wind power production measured relative to the total installed capacity.

(b) Daily spot electricity prices

Fig. 1: Historical daily observations for the DK1 price area in the period 1 January 2006 to 31 December 2014.

Before proceeding to the estimation of a joint model for prices and wind power production, two comments are in order. First, since the production series is bounded, with a lower bound at 0 and an upper bound at 1, we perform a logistic transformation in order to obtain data that can take values on the entire real line. Second, we split our data into an in-sample (IS) period spanning from 1 January 2006 to 31 December 2012, and an out-of-sample (OOS) spanning from 1 January 2013 to 31 December 2014. Estimation of marginal models and copulas is performed on the IS data.

3.1 Marginal specifications for spot electricity prices and wind power production

Prior to modeling the dependence structure of electricity prices and wind power production, we filter out the stylized facts affecting the marginal behavior of the individual variables. As a first step, we demean and correct

[5]The data is publicly available on *Energinet.dk* and on the web page of Nord Pool's Elspot market, *nordpoolspot.com*. Elspot is a day-ahead physical delivery market for electricity currently operating in the Nordic and Baltic region.

[6]We note that one observation has been truncated in the price data, corresponding to the date 7 June 13, since this is assessed to be an outlier. On this date, the hourly price reached Nord Pool's cap price due to a combination of low wind, reduced import possibilities caused by planned maintenance on both transmission cables and central power stations.

for deterministic seasonality by performing a regression on a constant and dummy variables. Specifically, we have used the dummy variable *month-of-year* to correct the (transformed) wind power production series for seasonality. For the price series both *day-of-week* and *month-of-year* dummy variables were used as regressors. To model the conditional mean and variance of the variables, we consider ARMA–GARCH models with different specifications for the error distribution. We consider ARMA models up to order (7,7), and GARCH models up to order (2,2). Based on the Bayesian Information Criterion, we find that the optimal model for the wind power production series is an ARMA(1,3)–GARCH(1,1), and use a skewed generalized error distribution for the standardized residuals. For the day-ahead electricity prices, we find the optimal model to be an ARMA(3,1)–GARCH(1,1), and use a skewed t distribution for the standardized residuals. Table 1 summarizes the estimation results, and Fig. 10 in Appendix A displays the autocorrelation functions, histograms and quantile plots for the standardized residuals resulting from the fitted models. A visual inspection of Fig. 10 shows that almost no autocorrelation is left in the standardized residuals. The specified distributions provide a reasonable fit, however we observe some deviations in the tails of both distributions.

	Daily wind power production ARMA(1,3) – GARCH(1,1)	Daily spot electricity prices ARMA(3,1) – GARCH(1,1)
	Conditional mean	
$\hat{\phi}_1$	0.8725 (0.0510)	1.4579 (0.0065)
$\hat{\phi}_2$	-	-0.5525 (0.0176)
$\hat{\phi}_3$	-	0.0897 (0.0261)
$\hat{\theta}_1$	-0.3578 (0.0550)	-0.8365 (0.0128)
$\hat{\theta}_2$	-0.2733 (0.0363)	-
$\hat{\theta}_3$	-0.0610 (0.0264)	-
	Conditional variance	
$\hat{\omega}$	0.0803 (0.1269)	2.4433 (0.7388)
$\hat{\alpha}$	0.0251 (0.0199)	0.1657 (0.0312)
$\hat{\beta}$	0.9022 (0.1333)	0.7832 (0.0410)
	Skewed general error dist./Skewed t dist.	
Shape $\hat{\nu}$	2.1348 (0.0967)	4.9967 (0.4711)
Skewness $\hat{\xi}$	0.8024 (0.0269)	0.9583 (0.0222)
	Goodness-of-fit tests	
KS (p-val.)	0.6293	0.7097
CvM (p-val.)	0.5882	0.5996

Table 1: The first panels display parameter estimates together with their std. errors in parenthesis. The last panel displays the results of GoF tests.

We complement these findings with GoF tests, where we consider the KS and CvM tests. The resulting p-values are given in Table 1 and indicate that

there is not sufficient evidence as to reject the null that the distributional assumptions are well-specified.[7] We note that finding suitable marginal models is of great concern when working with copula models, since the copula takes as input *iid* Unif$(0,1)$ variables that result from applying the probability integral transform to the standardized residuals. A violation of the assumptions will thus automatically lead to a misspecified copula model.

Because we condition with different information sets when specifying the marginal models, we need to investigate whether or not lagged values of wind power production help explain electricity prices and vice versa. To do this, we consider the specified models for the conditional mean with added explanatory variables consisting of seven lagged values of the "other" series, and test for the significance of cross-sectional effects by performing a Wald test. For the wind power production, we consider an ARMAX(1,3,7) model, and for the electricity prices, we consider an ARMAX(3,1,7) model. The tests yield a p-value of 0.25 for the wind power production model, and 0.09 for the electricity price model, thus suggesting no cross-equation effects at a 5% significance level.[8]

3.2 Symmetric vs. asymmetric dependence

Having decided upon the marginal models for price and wind power production, the remaining of this section focuses on the modeling of the dependence structure. First, we apply the probability integral transform to the standardized residuals resulting from the marginal models to obtain approximately uniformly distributed variables. To perform this transformation, we use the estimated parametric models for the distribution functions F, i.e. the estimated skewed generalized error distribution and skewed t distribution, see Table 1. We obtain

$$\hat{U}_{W,t} = F_{skew\ ged}(\hat{\eta}_{W,t}, \hat{\nu}_W, \hat{\xi}_W) \tag{25}$$

$$\hat{U}_{S,t} = F_{skew\ t}(\hat{\eta}_{S,t}, \hat{\nu}_S, \hat{\xi}_S), \tag{26}$$

where $\hat{U}_{W,t}$ and $\hat{U}_{S,t}$ denote the resulting uniforms corresponding to the wind power production time series and the spot price time series, respectively. Standardized residuals are denoted by $\hat{\eta}$, and estimated distribution parameters are denoted by $\hat{\xi}$ (skew parameter) and $\hat{\nu}$ (shape parameter).

As an introductory investigation of the dependence structure, we compute some measures of dependence for $\hat{\boldsymbol{U}}_W$ and $\hat{\boldsymbol{U}}_S$. Table 2 displays the

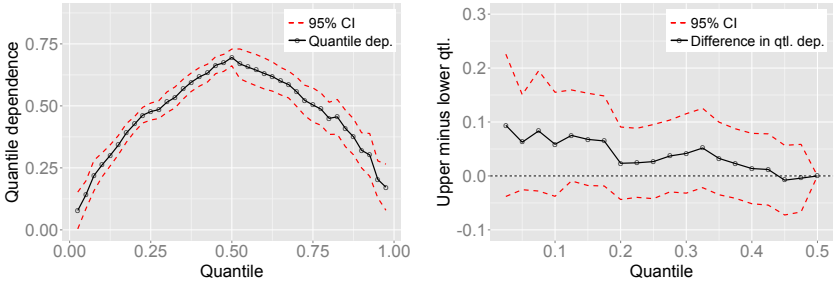
[7] We perform simulation-based GoF tests, that take the parameter estimation errors from the ARMA–GARCH models into account. Specifically, we test for whether or not the probability integral transforms implied by the estimated conditional densities are *iid* Unif$(0,1)$. The p-values for the tests are based on 999 simulations.

[8] We have tried testing for cross-sectional effects with different other specifications, and none of the results indicate cross-equation effects at a 5% significance level.

estimated coefficients for Spearman's ρ, Kendall's τ and linear correlation, implying (not surprisingly) that prices and wind power production are negatively correlated. Based on Eq. (13) we also compute quantile dependence measures, and the results, displayed in Fig. 2, show evidence for a symmetric dependence structure. When considering the farther right and left portions of Fig. 2(a), the results reveal a slightly larger probability of observing low prices given that the production is high than the opposite. However, according to Fig. 2(b), this difference is not statistically significant.

	Spearman's ρ	Kendall's τ	Linear correlation
Estimate	-0.5024	-0.3478	-0.5030
95% CI	$(-0.5716, -0.4332)$	$(-0.3987, -0.2969)$	$(-0.5714, -0.4347)$

Table 2: Estimated dependence measures with 95% confidence intervals based on the block-bootstrap procedure described in Section 2.4 and $M = 999$ bootstrap samples.



(a) Quantile dependence

(b) Difference in quantile dependence

Fig. 2: Fig. 2(a) displays estimated quantile dependence for quantile $q \in [0.025, 0.975]$ and a size step of 0.025, along with a 95% confidence interval based on the block-bootstrap procedure described in Section 2.4 and $M = 999$ bootstrap samples. The y-axis provides the probability of $\hat{U}_W$ lying below (above) its q quantile given that $\hat{U}_S$ lies above (below) its $1 - q$ quantile for $q \leq 1/2$ ($q > 1/2$). Fig. 2(b) shows the difference in corresponding left and right quantile dependence illustrated in Fig. 2(a) with a corresponding 95% confidence interval.

3.3 Constant copula models

Although we anticipate time-variation in the dependence structure, we consider six constant copula models, to have as benchmarks for later comparisons. A brief overview of these copula models is provided in Appendix B. The estimation results for the proposed constant copulas are given in Table 3, together with GoF test results. Among the constant copulas we consider, it is only the Gaussian and Student t that allow for negative dependence. To deal

with this issue, we have performed suitable rotations of our data when estimating the Clayton, Gumbel, Joe-Frank and Symmetrized Joe-Clayton (SJC) copulas. Furthermore, the Gaussian and the Student t copulas are symmetric, the Clayton and Gumbel are asymmetric, and the combinations Joe-Frank and SJC allow for more flexible dependence structures and nest the case of symmetric dependence.

Copula		Parameter estimates	(s.e.)	$\log \mathcal{L}$	GoF tests p-val.			
					Kolmogorov-Smirnov		Cramer-von Mises	
					$KS^{(C)}$	$KS^{(R)}$	$CvM^{(C)}$	$CvM^{(R)}$
Gaussian	$\hat{\rho}$	-0.4923	(0.0134)	355.07	0.4545	0.6096	0.6747	0.4044
Student t	$\hat{\rho}$	-0.4967	(0.0147)	357.15	0.4134	0.7427	0.7227	0.5435
	$\hat{\nu}^{-1}$	0.0318	(0.0170)					
Clayton	$\hat{\theta}$	0.7024	(0.0339)	288.94	*0.0020*	*0.0060*	*0.0000*	*0.0020*
Gumbel	$\hat{\theta}$	1.4455	(0.0220)	331.14	*0.0450*	0.0681	*0.0390*	0.0661
Joe-Frank	$\hat{\theta}$	8.8682	(4.4760)	368.77	0.7778	0.9359	0.7017	0.8278
	$\hat{\delta}$	0.3431	(0.1125)					
SJC	$\hat{\tau}^{U}$	0.3355	(0.0368)	320.18	*0.0070*	*0.0150*	*0.0230*	*0.0290*
	$\hat{\tau}^{L}$	0.2012	(0.0408)					

Table 3: Estimation and GoF test results for constant copula models. The p-values less than 0.05 are given in italics, highlighting that the dependence structure is not well-represented by the proposed copula model. The superscript (C) refers to the tests performed on the empirical copula of the standardized residuals and the superscript (R) refers to the tests performed on the empirical copula of the Rosenblatt transforms. The GoF tests are simulation based (999 bootstraps) and take parameter estimation errors into account.

The GoF results in Table 3 support our earlier findings in Section 3.2. The Gaussian and Student t copulas are, according to all tests, a good specification. Clayton is rejected by all tests, while Gumbel is only partly rejected. For the combination copulas, the test results are more surprising: The Joe-Frank specification is accepted by all tests, while the SJC specification is rejected by all tests. We attempt to understand these results by plotting the quantile dependence we observe in our data together with the quantile dependence implied by some of the fitted copulas in Fig. 3.

We observe that the quantile dependence implied by the Gaussian copula provides a reasonable fit to our data. So does the Joe-Frank copula, by providing a fit that generates almost no asymmetry. The Gumbel copula on the other hand is too asymmetric, producing large deviations as we approach one of the tails. Lastly, the SJC, although implying less asymmetry than Gumbel, assigns too much probability to extreme events compared to what we observe in the data, and thus produces large deviations as we approach both tails.[9]

[9] We have omitted the quantile dependence implied by the Student t and Clayton copulas in Fig. 3 for clarity reasons. The Student t copula implies quantile dependence coefficients that are almost indistinguishable from the Gaussian ones, which is due to the very high value we estimate for the degree of freedom of this copula. The Clayton copula implies even more asymmetry than Gumbel in the far right side of the quantile plot.

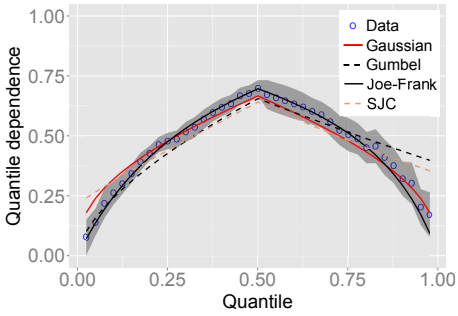


Fig. 3: Quantile dependence implied by some of the fitted constant copula models in Table 3.

3.4 Time-varying copula models

To confirm our suspicion that the dependence of spot electricity prices and wind power production is time-varying, we perform the two tests briefly described in Section 2.3. The results are given in Table 4, showing no evidence of a one-time break in the dependence structure, but strong evidence for the presence of autocorrelation in the rank correlations.

	One-time break	Time-varying dep. of autoreg. type		
		AR(1)	AR(5)	AR(7)
p-value	0.7898	0.0020	0.0110	0.0000

Table 4: Test results for time-varying dependence. To test for the presence of a one-time break in the rank correlation we use the "sup" test, and test the null of no one-time break. To test for the presence of time-varying dependence of autoregressive type we consider the regression $\hat{U}_{W,t}\hat{U}_{S,t} = \mu + \sum_{i=1}^{p} \phi_i \hat{U}_{W,t-i}\hat{U}_{S,t-i} + \varepsilon_t$, for $p = 1, 5, 7$; the null of a constant copula cannot be rejected if we find that $\phi_i = 0$, for $i = 1, \ldots, p$. For all tests, p-values are obtained by bootstrap testing (based on 999 bootstraps, where bootstrap samples are obtained by randomly drawing rows, with replacement, from $(\hat{\boldsymbol{u}}_W, \hat{\boldsymbol{u}}_S)'$).

In light of these findings, we consider three copula models where the transformed dependence parameter denoted by g evolves according to a GAS(1,1) model, see Eq. (12). The transformations applied to the copula dependence parameters, estimation and GoF test results are all displayed in Table 5. For the Gaussian copula, a closed form expression for the Fisher information can be derived (see e.g. Schepsmeier and Stöber (2014)). For the Gumbel and the Joe-Frank copulas, the Fisher information is computed numerically by performing the steps in Section 2.3. The Joe-Frank copula has two dependence parameters, and we consider the case where one parameter evolves according to the GAS specification, while the other is kept constant. It should however be mentioned that letting both parameters vary through time provides very little improvement. Regarding the parameter estimates,

α is high in all models, implying a very persistent time-varying correlation process. Also, the intercept parameter ω is not significant in any model. As far as the GoF test results are concerned, the Joe-Frank and Gaussian GAS models are accepted at a 5% level, while the Gumbel GAS model is only partially accepted.

	Transf. h	Parameter estimates (s.e.)				log $\mathcal{L}$	GoF tests p-val.	
		$\hat{\omega}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\delta}$		$KS^{(R)}$	$CvM^{(R)}$
Gaussian	$\log\left(\frac{1+\rho}{1-\rho}\right)$	-0.0196 (0.0268)	0.9820 (0.0246)	0.0390 (0.0159)	-	373.90	0.5354	0.4646
Gumbel	$\log(\theta-1)$	-0.0272 (0.1527)	0.9672 (0.1849)	0.0420 (0.0334)	-	342.90	*0.0404*	0.0707
Joe-Frank	$\log(\theta-1)$	0.0314 (0.0509)	0.9860 (0.0208)	0.0403 (0.0127)	0.2944 (0.0894)	388.77	0.7959	0.8571

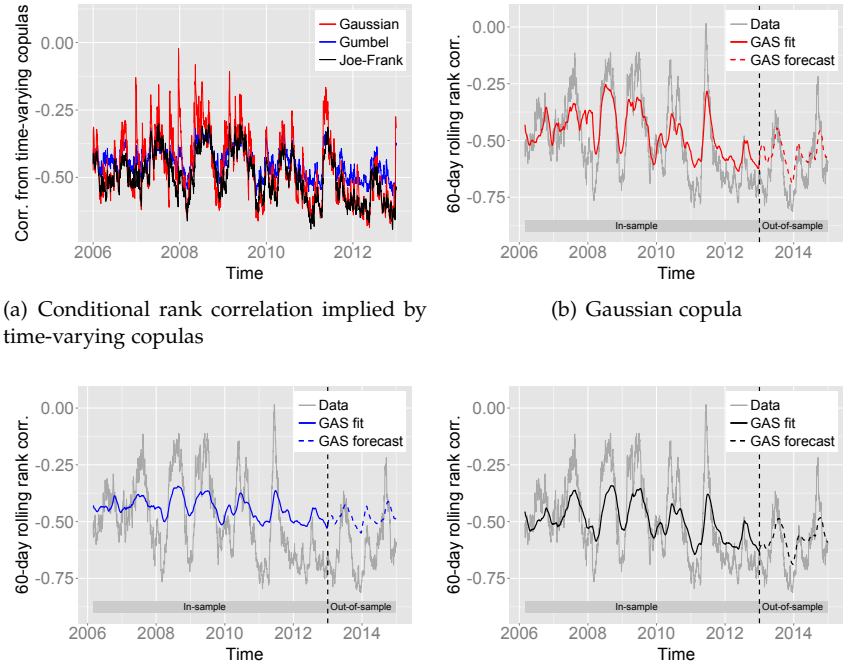
Table 5: Estimation and GoF results for time-varying copulas. Due to the high computational time, the GoF tests and standard errors are based on 99 bootstraps. The superscript (R) indicates that the GoF tests are based on the Rosenblatt transform.

To visualize and compare the fits of the proposed GAS models, we plot the conditional rank correlation implied by the fitted time-varying copula models in Fig. 4(a). The numbers are obtained by mapping the copula parameter(s) to a rank correlation coefficient[10]. In Figs. 4(b)-(d) we plot actual 60-day rolling rank correlations of the data $(\hat{\boldsymbol{U}}_W, \hat{\boldsymbol{U}}_S)'$ together with the in-sample fit of the proposed time-varying models. To perform the same comparison for the out-of-sample period, we obtain the approx. uniforms $(\hat{\boldsymbol{U}}_W^{OOS}, \hat{\boldsymbol{U}}_S^{OOS})'$ by first applying the estimated function for removing seasonality and then the ARMA–GARCH filters, without re-estimating any parameters, to the out-of-sample wind power production data and the out-of-sample spot electricity price data. The 60-day rolling rank correlations of $(\hat{\boldsymbol{U}}_W^{OOS}, \hat{\boldsymbol{U}}_S^{OOS})'$ are then computed and compared to one-step-ahead forecasts from the time-varying copulas. Due to the elevated computational cost of using a rolling estimation window to produce forecasts, we restrict ourselves to considering a fixed estimation window corresponding to the in-sample period, but enlarge the conditioning set as information becomes available.

One first and surprising remark regarding Fig. 4 is related to the data itself and implicitly the fits produced by the GAS models, namely that the correlation is generally stronger during winter than during summer. There

[10]Specifically, we follow the procedure described in Patton (2013): (1) construct a grid of copula parameters, (2) perform 100,000 simulations from the copula model at each point in the grid, (3) compute the rank correlation of the simulations, and finally (4) use linear interpolation to obtain the correlation at intermediate points. We also mention that the functions mapping the copula parameters to rank correlation are smooth.

are many factors that can help explain this finding since price formation is a complex process that is not only influenced by supply and demand (which in turn have strong seasonal components), but also transmission capacity. To provide a few facts that can help explain our findings, we mention that the wind power production relative to the consumption in the DK1 price area has been higher for winter periods than summer periods, during the sample period we consider in this paper. Also, we can expect that situations with little wind during summer do not always push the prices upwards. This is (aside from consumption being lower during summer) due to the fact that DK1 is well connected with cables to Norway, Sweden and Germany, which are all heavy producers of renewable energy, and hence electricity could be imported at a lower price compared to the cost of having to turn on the more costly power stations in DK1.



(a) Conditional rank correlation implied by time-varying copulas

(b) Gaussian copula

(c) Gumbel copula

(d) Joe-Frank copula

Fig. 4: Fits and forecasts produced with the three time-varying copulas from Table 5.

Considering now the fits implied by the proposed time-varying copulas, Fig. 4 reveals that the Gaussian GAS implies most variation in the correlation and is able to capture periods with weaker dependence the best. The Gumbel GAS specification is the one that least captures the variation that we observe in the data. The Joe-Frank GAS specification is superior at reaching

the stronger correlations, but does not produce correlations that are weaker than around -0.3. The plots clearly help establish that the Gumbel GAS specification is the inferior choice. However, it is difficult to choose the better copula when considering the Gaussian GAS against the Joe-Frank GAS.

From fitting not only time-varying copulas but also constant ones, we have so far obtained many different models that are actually well-specified according to the GoF tests. To help choose among all the considered copulas, we perform the pairwise comparison tests described in Section 2.5.[11] The results are summarized in Table 6. We find that The Joe-Frank GAS specification outperforms all other specifications in-sample, however its superiority over the Gaussian GAS specification is not statistically significant. When considering the out-of-sample results, the situation reverses, with the Gaussian GAS specification performing the best, but not significantly better than the Joe-Frank GAS. Since the Gaussian GAS is the smaller model, we will choose this specification as our preferred one, and continue our investigations using this model to describe the dependence between wind power production and spot electricity prices.

4 A simulation study

Performing simulations from a copula model is straightforward. The basic steps are (1) at time t, generate the pair $(U_{W,t}, U_{S,t})$ from the Gaussian copula with dependence parameter ρ_t, (2) perform the inverse of the transformations given in Eqs. (25) and (26) to obtain standardized residuals $(\eta_{W,t}, \eta_{S,t})$, (3) insert the standardized residuals in the marginal models from before (see Table 1) to obtain a deseasonalized pair $(\tilde{Y}_{S,t}, \tilde{Y}_{W,t})$, (4) use the estimated seasonal function to obtain a pair $(Y_{S,t}, Y_{W,t})$ of spot electricity price and wind power production, (5) compute ρ_{t+1} using the Gaussian GAS update equation and (6) repeat steps (1)–(5). Using this procedure one day at a time, we can construct spot electricity price series and wind power production series; and by repeating the process many times, an empirical distribution is produced. Such an empirical distribution is shown in Fig. 5.

Fig. 5 illustrates the simulated conditional joint distribution for December 2013 obtained by simulating 10,000 random paths for a one-month horizon. Note that although we have chosen a Gaussian copula model for the dependence structure, the marginal distributions were chosen to be a skewed generalized error distribution and a skewed t distribution for the wind power production and spot electricity prices, respectively. Therefore, the resulting joint distribution is not bivariate normal; as illustrated in Fig. 5, the simulated distribution exhibits asymmetry and heavy tails.

[11] The in-sample pairwise comparison between the Gaussian and the Student t copula is based on a simple t-test, since these models are nested.

	In-sample model comparisons								
	Gaussian	Student t	Clayton	Gumbel	Joe-Frank	SJC	Gaussian^{GAS}	Gumbel^{GAS}	Joe-Frank^{GAS}
Gaussian									
Student t	1.88^{*}								
Clayton	-4.61^{***}	-5.04^{***}							
Gumbel	-2.20^{**}	-2.71^{***}	7.07^{***}						
Joe-Frank	1.40	1.28	5.27^{***}	3.10^{***}					
SJC	-3.91^{***}	-4.91^{***}	3.32^{***}	-1.74^{*}	-3.54^{***}				
Gaussian^{GAS}	2.94^{***}	2.54^{***}	5.71^{***}	3.63^{***}	0.44	5.30^{***}			
Gumbel^{GAS}	-1.12	-1.42	8.08^{***}	2.91^{***}	-2.00^{**}	3.07^{***}	-2.91^{***}		
Joe-Frank^{GAS}	3.14^{***}	3.12^{***}	6.37^{***}	4.44^{***}	3.20^{***}	4.88^{***}	1.57	3.66^{***}	
log $\mathcal{L}$	355.06	357.13	288.93	331.14	368.77	320.17	373.90	342.90	388.77
	Out-of-sample model comparisons								
	Gaussian	Student t	Clayton	Gumbel	Joe-Frank	SJC	Gaussian^{GAS}	Gumbel^{GAS}	Joe-Frank^{GAS}
Gaussian									
Student t	1.41								
Clayton	-4.36^{***}	-4.95^{***}							
Gumbel	-1.58	-2.04^{**}	8.18^{***}						
Joe-Frank	0.16	-0.22	3.64^{***}	1.32					
SJC	-2.67^{***}	-3.43^{***}	5.19^{***}	-1.99^{**}	-1.99^{**}				
Gaussian^{GAS}	2.85^{***}	2.76^{***}	5.61^{***}	3.54^{***}	2.04^{**}	4.56^{***}			
Gumbel^{GAS}	-0.67	-0.96	6.53^{***}	2.22^{**}	-0.63	2.53^{***}	-2.75^{***}		
Joe-Frank^{GAS}	1.76^{*}	1.58	4.21^{***}	2.38^{***}	2.46^{***}	2.90^{***}	-0.29	1.82^{*}	
log $\mathcal{L}$	160.10	161.48	129.55	150.52	160.67	146.11	171.96	155.25	170.56

Table 6: t-statistics from in-sample and out-of-sample pairwise model comparisons. A positive (negative) value means that the model in the row is superior (inferior) to that in the column. t-statistics are followed by *, ** or *** if one model is significantly better than the other at a 0.1, 0.05 or 0.01 significance level, respectively. The in-sample model comparisons are based on Rivers and Vuong (2002) and the out-of-sample model comparisons are based on Diks et al. (2010).

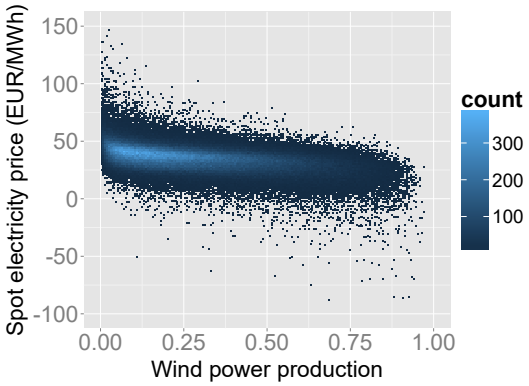
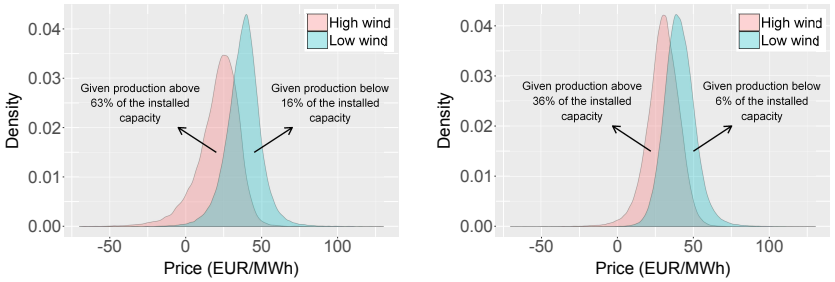


Fig. 5: Simulated joint distribution for the daily spot electricity prices and the wind power production in December 2013. The results are based on 10,000 simulations (for each day of December) and a Gaussian GAS model for the dependence structure.

Next, we use our model to study how different wind scenarios affect the distribution of prices. To this end, we perform one-month ahead simulations for all OOS months, i.e. a total of 24 months. Due to the elevated computational cost, we do not re-estimate the parameters of our joint model; we do however enlarge the conditioning set one month at the time. In Figs. 6(a)–(b), we display simulated empirical price distributions conditional on different levels of low/high wind scenarios. The simulations are grouped into winter (Dec., Jan., Feb.) and summer (Jun., Jul., Aug.) months. To define what a low/high wind scenario is during winter, we have considered the 20% and 80% quantiles of our actual OOS wind power production data during the specified winter months; the same procedure was followed for the summer months.

For both the winter and the summer period, we observe that the different wind scenarios shift the simulated price distributions. Moreover, the simulated distributions are left-skewed for the high wind scenarios (the skewness parameter is -1.98 for the winter months and -1.03 for the summer months), implying that extreme low prices are more likely than extreme high prices. For the low wind scenarios, the estimated distributions are right-skewed (the skewness is 0.64 and 1.06 for the winter and summer months respectively), thus implying the opposite compared to the high wind cases. We also notice that the low/high wind scenarios push the price distributions further apart for the winter months than the summer months, which we have confirmed by measuring the Kullback-Leibler distance between distributions. This can be explained by the fact that during summer periods, the dependence between electricity prices and wind power production is weaker than during winter periods, as earlier illustrated in Fig. 4. All these features are present



(a) Simulated prices: winter period

(b) Simulated prices: summer period

(c) Actual prices: winter period

(d) Actual prices: summer period

Fig. 6: Distributions for the daily spot electricity prices during winter and summer months, for the out-of-sample period 1 January 2013 to 31 December 2014, under the assumption of high and low wind power production. The simulated predictive distributions are based on 10,000 one-month-ahead simulations, using a Gassian GAS model for the dependence structure. Fig. 6(c) and Fig. 6(d) are based on 37 observations. To define the percentage corresponding to high/low wind scenarios during winter and summer, we used the 0.20 and 0.80 quantiles of the actual out-of-sample wind power production data.

when performing the same calculations on the actual data, which we show in Figs. 6(c)–(d), confirming that our empirical model captures the dynamics between daily spot electricity prices and wind power production.

5 Application to pricing and risk management

In the following we present applications of the proposed joint model for spot electricity prices and wind power production. We start by consider an energy trading company that enters into agreements with wind power producers, where a predetermined fixed price R is paid for the fluctuating wind power production. Since the production will first become known through the delivery period of the agreements, these products imply a volumetric risk. Furthermore, we assume that the trading company sells the production

it receives from the wind power producers on the day-ahead market, at a spot price we denote by S. Hence, the company will also be exposed to price risk. In the remaining of this section, we will refer to such agreements as fixed price for fluctuating wind power production agreements. With such a formulation, we can express the profit of the trading company as

$$\sum_{t=T_1}^{T_2} Q_t(S_t - R), \tag{27}$$

where time is measured in days, Q_t is the wind power production at time period t, S_t is the daily spot electricity price valid at t, and R is a fixed price set at the inception of the contract, which we denote t_0. Furthermore, the contract length spans from T_1 to T_2, where $t_0 < T_1 \leq T_2$. We note that to participate in the day-ahead electricity auction market, buy or sell bids have to be made to the exchange one day before delivery takes place. By working with the payoff in Eq. (27), we implicitly assume that the quantity we bid one day before equals the actual wind power production, i.e.

$$Q_t = \mathbb{E}_{t-1}\left[Q_t\right], \tag{28}$$

where $\mathbb{E}_{t-1}\left[Q_t\right]$ denotes the expectation at time $t-1$ for the production at time t. Thus, we assume no balancing risk.

What differentiates the product described above with payoff given in Eq. (27) from a standard forward contract is the production uncertainty associated with the former, and hence the presence of an additional risk due to the correlation between S and Q. If we express the price R in terms of the forward price F, Eq. (27) becomes

$$\sum_{t=T_1}^{T_2} Q_t(S_t - (F - c)), \tag{29}$$

where $F \equiv F(t_0, T_1, T_2)$ denotes the forward price at time t_0, for the delivery period from T_1 to T_2 and $c \equiv c(t_0, T_1, T_2)$ denotes the compensation that is to be subtracted from the forward price due to the negative correlation between prices and volume. So c can be thought of as the price of correlation risk. The fair value of c can be obtained by the usual practice of setting the discounted conditional expectation of the payoff given in Eq. (29) equal to zero.[12] To ease the presentation, we will assume a risk-free rate of zero, thus obtaining:

$$\mathbb{E}_{t_0}^{\mathbb{Q}}\left[\sum_{t=T_1}^{T_2} Q_t(S_t - (F - c))\right] = 0, \tag{30}$$

[12]Risk preferences could easily be included by e.g. introducing a simple volumetric risk aversion rule like in Kolos and Mardanov (2008).

$$c = F - \frac{\mathbb{E}_{t_0}^{\mathbb{Q}}\left[\sum_{t=T_1}^{T_2} Q_t S_t\right]}{\mathbb{E}_{t_0}^{\mathbb{Q}}\left[\sum_{t=T_1}^{T_2} Q_t\right]}. \tag{31}$$

With our framework, an estimate for c can easily be obtained by performing Monte Carlo simulations from the proposed copula model. However, this estimate will reflect the price of correlation risk under the physical or objective measure $\mathbb{P}$, since the model is fitted to historical spot electricity price and wind power production data. According to Eqs. (30) and (31), the expectations must be taken under a pricing measure $\mathbb{Q}$, that will reflect the risk premium charged by, in our context, the energy trading company offering the "insurance" to the wind power producer. Following Benth et al. (2008), the pricing measure $\mathbb{Q}$ is equivalent to $\mathbb{P}$, but needs not be an equivalent *martingale* measure due to the non-storability of our underlying "assets". Since neither electricity nor wind can be stored, they are not tradable assets in the classical sense. This implies that the spot–forward relation, for example, cannot be derived based on a buy-and-hold hedging argument. Instead, the usual practice is to simply define the forward price as the conditional expectation of the spot electricity price under the risk-neutral probability measure $\mathbb{Q}$, thereby turning the discounted spot price into a martingale (see Benth and Meyer-Brandis (2009) and Benth and Šaltytė Benth (2012)). Indeed, by defining

$$F(t_0, T_1, T_2) = \mathbb{E}_{t_0}^{\mathbb{Q}}\left[\frac{1}{T_2 - T_1 + 1}\sum_{t=T_1}^{T_2} S_t\right], \tag{32}$$

for the case of electricity, one can compute the implied market price of risk by considering the difference between quoted forward prices in the market and forward prices obtained by simulation with our model under $\mathbb{P}$.[13] In theory, the same could be done to estimate the risk premium associated with wind, however forwards with wind index as underlying are not currently traded in most European energy markets – and if they are, they are highly illiquid.

The fact that our setting is a bivariate one complicates the question of measure change even further, since aside from the marginal behavior of spot electricity price and wind power production under $\mathbb{Q}$, implied information regarding the market price of dependency risk must also be provided. A parametrization of this is not straightforward; in fact, the discussion can become quite extensive in the context of copulas and incomplete markets. Such a discussion is outside the scope of this paper, and we refer instead to Cherubini et al. (2004) for more details. Moreover, even if a theoretical procedure to

[13] For further discussions and empirical studies regarding pricing in electricity markets we refer to Benth et al. (2013), Burger et al. (2004), Kolos and Ronn (2008), and Lucia and Schwartz (2000).

calibrate the market price of dependency risk were to be established, the lack of exchange-traded instruments written on spot times wind would impede applying this in practice.

In light of the above discussion, we turn to the rational expectation hypothesis, which is a valid choice and a common assumption in this context (see e.g. Benth and Kettler (2011), Coulon et al. (2013), and Oum and Oren (2010)). This implies that we set $\mathbb{P} = \mathbb{Q}$, i.e. set the market price of risk to zero. Since we suspect a measure change to yield different prices, but not to alter the overall conclusions in our following empirical analysis, we find this assumption to be a reasonable one.

According to the payoff in Eqs. (27) or (29), it is clear by now that we are dealing with two sources of risk simultaneously: one is related to price uncertainty, and the other is related to production uncertainty; and since the market is incomplete, a perfect hedge cannot be performed. However, the price risk can be hedged. Here, we construct a simple hedging portfolio by taking a short position in a quantity H^* of standard forward power contracts. We assume that the hedge is static and performed at time t_0. The payoff of the hedge for the entire delivery period is given by

$$H^* \left(\mathbb{E}^{\mathbb{Q}}_{t_0} \left[\frac{1}{T_2 - T_1 + 1} \sum_{t=T_1}^{T_2} S_t \right] - \frac{1}{T_2 - T_1 + 1} \sum_{t=T_1}^{T_2} S_t \right), \tag{33}$$

or in a compact form

$$H^*(F - \overline{S}), \tag{34}$$

where F denotes the same forward price as in Eq. (29), and $\overline{S}$ denotes the average day-ahead electricity price for the same delivery period. To obtain H^*, we fix c to its value obtained from Eq. (31) and follow the standard procedure of minimizing the variance of the portfolio payoff:

$$\min_{H^*} \mathbb{V}\mathrm{ar}_{t_0} \left[\sum_{t=T_1}^{T_2} \tilde{Q}_t (S_t - (F - c)) + H^*(F - \overline{S}) \right]. \tag{35}$$

In Eq. (35), $\tilde{Q}_t = 24 \cdot Q_t \cdot \Lambda$, with Λ being the total installed capacity under the agreement that pays out a predetermined fixed price in return for the fluctuation wind power production. Since Q_t corresponds to daily wind power production relative to the total installed capacity in the entire DK1 price area, we need to transform this number to daily wind power production measured in MWh corresponding to the total installed capacity that the energy trading company actually has under agreement. By performing this transformation, we imply that our joint model is a good representation on a smaller scale. This is a realistic assumption as long as the energy trading

company manages a portfolio of diversified wind turbines in terms of type and location. Solving for H^* in Eq. (35) yields

$$H^* = \frac{\mathbb{C}\text{ov}_{t_0}\left[\overline{S}, \sum_{t=T_1}^{T_2} \tilde{Q}_t S_t\right] - (F-c)\mathbb{C}\text{ov}_{t_0}\left[\overline{S}, \sum_{t=T_1}^{T_2} \tilde{Q}_t\right]}{\mathbb{V}\text{ar}_{t_0}\left[\overline{S}\right]}. \tag{36}$$

It is clear that by hedging a quantity that is equal to H^*, we are protected on average and not against worst case scenarios, such as the combination of extremely low prices / high wind power production, which is a probable outcome in the DK1 price area. We could remedy the situation to a large extent by adding options to our portfolio, however this is outside the scope of the present paper. Work related to the optimal hedging of volumetric risk associated with wind power production is, to the best of our knowledge, not yet available. However, energy related discussions regarding the hedging of volumetric risk associated with consumers' load are presented in e.g. Oum and Oren (2009) and Oum et al. (2006), where many of the ideas can be transferred to our application. Nonetheless, our simple hedge is actually realistic since the market for options is very illiquid in DK1.

5.1 Example 1

Having developed a joint model for day-ahead electricity prices and wind power production, we can perform Monte Carlo simulations and use Eq. (31) to find the fair fixed price/compensation of a contract with any given specifications. Assume that we stand on the last trading day of November 2013, denoted t_0, and wish to find the fixed price for a front month contract, namely a December 2013 contract. Given all information available up to and including the valuation date t_0, we perform 10,000 simulations for price and quantity from our proposed joint model, where for each simulation we keep a path of length 31 (since we work with daily data) corresponding to the number of days in December. We note that we work with a fixed estimation window corresponding to the IS period, but enlarge the filtration, conditioning on the information up to and including the valuation date t_0. The contract specifications and results are summarized in Table 7, and we see that due to the negative correlation between prices and production, the compensation c that is to be subtracted from the forward price equals 3.24 EUR/MWh.

In addition to calculating the fixed price of a contract with fluctuating wind power production, we can extract information from the performed simulations that can be useful in a risk management context. We assume that agreements corresponding to an installed capacity of 500 MW are entered into on the last trading day of November 2013, with delivery December 2013. The price of a standard forward contract is fixed to its estimated value of

Contract information		
Time of valuation	t_0	29/11/2013
Contract length	T_1 to T_2	01/12/2013 to 31/12/2013
Pricing results		
Simulated forward price	$F \equiv F(t_0, T_1, T_2)$	35.26 EUR/MWh
Price of correlation risk	$c \equiv c(t_0, T_1, T_2)$ cf. Eq.(31)	3.24 EUR/MWh
Fixed price for fluctuating wind power production	$R \equiv R(t_0, T_1, T_2)$	32.02 EUR/MWh
Risk management results (500 MW installed capacity under agreement)		
5 % VaR without price hedge	See Fig. 7	1,099,248 EUR
5 % VaR with price hedge	See Fig. 7	465,485 EUR

Table 7: Simulation results for a December 2013 contract, with valuation date 29 November 2013, i.e. one business day before start delivery of the contract. The results are based on 10,000 simulations and using a Gassian GAS model for the dependence structure.

35.26 EUR/MWh, and the price of an agreement with a fluctuating wind power production is set to 32.02 EUR/MWh cf. Table 7. Given these specifications, we estimate the distribution of the portfolio payoff (see Fig. 7) and calculate the 5% Value-at-Risk (see Table 7) in two cases: One where the portfolio includes a price hedge, and one without a price hedge. When covering our price exposure in the forward market by assuming a short position corresponding to a quantity of H^* forwards, we observe that the variance of the profit distribution reduces significantly. In this example, the 5% Value-at-Risk is reduced from approximately EUR 1.1 million to EUR 0.5 million. It is also important to notice that the profit distribution is in both cases asymmetric, with a heavy-tail to the left, translating to the fact that expected losses are greater than expected gains.

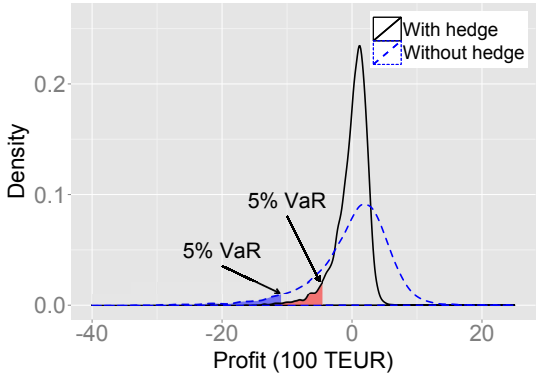
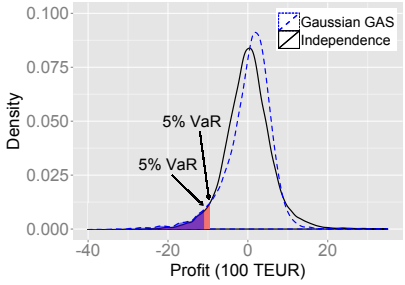


Fig. 7: Profit distributions for a December 2013 contract. The results are based on 10,000 simulations of price and quantity, using a Gaussian GAS model for the dependence structure. The forward price is fixed to 35.26 EUR/MWh, the compensation is fixed to 3.24 EUR/MWh and the total installed capacity of the portfolio equals 500 MW. The variance minimizing hedge quantity H^* is obtained by performing the calculation in Eq. (36).

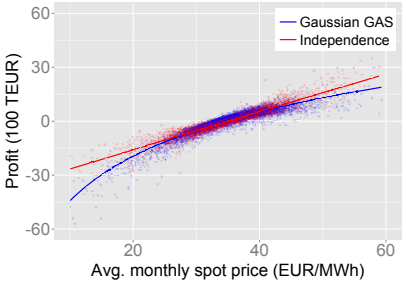
Revisiting the issue of pricing and considering the profit distributions in Fig. 7, alternative approaches to that of performing a measure change can be applied. An example can be to consider an a priori given 5% Value-at-Risk level that is acceptable, and solve for the correlation risk premium that satisfies this level.

To stress the effect of correlation on the profit distribution, we perform additional simulations, where all but the copula model remains unchanged. Specifically, we assume the independence copula and thus a zero compensation, instead of the Gaussian GAS model which we have established reflects the reality to a much greater extent. Fig. 8(a) illustrates the estimated profit distributions of the portfolio (with no hedge), and shows that the negative correlation implies a distribution that is more asymmetric. If prices and production were independent, we estimate a 5% Value-at-Risk of EUR

0.93 million corresponding to a reduction of approximately 15% compared to the 5% Value-at-Risk of EUR 1.1 million we obtain with the Gaussian GAS copula. Assuming independence would thus lead to an underestimation of risk. We also display the average spot electricity price for the period of the contract as a function of the estimated profit in Fig. 8(b). Under independence, we observe that the payoff becomes linear, and hence forwards would suffice as hedging instruments. Under negative dependence, the payoff becomes non-linear, emphasizing the need for options in the hedging portfolio. Furthermore, we observe that a larger profit (smaller loss) can be obtained if prices and production are independent as we move away from the mean average price of 35.26 EUR/MWh. This is also supported by Fig. 8(a), where we observe that the negative correlation implies that a smaller probability is assigned to large profits, and a higher probability is assigned to large losses.



(a) Profit distributions for a portfolio of fixed price December 2013 contracts under different assumptions for the dependence structure.



(b) Estimated payoffs for a portfolio of fixed price December 2013 contracts, as a function of the avg. monthly spot electricity price.

Fig. 8: Illustration of the importance of correlation in the analysis of profit. The results are based on 10,000 Monte Carlo simulations with a Gaussian GAS copula (c is fixed to 3.24 EUR/MWh) and the independence copula (c is fixed to 0 EUR/MWh), respectively. The total installed capacity of the portfolio is set to 500 MW, and the same marginal models for prices and wind power production are used.

5.2 Example 2

In Section 3.4, we have established that some of the fitted time-varying copula models are superior to the constant ones, see e.g. Table 6. Here, we wish to investigate if this also holds when comparing the actual profits or losses generated with different copula models. For this, we consider the OOS period corresponding to the years 2013 and 2014. We assume the following trading strategy: On the last trading day of each month (Dec. 2012 - Nov. 2014), we enter into front month agreements with wind power generators, where a fixed price is paid for the fluctuating wind power production. The total

installed capacity of each monthly portfolio is fixed to 500 MW. We perform 10,000 simulations from joint models with the different copula specifications that we wish to compare against each other (marginal models are kept unchanged), and estimate compensations c and hedge quantities H^* for each month at a time using Eqs. (31) and (36). For each monthly portfolio, we then calculate the realized profit using the actual daily electricity prices, actual daily wind power production[14] and actual forward prices.

For clarity, let us consider a concrete example: We stand on the last trading day of December 2012, and wish to enter into fixed price agreements with fluctuating wind power production for the January 2013 month. To enter the contract, we first estimate the fixed price that we are willing to pay for the production that we will receive during January. Since we also perform a hedge in the forward market, we estimate the quantity of forwards we are to short. In this example, we will use a constant Clayton copula to describe the dependence between prices and wind power production, and hence we obtain an estimated compensation denoted by $\hat{c}^{\text{Clayton}}_{t_0,\text{Jan}}$ and an estimated hedge quantity $\hat{H}^{*,\text{Clayton}}_{t_0,\text{Jan}}$. On the last trading day of December 2012, we can observe the actual forward price $F^{\text{Obs}}_{t_0}$, and thus the fixed price we offer the wind power producers is

$$\hat{R}^{\text{Clayton}}_{t_0,\text{Jan}} = F^{\text{Obs}}_{t_0} - \hat{c}^{\text{Clayton}}_{t_0,\text{Jan}}. \tag{37}$$

By the end of January 2013, we will also have observed the actual daily spot electricity prices S^{Obs} and the actual daily wind power production Q^{Obs} for the DK1 price area. With this information, we can now approximate the actual profit resulting from the trades we have performed:

$$\text{Actual profit}_{\text{Jan}} = \underbrace{\sum_{t=T_1}^{T_2} \tilde{Q}^{\text{Obs}}_t (S^{\text{Obs}}_t - \hat{R}^{\text{Clayton}}_{t_0,\text{Jan}})}_{\text{Agreement payoff}} + \underbrace{\hat{H}^{*,\text{Clayton}}_{t_0,\text{Jan}} (F^{\text{Obs}}_{t_0} - \overline{S}^{\text{Obs}})}_{\text{Hedge payoff}} \tag{38}$$

where $t_0 = 31/12/2012$, $T_1 = 01/01/2013$, $T_2 = 31/01/2013$ and $\tilde{Q}^{\text{Obs}}$ is the approximation

$$\tilde{Q}^{\text{Obs}}_t = Q^{\text{Obs}}_t \cdot 24\ (\text{h}) \cdot 500\ (\text{MW}). \tag{39}$$

The results obtained by performing the above calculations for all OOS months with different copula specifications are presented in Table 8. The numbers

[14]The actual daily wind power production is given in % for the entire price area, but only a subset of the existing wind turbines in DK1 is part of our portfolio. Therefore, we note that the realized profit we calculate is an approximation; We obtain the actual production of the wind turbines under agreement by multiplying the actual daily wind power production for the entire price area with the assumed installed capacity of the portfolio of 500 MW and 24 hours.

show that the joint model with a Gaussian GAS copula provides the highest (lowest) monthly profit (loss) in 15 out of the 24 months, corresponding to 62.50%. Considering the second column block of Table 8, we see that it is indeed the Gaussian GAS and the Joe-Frank GAS that yield the lowest losses in average, which supports the results we obtained in Section 3.4. Hence, allowing for time variation in a suitable copula model is beneficial. The constant Clayton specification performs the poorest, generating the largest average loss. This is again in accordance with earlier findings, where we have established that the constant Clayton specification is not suitable for the dependence of prices and wind power production, and also least suitable among the copula models we consider in Table 8. The time-varying Gaussian and Joe-Frank copulas outperform the other copulas since they are able to capture the increasingly negative correlation we observe towards the last years of our sample (see Fig. 4); and thus, they are able to generate larger compensations. For instance, the constant Gaussian copula yields an average compensation for the OOS period of 2.69 EUR/MWh, while the Gaussian GAS copula yields a value of 2.98 EUR/MWh.

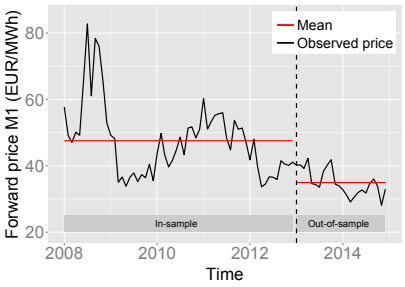
	Highest profit (lowest loss) per month for the OOS period		Realized average profit for the OOS period (EUR/MWh)	
	Constant	Time-varying (GAS)	Constant	Time-varying (GAS)
Gaussian	4.16%	**62.50%**	-0.9144	**-0.6103**
Gumbel	16.67%	0.00%	-0.8693	-0.7961
Joe-Frank	0.00%	16.67%	-1.0294	-0.6666
Clayton	0.00%	-	-1.1465	-

Table 8: OOS model comparisons based on realized monthly profit/loss. In the first column block, we calculate how often each copula model yields the lowest monthly loss or the highest monthly profit. The second column block presents the realized average profit/loss (in EUR/MWh) for selected copula models, obtained by dividing the total realized cash-flow for the period by the total realized wind power production. All results are based on the same trading strategy and 10,000 simulations.

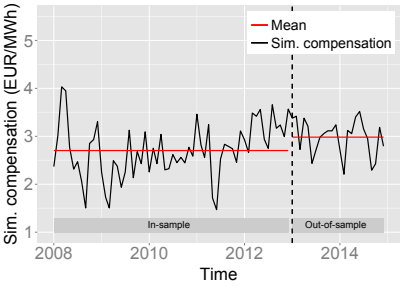
Lastly, we illustrate in Fig. 9 the evolution of actual forward prices and also the evolution of compensations estimated with our proposed joint model for electricity prices and wind power production, i.e. the one with the Gaussian GAS copula specification for the dependence structure.

Overall, compensations amount to an increasing percentage of the forward price during the period of our study. Clearly, this is mainly due to the decreasing tendency in forward prices, but also due to the slight increase in compensations if we consider the IS and OOS average compensations. The slight increase in compensations can be justified by the increasing installed capacity of wind power that Denmark has experienced over the past years - and hence the stronger dependence between wind power production and electricity prices. This also explains the decrease in forward prices, but only

to a small extent; the major contributing factor here has been the decreasing raw material prices. The reduction in forward price due to the correlation risk amounts to an average of 7%, and can reach as high as 11%. A similar conclusion is reached by Elberg and Hagspiel (2015), where the authors study the market value of wind power at different locations in Germany, and show that this value is reduced compared to the average spot price as a result of increasing wind power penetration.



(a) Actual forward prices for a front month (M1) contract, valid the last trading day before delivery start. Total of 84 prices, one for each month in our IS and OOS sample.



(b) Simulated compensations for a front month contract, valued the last trading day before delivery start. The results are based on 10,000 simulations, with a Gaussian GAS models for the dependence structure.

Fig. 9: Evolution of actual forward prices and estimated compensations.

6 Conclusion

This work concentrates on the dependency between daily spot electricity prices and wind power production, and its role regarding the pricing and the risk distributions associated with contracts exposed to both price and volumetric risk. The analysis is carried out on data from the Danish power market, which is characterized by a high penetration of wind power in the system. We propose a copula approach since we wish to concentrate on the dependence in more detail. We employ marginal models of the ARMA–GARCH type and parametric error distributions for each individual variable, and then link the innovations through various constant and time-varying copulas. Based on statistical tests concerning copula selection, we choose a time-varying Gaussian copula as our preferred specification for the dependence structure. By performing Monte Carlo simulation studies, we are able to visualize the joint empirical distribution implied by our model, and see how this deviates from the Gaussian benchmark. Also, we study the distribution of prices conditional on different levels of wind power penetration,

and show that prices decrease (increase), on average, at times of high (low) levels of wind power production; the shape of the conditional distribution of prices is also affected by the different levels of wind power production. These findings confirm what previous studies concerned with the impact of wind power – or predicted wind power penetration – on electricity prices have shown (e.g. Gelabert et al. (2011), Jónnson et al. (2010)).

We apply the developed empirical model in the context of an energy trading company offering wind power producers a predetermined fixed price for their fluctuating wind power production. We find that the correlation risk premium that the energy trading company should charge when entering such agreements is significant, amounting to 7% of the price of a standard forward power contract on average. Furthermore, our results indicate that the choice of copula impacts the price of correlation risk: An out-of-sample study based on comparing realized profits generated by different copulas shows that introducing time-variation in the copula model is beneficial. When considering the profit distribution, we find that under independence, the risk is underestimated. Additionally, we show that a simple hedge in the forward market can reduce e.g. the 5% Value-at-Risk of the profit distribution significantly. However, due to the non-linearity of profit, options should be included in the hedging portfolio in order to reduce the risk even further; this could be an interesting subject for further research.

Finally, although our empirical study concentrates on the Danish power market, the mechanism of spot price formation in e.g. other European electricity markets is also based on matching supply and demand. Further, wind power production has a very low marginal cost, ensuring that it will always be represented in the merit order stack. Due to the physical conditions upon which the day-ahead electricity markets are based, we believe that the proposed modeling framework is relevant and can be applied to other electricity markets that, like Denmark, rely heavily on wind power production. Such extensions are left for future research.

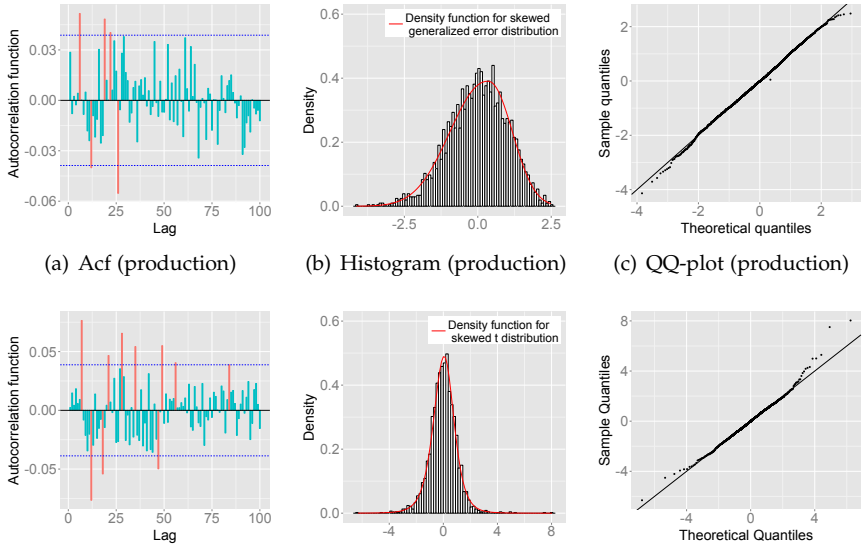
Acknowledgments

The authors would like to express their gratitude to Rikke Preisler Vilstrup for her contributions to this paper through a mutual unpublished master thesis project, and Thomas Aalund Fredholm, Christian Sønderup and Jakob Vive Munk at Neas Energy for posing the problem and providing helpful comments and suggestions. The authors also thank the referees for providing constructive criticism and suggestions that enhanced the quality of this paper.

Funding

Anca Pircalabu acknowledges support from Innovation Fund Denmark, grant number 4135-00082B.

A Additional figures



(a) Acf (production) (b) Histogram (production) (c) QQ-plot (production)

(d) Acf (price) (e) Histogram (price) (f) QQ-plot (price)

Fig. 10: Diagnostics for marginal models for spot electricity price and wind power production. Figs. 10(a) - 10(c) display the autocorrelation function, histogram and quantile plot for the standardized residuals resulting from the marginal model for wind power production. Correspondingly, Figs. 10(d) - 10(f) display the same diagnostics for the standardized residuals resulting from the marginal model for spot electricity prices.

B Properties of selected copula models

Elliptical copulas

A bivariate elliptical copula is defined as

$$C(u_1, u_2) = F(F_1^{-1}(u_1), F_2^{-1}(u_2)), \tag{40}$$

where $u_1, u_2 \in [0, 1]$. The elliptical copulas we consider in this paper are the Gaussian copula and the Student t copula. In the case of the Gaussian copula, F corresponds to the bivariate standard normal *cdf*, and F_1^{-1} and F_2^{-1} denote the inverse of the univariate standard normal *cdf*. In the case of the Student t copula, F corresponds to the bivariate Student t *cdf*, and F_1^{-1} and F_2^{-1} denote the inverse of the univariate Student t *cdf*.

Archimedian copulas

A bivariate Archimedian copula is defined as

$$C(u_1, u_2) = \phi^{-1}(\phi(u_1) + \phi(u_2)), \tag{41}$$

where $u_1, u_2 \in [0, 1]$ and $\phi : [0, 1] \to [0, \infty)$ is a generator function satisfying that ϕ^{-1} is monotone on $[0, \infty)$. The Archimedian copulas we consider in this paper are Clayton and Gumbel, and also the combinations Joe-Frank and a symmetrized version of Joe-Clayton.

Properties of selected copulas are collected in Table 9.

	Generator function	Parameter range	Sym.	Neg. dep.	Tail dependence
Elliptical family					
Gaussian	-	$\rho \in (-1, 1)$	Yes	Yes	0
Student t	-	$\rho \in (-1, 1)$	Yes	Yes	$2t_{\nu+1}\left(-\sqrt{\nu+1}\sqrt{\frac{1-\rho}{1+\rho}}\right)$
	-	$\nu > 2$			
Archimedian family					
Clayton	$\frac{1}{\theta}(u^{-\theta} - 1)$	$\theta > 0$	No	No	$(2^{-\frac{1}{\theta}}, 0)$
Gumbel	$(-\log u)^{\theta}$	$\theta \geq 1$	No	No	$(0, 2 - 2^{\frac{1}{\theta}})$
Joe-Clayton	$(1 - (1-u)^{\theta})^{-\delta} - 1$	$\theta \geq 1$	No	No	$(2^{-\frac{1}{\delta}}, 2 - 2^{\frac{1}{\theta}})$
		$\delta > 0$			
Joe-Frank	$-\log\left(\frac{1-(1-\delta u)^{\theta}}{1-(1-\delta)^{\theta}}\right)$	$\theta \geq 1$	No	No	$(0, 0)$
		$\delta \in (0, 1]$			

Table 9: Overview over the properties of selected copulas. Copula features and notation coincide with the R-package **CDVine**. t denotes the univariate Student t *pdf* with $\nu + 1$ degrees of freedom. We note that we consider the symmetrized Joe-Clayton copula proposed in Patton (2006), which is an even more flexible version of the Joe-Clayton copula, since is allows for both symmetry and asymmetry. See Patton (2006) for details regarding this copula.

References

C. Alexander. Correlation in crude oil and natural gas markets. In Vincent Kaminski, editor, *Managing Energy Price Risk: The New Challenges and Solutions*. Risk Publications, 2004.

K. Avdulaj and J. Barunikl. Are benefits from oil–stocks diversification gone? New evidence from a dynamic copula and high frequency data. *Energy Economics*, 51:31 – 44, 2015.

F. E. Benth and P. C. Kettler. Dynamic copula models for the spark spread. *Quantitative Finance*, 11(3):407 – 421, 2011.

F. E. Benth and T. Meyer-Brandis. The information premium for non-storable commodities. *Journal of Energy Markets*, 2(3):111 – 140, 2009.

F. E. Benth and J. Šaltytė Benth. *Modelling and Pricing in Financial Markets for Weather Derivatives*. World Scientific, 2012.

F. E. Benth, J. Šaltytė Benth, and S. Koekebakker. *Stochastic Modelling of Electricity and Related Markets*. World Scientific, 2008.

F. E. Benth, R. Biegler-König, and R. Kiesel. An empirical study of the information premium on electricity markets. *Energy Economics*, 36:55 – 77, 2013.

D. Berg. Copula goodness-of-fit testing: An overview and power comparison. *European Journal of Finance*, 15(7-8):675 – 701, 2009.

T. Bollerslev. Glossary to ARCH (GARCH). CREATES Research Papers 2008-49, School of Economics and Management, University of Aarhus, Sep 2008. URL `https://ideas.repec.org/p/aah/create/2008-49.html`.

M. Burger, B. Klar, A. Muller, and G. Schindlmayr. A spot market model for pricing derivatives in electricity markets. *Quantitative Finance*, 4(1):109 – 122, 2004.

U. Cherubini and E. Luciano. Bivariate option pricing with copulas. *Applied Mathematical Finance*, 9:69 – 85, 2002.

U. Cherubini, E. Luciano, and W. Vecchiato. *Copula Methods in Finance*. John Wiley & Sons, 2004.

M. Coulon, W. B. Powell, and R. Sircar. A model for hedging load and price risk in the Texas electricity market. *Energy Economics*, 40:976 – 988, 2013.

D. Creal, S. J. Koopman, and A. Lucas. Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, 28(5):777 – 795, 2013.

A. Dias and P. Embrechts. Dynamic copula models for multivariate high-frequency data in finance. Warwick Business School, Finance Group, Working Paper, 2004.

C. G. H. Diks, V. Panchenko, and D. J. C. van Dijk. Out-of-sample comparison of copula specifications in multivariate density forecasts. *Journal of Economic Dynamics and Control*, 34(9):1596 – 1609, 2010.

C. Elberg and S. Hagspiel. Spatial dependencies of wind power and interrelations with spot price dynamics. *European Journal of Operational Research*, 241(1):260 – 272, 2015.

P. Embrechts, A. J. McNeil, and D. Straumann. Correlation and dependence in risk management: Properties and pitfalls. In M. A. H. Dempster, editor, *Risk Management: Value at Risk and Beyond*, pages 176 – 223. Cambridge University Press, 1999.

J.-D. Fermanian and M. Wegkamp. Time dependent copulas. *Journal of Multivariate Analysis*, 110:19 – 29, 2012.

L. Gelabert, X. Labandeira, and P. Linares. An ex-post analysis of the effect of renewables and cogeneration on Spanish electricity prices. *Energy Economics*, 33:59 – 65, 2011.

C. Genest, B. Remillard, and D. Beaudoin. Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and Economics*, 44(2): 199 – 213, 2009.

R. Giacomini and H. White. Tests of conditional predictive ability. *Econometrica*, 74(6):1545 – 1578, 2006.

C. González-Pedraz, M. Moreno, and I. Peña. Portfolio selection with commodities under conditional copulas and skew preferences. *Quantitative Finance*, 15(1):151 – 170, 2015.

H. Joe. *Multivariate Models and Dependence Concepts*. Chapman and Hall, 1 edition, 1997.

T. Jónnson, P. Pinson, and H. Madsen. On the market impact of wind energy forecasts. *Energy Economics*, 32(2):313 – 320, 2010.

S. P. Kolos and K. Mardanov. Pricing volumetric risk. *Energy risk*, pages 54 – 60, 2008.

S. P. Kolos and E. I. Ronn. Estimating the commodity market price of risk for energy prices. *Energy Economics*, 30(2):621 – 641, 2008.

J. J. Lucia and E. S. Schwartz. Electricity prices and power derivatives: Evidence from the Nordic Power Exchange. *Review of Derivatives Research*, 5:5 – 50, 2000.

R. I. McKinnon. Future markets, buffer stocks, and income stabiliy for primary producers. *Journal of Political Economy*, 75(6):844 – 861, 1967.

G. Moschini and H. Lapan. The hedging role of options and futures under joint price, basis, and production risk. *International Economic Review*, 36(4): 1025 — 1049, 1995.

R. B. Nelsen. *An Introduction to Copulas*. Springer, 1999.

D. B. Nelson and C. Q. Cao. Inequality constraints in the univariate GARCH model. *Journal of Business & Economic Statistics*, 10(2):229 – 235, 1992.

Y. Oum and S. S. Oren. VaR constrained hedging of fixed price load-following obligations in competitive electricity markets. *Risk and Decision Anaysis*, 1 (1):43 – 56, 2009.

Y. Oum and S. S. Oren. Optimal static hedging of volumetric risk in a competitive wholesale electricity market. *Decision Analysis*, 7(1):107 – 122, 2010.

Y. Oum, S. S. Oren, and S. Deng. Hedging quantity risks with standard power options in a competitive wholesale electricity market. *Naval Research Logistics*, 53(7):697 – 712, 2006.

F. Paraschiv, D. Erni, and R. Pietsch. The impact of renewable energies on EEX day-ahead electricity prices. *Energy Policy*, 73:196 – 210, 2014.

A. J. Patton. Modelling asymmetric exchange rate dependence. *International Economic Review*, 47(2):527 – 556, 2006.

A. J. Patton. Copula methods for forecasting multivariate time series. In Graham Elliott and Allan Timmermann, editors, *Handbook of Economic Forecasting*, volume 2B, pages 899 – 960. Elsevier B.V., 2013.

D. N. Politis and J. P. Romano. The stationary bootstrap. *Journal of the American Statistical Association*, 89(428), 1994.

D. N. Politis and H. White. Automatic block-length selection for the dependent bootstrap. *Econometric Reviews*, 23(1):53 – 70, 2004.

D. N. Politis, H. White, and A. J. Patton. Correction: Automatic block-length selection for the dependent bootstrap. *Econometric Reviews*, 28(4):372 – 375, 2009.

D. Rivers and Q. Vuong. Model selection tests for nonlinear dynamic models. *Econometrics Journal*, 5:1 – 39, 2002.

I. A. De Lira Salvatierra and A. J. Patton. Dynamic copula models and high frequency data. Economic Research Initiatives at Duke (ERID), Working Paper No. 165, June 2013.

U. Schepsmeier and J. Stöber. Derivatives and Fisher information of bivariate copulas. *Statistical Papers*, 55(2):525 – 542, 2014.

A. Sklar. Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de L'Université de Paris*, 8:229 – 231, 1959.

X. Wen, Y. Wei, and D. Huang. Measuring contagion between energy market and stock market during financial crisis: A copula approach. *Energy Economics*, 34:1435 – 1446, 2012.

Paper II

A mixed C-vine copula model for hedging price and volumetric risk in wind power trading

List of authors: Anca Pircalabu[*,†] and Jesper Jung[†]
[*] Department of Mathematical Sciences, Aalborg University
[†] Quantitative Analytics, Neas Energy

Published in
Quantitative Finance Vol. 17(10), pp. 1583–1600, 2017.
DOI: http://dx.doi.org/10.1080/14697688.2017.1307511

The layout has been revised.

ABSTRACT

When energy trading companies enter into long-term agreements with wind power producers, where a fixed price is paid for the fluctuating production, they are facing a joint price and volumetric risk. Since the payoff of such agreements is non-linear, a hedging portfolio would ideally consist of not only forwards, but also a basket of e.g. call and put options. Illiquidity and an almost non-existent market for options challenge however the optimal hedging of joint price and volumetric risk in many market places. Here, we consider the case of the Danish power market, and exploit its strong positive correlation with the much more liquid German market to construct a proxy hedge. We propose a three-dimensional mixed vine copula to model the evolution of the Danish and German spot electricity prices and the Danish wind power production. We construct a realistic hedging portfolio by identifying various instruments available in the market, such as real options in the form of the right to transfer electricity across the border and the right to convert electricity to heat. Using the proposed vine copula to determine optimal hedging decisions, we show that significant benefits are to be drawn by extending the hedging portfolio with the proposed instruments.

1 Introduction

In recent years, the share of wind power production has increased significantly in many market places. As a consequence, considering in more detail the joint price and volumetric risk associated with wind power trading has become highly relevant. Not only are both spot electricity prices and wind power production notoriously volatile and unpredictable, but the mechanism of spot price formation implies a negative relation between the two that adds another dimension to the problem. Since spot prices are set by matching supply and demand curves, with supply curves prioritizing the cheapest generation sources (hereby wind turbines), the negative relation intensifies with a growing share of wind power production in the electrical grid. Wind power producers can rarely manage such involved risk, which would require constructing and rebalancing hedging portfolios through trading different suitable derivative instruments on the exchange, and they often seek to transfer their risk to another party. Alternatively in some market places, the price risk of wind power producers might automatically be removed (or partially removed) by support mechanisms initiated to promote the growth in renewable generation.

In this paper, we consider the problem of hedging joint price and volumetric risk in the context of an energy trading company entering into longer-term financial agreements with wind power producers, where a pre-determined

fixed price is paid for the fluctuating wind power production. Such financial agreements are becoming common in e.g. Denmark, Sweden, and the UK, which all rely to some extent on wind power generation.

The problem of joint price and volumetric risk was first addressed by McKinnon (1967) in relation to the classical farmer's problem, and has since been studied in different contexts. In the case of the energy markets, the main focus has been directed towards the joint price and volumetric risk associated with the customers' electricity demand. Some of the first to address many aspects of this problem, including hedging, were Oum et al. (2006) and Oum and Oren (2009, 2010). However, as acknowledged by Oum and Oren (2010), their proposed hedge had limited application in practice, as they allowed for trading in calls and puts written on the spot electricity price with a full spectrum of strikes, which was not possible at the time - and neither is it now. Later, the work of Coulon et al. (2013) focused on constructing a structural model for hedging the same type of risk, and used hedging instruments that are more available.

Our problem differentiates itself from the existing literature in different aspects. First, we consider volumetric risk on the supply rather than the demand side, which raises different modeling challenges. Second, we choose to analyze data from the Danish power market, since Denmark is among the top wind power countries in the world. This comes however at the cost of considering a market place that is characterized as rather illiquid, and where in addition the availability of options is limited. Nonetheless, this is the situation in the majority of European power markets. Since the hedging possibilities are naturally limited, we extend the modeling framework to a three-dimensional case, where we exploit the strong positive relation between the Danish and the German spot electricity prices to enlarge the set of available hedging instruments.

To model the joint behavior of electricity prices in Denmark and Germany and wind power production in Denmark, we consider a vine copula, which is essentially a construction formed using pairs of bivariate copulas. J. (1997) was the first to propose such a construction, and more general settings were later proposed by Bedford and Cooke (2001, 2002), who introduced the regular vines with the subclasses known as D-vines and canonical (C)-vines. Vine copula models allow for very flexible multivariate distributions, which motivates our model choice.

Copula models have found various applications in economics, finance and risk management. The literature is extensive for the bivariate case, with some examples being Cherubini and Luciano (2002) who consider bivariate option pricing, Patton (2006) who provides an application to exchange rates, Benth and Kettler (2011) who consider the spark spread, Avdulaj and Barunikl (2015) who investigate oil-stock diversification, and Elberg and Hagspiel (2015) who study spatial dependencies of wind power and interrelations

with the spot price in Germany. Copulas are also a popular choice when it comes to higher dimensional problems, and the range of applications is again broad: Grothe and Schnieders (2011) investigate the optimal allocation of wind farms, Gatfaoui (2016) links the gas, oil and stock markets through trivariate copulas, Brechmann and Czado (2013) consider vines in the context of financial risk management, and Reboredo and Ugolini (2015) model systemic sovereign debt risk using vine copulas.

We offer two main contributions in this paper: First, we construct a flexible empirical model that captures the marginal behavior of the individual variables accurately, as well as the dependency structure between the variables. Second, we identify instruments that can be used in practice to hedge the multiplicative price and volumetric risk, and study their hedging benefits. Based on empirical examples, we find that the variance of the revenue distribution can be significantly reduced by including these instruments, compared to the common strategy of using solely power forwards in a hedging portfolio. Furthermore, we employ different alternative models to facilitate hedging decisions, with the purpose of highlighting the effects of heavy tails in the margins, tail dependence and time variation in the dependence structure.

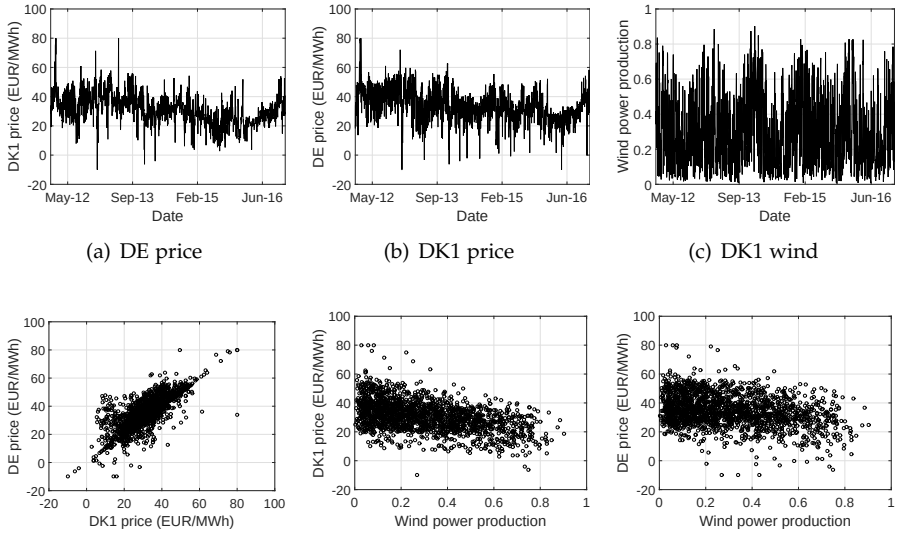
The paper is structured as follows: In section 2 we present the data. Section 3 introduces the modeling framework and presents empirical results. In Section 4, we study hedging applications and compare different vine copula specifications against each other. Finally we conclude in Section 5.

2 Data

Our empirical study relies on data from the Danish and the German power market for the period 01 January 2012 to 12 December 2016, corresponding to 1808 daily observations. Since the Danish power market is divided into two price areas, namely Western Denmark (DK1) and Eastern Denmark (DK2), where the share of wind power is significantly higher in DK1, we restrict our attention to this area. Hence, the first data input consists of the daily spot electricity price in DK1. Next, we consider the DK1 wind power production as a percentage of the installed capacity (wind power production index) on a daily basis, and the last data input consists of the daily spot electricity price in Germany (DE). The three time series are illustrated in Figs. 1(a) - 1(c). Regarding the price data, we have truncated a few extreme observations as they are considered outliers. Specifically, 4 extreme observations were truncated in the DK1 price series, and 6 extreme observations were truncated in the case of DE.

In Figs. 1(d) - 1(f), the data are plotted against each other pairwise. The first plot reveals a strong positive relation between the two prices, substan-

tiating the idea of a proxy hedge, that is using instruments from the much more liquid DE market to hedge price and volume risk in DK1. The two remaining plots show a negative relation between wind power production and prices in the two areas we consider. As mentioned previously, this relation is expected due to the mechanism of price formation, and implies an additional correlation risk when dealing with a simultaneous exposure to spot electricity price and wind power production.



(a) DE price (b) DK1 price (c) DK1 wind

(d) DK1 price vs. DE price (e) DK1 price vs. DK1 wind (f) DE price vs. DK1 wind

Fig. 1: Historical data

3 A vine copula model for spot electricity prices and wind power production

Let us denote by $\boldsymbol{x}_t = (x_{1,t}, x_{2,t}, x_{3,t})$ for $t = 1, \ldots, T$, the three-dimensional time series consisting of the spot electricity price in DK1, the spot electricity price in DE and the wind power production in DK1. We are interested in modeling the joint (conditional) distribution of $\boldsymbol{x}_t$ using a vine copula approach. In the following, we elaborate on this approach in the context of model construction and estimation. Since our problem is three-dimensional, we will concentrate on such a case.

3.1 Decomposing a trivariate distribution function using pair-copulas

In a joint distribution function, information regarding both the marginal behavior of each random variable and the dependence structure between the random variables is generally embedded. The copula is introduced as a tool to isolate the study of the dependence structure from that of the marginal behavior of the individual variables. A copula is a multivariate distribution function C defined on the unit cube, with standard uniform margins. The central result when working with copulas, Sklar's Theorem (Sklar (1959)), shows how the marginal distributions and the copula are connected to the joint distribution. For the three-dimensional case, the theorem states that if we let F be the joint distribution function of the random vector $\boldsymbol{X} = (X_1, X_2, X_3)$ with marginal distribution functions F_1, F_2, F_3, then there exists a three-dimensional copula C such that

$$F(x_1, x_2, x_3) = C\{F_1(x_1), F_2(x_2), F_3(x_3)\}. \tag{1}$$

If the marginals are continuous, then the copula is unique and defined through

$$C(u_1, u_2, u_3) = F\{F_1^{-1}(u_1), F_2^{-1}(u_2), F_3^{-1}(u_3)\}, \tag{2}$$

where u_i are standard uniform variables and F_i^{-1} are the ordinary inverses of the marginal distribution functions, for $i = 1, 2, 3$. Increasing further the practical applicability of this representation is the fact that the converse of Sklar's Theorem holds: If we start with a copula C and margins F_1, F_2, F_3, then F defined in equation (1) is a joint distribution function with margins F_1, F_2, F_3. This allows us to construct multivariate distributions in a flexible way, using arbitrary margins and copulas as building blocks.

Next, we turn our attention to density functions. Assuming densities exist, the joint probability density function f corresponding to the joint distribution F defined in equation (1) is given by

$$f(x_1, x_2, x_3) = f_1(x_1) \cdot f_2(x_2) \cdot f_3(x_3) \cdot c\{F_1(x_1), F_2(x_2), F_3(x_3)\}, \tag{3}$$

where f_i denote the marginal density functions, for $i = 1, 2, 3$, and c denotes a trivariate copula density.

Using that a joint density function $f(x_1, x_2, x_3)$ can be factorized as e.g.

$$f(x_1, x_2, x_3) = f_1(x_1) \cdot f_{2\,|\,1}(x_2 \,|\, x_1) \cdot f_{3\,|\,2,1}(x_3 \,|\, x_2, x_1), \tag{4}$$

and the fact that each conditional density in equation (4) can be written as the appropriate pair-copula multiplied by a (conditional) marginal density, an iterative decomposition can be used to obtain the expression for the three-

dimensional vine structure:

$$\begin{aligned} f(x_1, x_2, x_3) = f_1(x_1) \cdot f_2(x_2) \cdot f_3(x_3) \\ \cdot c_{12}\{F_1(x_1), F_2(x_2)\} \cdot c_{13}\{F_1(x_1), F_3(x_3)\} \\ \cdot c_{23\,|\,1}\{F_{2\,|\,1}(x_2 \,|\, x_1), F_{3\,|\,1}(x_3 \,|\, x_1)\}. \end{aligned} \tag{5}$$

For more details on the recursive conditioning used to obtain equation (5) and more general constructions, see Aas et al. (2009) and Czado et al. (2012). We note here that the decomposition given in equation (5) is not unique since different orderings of the variables are possible, but will elaborate on the procedure used to select the vine copula model in Section 3.2. Furthermore, the marginal conditional distributions $F_{2\,|\,1}(x_2 \,|\, x_1)$ and $F_{3\,|\,1}(x_3 \,|\, x_1)$ entering the decomposition in equation (5) are often referred to as $h-$functions, and are defined as

$$h(u_i \,|\, u_j; \Theta) = F(x_i \,|\, x_j) = \frac{\partial C\{F(x_i), F(x_j); \Theta\}}{\partial F(x_j)}, \tag{6}$$

where $u_i = F(x_i)$ and $u_j = F(x_j)$ are standard uniform, and C is a parametric bivariate copula with parameters Θ.

Keeping in mind the decomposition given in equation (5), we describe below the two-step procedure we follow to model the joint behavior of prices in DK1, prices in DE and wind power production in DK1.

In the first step, we consider each time series individually to filter out seasonality and serial time dependence. To correct each series for deterministic seasonality, we consider appropriate seasonal functions which we fit by simple linear regression. Then, we fit ARMA-GARCH type models to the deseasonalized data in order to filter out the serial dependence in the conditional mean and the conditional variance. To allow for more flexibility, we relax the usual Gaussian assumption for the standardized innovation processes.

In the second step, we consider a vine copula model for the approximately *i.i.d.* standardized innovations obtained in the first step. That is, we model the dependency of processed data, after seasonality and marginal time dependencies have been removed. Specifically, the type of vine we consider is the so-called mixed C-vine. The *C* refers to the graphical representation of a sequence of trees where each tree has a central node, and the term *mixed* refers to having no restrictions on the copula family chosen for each individual pair.

3.2 Inference for the three-dimensional model

When considering full inference in the context of a vine copula, three issues must be addressed: 1) the selection of a specific decomposition, 2) the

selection of bivariate copula families in the vine - given that an estimation procedure is in place, and 3) estimation of the full model.

Regarding the selection of a specific decomposition, we follow the procedure in Czado et al. (2012) and order variables according to their importance. To achieve this, we compute Kendall's τ for all possible pairings of the variables. Then, the variable most related to all other variables is set to be the central node of the first tree, and so forth.

Selection of the bivariate copulas in the vine is based on the Akaike information criterion (AIC), which is a commonly used criterion in this context, see e.g. Czado et al. (2012) and Elberg and Hagspiel (2015). The goodness-of-fit (GoF) of the chosen copulas is verified by performing the Cramér-von Misses (CvM) test, see Genest et al. (2009) and Berg (2009) for a review and power study of many GoF tests available for copulas.

Estimation of the model is based on the numerical maximization of the log-likelihood. In the three-dimensional case, the log-likelihood for the full model is given by

$$\begin{aligned}\log \mathcal{L} &= \sum_{t=1}^{T} \log f(y_{1,t}, y_{2,t}, y_{3,t} \mid \mathcal{F}_{t-1}; \Theta) \\ &= \sum_{i=1}^{3} \sum_{t=1}^{T} \log f_i(y_{i,t} \mid \mathcal{F}_{t-1}; \Theta_i) + \sum_{t=1}^{T} \log c(u_{1,t}, u_{2,t}, u_{3,t} \mid \mathcal{F}_{t-1}; \Theta^c),\end{aligned} \tag{7}$$

where $\Theta = (\Theta_1, \Theta_2, \Theta_3, \Theta^c)$ denotes the parameters for the full model, and $\mathcal{F}_{t-1}$ is the filtration. Furthermore, the trivariate copula log-likelihood in Eq. (7) is decomposed as

$$\begin{aligned}\sum_{t=1}^{T} \log c(u_{1,t}, u_{2,t}, u_{3,t} \mid \mathcal{F}_{t-1}; \Theta^c) = &\sum_{t=1}^{T} \log c_{12}(u_{1,t}, u_{2,t} \mid \mathcal{F}_{t-1}; \Theta_1^c) \\ &+ \sum_{t=1}^{T} \log c_{13}(u_{1,t}, u_{3,t} \mid \mathcal{F}_{t-1}; \Theta_2^c) \\ &+ \sum_{t=1}^{T} \log c_{23\,|\,1}\{h(u_{2,t} \mid u_{1,t}; \Theta_1^c), h(u_{3,t} \mid u_{1,t}; \Theta_2^c) \mid \mathcal{F}_{t-1}; \Theta_3^c\},\end{aligned} \tag{8}$$

where $\Theta^c = (\Theta_1^c, \Theta_2^c, \Theta_3^c)$ denotes the parameters for the entire copula vine. As in most applications, we consider here a stepwise procedure, where the three marginal models are estimated independently of the copula. Estimation of the copula vine is additionally split up into two steps, as in Aas et al. (2009). First, we consider a sequential procedure where the parameters of each bivariate copula in the vine are estimated. Second, a joint maximization of the full copula log-likelihood is performed, using as start values the parameters obtained in the sequential estimation.

3.3 Empirical results

Marginal models

We start our empirical study by considering the individual behavior of our three variables. In the case of the wind power production time series, we are dealing with data that is bounded between 0 and 1, cf. Fig. 1(c). A standard time series model such as the ARMA model is clearly not suitable for this data, since nothing in the model ensures that we remain within the natural bounds when performing simulations. Therefore, we perform the logit transformation defined as $\Lambda(x) = \ln(x/(1-x))$ to the wind data in order to obtain a time series that takes values on the entire real line. Regarding the price data, the log transformation usually used in the literature when modeling electricity spot price data is unfeasible here, since we observe negative prices in both the DK1 and the DE price areas. As a result, we work with the raw electricity prices.

To model the seasonal component of the wind power production in DK1, we consider the function

$$f_t^{Wind} = a + c_1 \sin(2\pi t/365) + c_2 \cos(2\pi t/365), \tag{9}$$

where a is a constant, and c_1 and c_2 denote the coefficients for a yearly cycle. To model the seasonal component of the electricity prices in DK1 and DE, we consider the slightly more involved function

$$\begin{aligned} f_t^{Price} = a + bt + c_1 \sin(2\pi t/365) + c_2 \cos(2\pi t/365) \\ + d_1 \sin(4\pi t/365) + d_2 \cos(4\pi t/365) + \sum_{j=1}^{6} w_j W_t^j, \end{aligned} \tag{10}$$

where a again denotes a constant, b is the trend coefficient, c_1, c_2, d_1 and d_2 represent coefficients for the yearly and half-yearly cycles, respectively, and lastly w_j for $j = 1, \ldots, 6$ are coefficients corresponding to the day-of-week dummies denoted W^j. Parameter estimates obtained by fitting the proposed seasonal functions to the data are displayed in Table 1.

Next, we fit ARMA(p,q), ARMA(p,q) – GARCH(1,1) and ARMA(p,q) – GJR(1,1) models to the deseasonalized data, where we let $p = 0, \ldots, 5$ and $q = 0, \ldots, 5$. Also, we relax the normality assumption for the residuals, and allow for the more flexible skew t distribution instead, as in Patton (2013). The order of the ARMA model and the type of variance model are chosen based on the Bayesian information criterion (BIC). In Table 2, the specifications for the marginal models and the corresponding parameter estimates are displayed.

We also report p-values resulting from performing the Ljung-Box (LB) Q-test of serial independence on the standardized residuals, which we denote r, and the squared standardized residuals, r^2. The LB Q-tests all show

	Electricity price DK1	Electricity price DE	Logit wind DK1
$\hat{a}$	40.980 (0.640)	45.347 (0.597)	-1.110 (0.028)
$\hat{b}$	-0.010 (0.000)	-0.010 (0.000)	-
$\hat{c}_1$	-2.184 (0.287)	-3.195 (0.268)	0.068 (0.039)
$\hat{c}_2$	0.814 (0.286)	2.133 (0.267)	0.489 (0.039)
$\hat{d}_1$	-0.571 (0.285)	0.222 (0.266)	-
$\hat{d}_2$	-1.400 (0.286)	-1.173 (0.267)	-
$\hat{w}_1$	1.025 (0.754)	2.098 (0.704)	-
$\hat{w}_2$	1.350 (0.754)	1.869 (0.704)	-
$\hat{w}_3$	0.955 (0.754)	1.572 (0.704)	-
$\hat{w}_4$	-0.266 (0.754)	0.437 (0.704)	-
$\hat{w}_5$	-4.535 (0.754)	-6.373 (0.704)	-
$\hat{w}_6$	-7.364 (0.754)	-13.219 (0.704)	-

Table 1: OLS estimates for the parameters of the seasonal functions, with naive standard errors in parenthesis. Reference day-of-week is *Monday*.

strong evidence of no serial dependence left in the standardized residuals. Lastly, we report the p-values resulting from simulation-based Kolmogorov-Smirnov (KS) and Cramér-von Mises (CvM) goodness-of-fit tests. The tests are based on 999 bootstraps, and we examine the adequacy of two distributions for the standardized residuals, namely the normal distribution and the skew t (or skew normal) distribution.[1] The results show that the skew t (or skew normal) specification is suitable in all marginal models, whereas the normal distribution is rejected at a 5% significance level. Having verified the adequacy of the three proposed marginal models, we proceed to modeling the dependence structure with copulas.

Copula models

By applying the probability integral transform to the standardized residuals resulting from the marginal models, we obtain the approximately standard uniforms U_{DK1}, U_{DE}, and U_W, which are the input variables in our copula. The variable U_{DK1} refers to the price in DK1, U_{DE} refers to the price in DE, and lastly U_W refers to the wind power production in DK1. In Fig. 2, we plot the empirical copula densities for the three possible pairings of our variables.

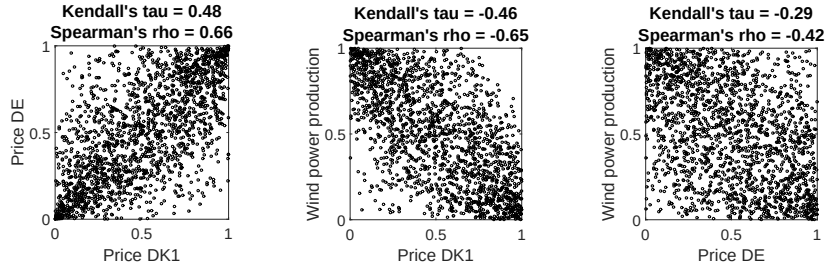
Since we find the variable with *most influence* (following the decomposition procedure in Czado et al. (2012)) to be the electricity price in DK1, we let this variable be the central node of the first tree. The pair-copula decomposition we obtain is illustrated in Fig. 3. We note that a C-vine coincides with a

[1]For the wind data, the skew t distribution converges towards the skew normal distribution since we obtain an estimate for the degrees of freedom above 300. Therefore, we replace the skew t with the skew normal distribution in this case.

	Electricity price DK1		**Electricity price DE**		**Logit wind DK1**	
Marginal model	ARMA(2,1) – GARCH(1,1)		ARMA(2,1) – GARCH(1,1)		ARMA(1,1)	
			Conditional mean			
AR1 $\hat{\phi}_1$	1.441 (0.045)		1.409 (0.080)		0.303 (0.043)	
AR2 $\hat{\phi}_2$	-0.460 (0.040)		-0.456 (0.063)		-	
MA1 $\hat{\theta}_1$	-0.837 (0.032)		-0.804 (0.070)		0.263 (0.042)	
Variance $\hat{\sigma^2}$	-		-		1.014 (0.036)	
			Conditional variance			
Constant $\hat{\omega}_G$	2.520 (0.364)		3.633 (0.637)		-	
ARCH $\hat{\alpha}_G$	0.142 (0.015)		0.120 (0.014)		-	
GARCH $\hat{\beta}_G$	0.795 (0.021)		0.772 (0.030)		-	
			Marginal distribution			
	Skew t		Skew t		Skew normal	
Shape $\hat{\nu}$	5.663 (0.833)		6.050 (0.929)		-	
Skew $\hat{\eta}$	-0.053 (0.026)		-0.174 (0.031)		-0.218 (0.047)	
			Testing model assumptions (p-values)			
LB Q-test (5 lags of r)	0.467		0.630		0.679	
LB Q-test (10 lags of r)	0.285		0.194		0.256	
LB Q-test (5 lags of r^2)	0.169		0.327		0.238	
LB Q-test (10 lags of r^2)	0.402		0.724		0.585	
	Normal	Skew t	Normal	Skew t	Normal	Skew Normal
KS test	0.004	0.225	0.001	0.828	0.011	0.426
CvM test	0.003	0.546	0.001	0.687	0.014	0.544

Table 2: Type and order of marginal models, parameter estimates with standard errors in parenthesis, and p-values.

D-vine in the three-dimensional case; hence, the *C* in C-vine is solely used to allude to the decomposition method.



(a) U_{DK1} vs. U_{DE} (b) U_{DK1} vs. U_W (c) U_{DE} vs. U_W

Fig. 2: Empirical copula densities obtained by applying the probability integral transform to the standardized residuals resulting from the marginal models for DK1 prices, DE prices and wind power production in DK1, respectively.

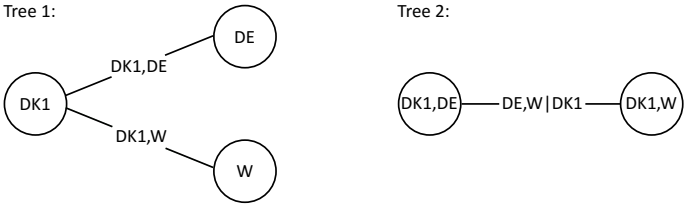


Fig. 3: Proposed C-vine structure for the electricity price in DK1, the electricity price in DE, and the wind power production in DK1.

According to Fig. 3, we must specify two bivariate copulas for the unconditional dependence between the pair U_{DK1}, U_{DE} and the pair U_{DK1}, U_W. Subsequently, we must specify one bivariate copula for the conditional dependence of U_{DE}, U_W given U_{DK1}. This is done in a recursive manner: First, we select the most suitable copulas for the pairs in the first tree, and then conditional on this selection, we move on to selecting a copula family that characterizes the conditional dependence in the second tree. For all pairs in the C-vine, we fit nine different bivariate copulas, and base selection on the AIC. The results are illustrated in Table 3, showing that the Student t copula is preferred for the dependence between prices in DK1 and DE, and that the Gaussian copula is preferred for the dependence between wind power production in DK1 and price in DK1. Given this selection and the associated estimated copula parameters, we apply the h-functions, cf. Eq. (6), to obtain the variables $U_{DE|DK1}$ and $U_{W|DK1}$. We then fit the nine different copulas to this pair and find that given prices in DK1, wind power production in DK1 and prices in DE are likely to be independent, and hence we choose the

independence copula for this pair.

We note that since only the Gaussian and the Student t copulas allow for negative dependence, a rotation was performed in order to fit the remaining seven copulas to the pair U_{DK1}, U_W in Table 3. Specifically, we fit these copulas to the pair U_{DK1}, $1 - U_W$ instead of the pair U_{DK1}, U_W. Regarding the standard errors for the estimates of the copula parameters stated in Table 3, these are simulation-based and take into account the estimation error coming from the ARMA–GARCH models, as in Patton (2013).

To test for the GoF of the selected copula families, we perform the CvM test, which is based on a comparison with the empirical copula. We obtain p-values of 0.584, 0.157 and 0.348 for the Student t copula, the Gaussian copula, and the independence copula, respectively, meaning that we cannot reject the null that the copula family is well-specified in neither one of the three cases. Having the independence copula describe the conditional dependence in the second tree simplifies our problem somewhat, since a joint estimation would yield the same results as the sequential estimation we have performed, and is therefore redundant.

Since the dimension of our problem is so small, one naturally wonders why not just use a trivariate copula instead of what might seem like an intricate decomposition of the problem. Also, is mixing different copula families really necessary? To provide some evidence for the benefit of the proposed mixed C-vine, we consider three other copula specifications: First, we consider the most standard case, i.e. fitting a trivariate Gaussian copula to our data. Second, we fit a trivariate Student t copula, and third, we fit a C-vine consisting of three Student t bivariate copulas, that is no mixing is allowed. The results are given in Table 4, and show that the mixed C-vine is the preferred copula model based on the AIC and the BIC.

Although the Student t C-vine has a slightly larger log-likelihood, the number of parameters in this model doubles compared to our proposed mixed C-vine. Also, estimation of this model is more involved, since a simultaneous estimation of the full copula log-likelihood is required. The trivariate Student t copula is less appropriate because it has only one degree of freedom. This is especially a drawback in our situation, where the dependency between the two prices exhibits significant tail dependence, whereas the dependence between each price and the wind power production exhibits no tail dependence, and thus pulling the one common degree of freedom in opposite directions. Consequently, although the difference in log-likelihoods between the proposed mixed C-vine and the trivariate Student t copula model is not striking, the latter would produce misleading tail dependence estimates which is especially problematic from a risk management perspective. Finally, the poorer fit of the trivariate Gaussian copula is due to similar arguments as in the case of the trivariate Student t, since the Gaussian copula has no tail dependence.

		Tree 1						Tree 2		
			$c_{DK1,DE}$			$c_{DK1,W}$			$c_{DE,W\|DK1}$	
Copula model		Param. (s.e.)	$\log \mathcal{L}_c$	AIC	Param. (s.e.)	$\log \mathcal{L}_c$	AIC	Param. (s.e.)	$\log \mathcal{L}_c$	AIC
Gaussian	$\hat{\rho}$	0.667	531.72	-1,061.44	-0.640	**476.58**	**-951.16**	0.018	0.29	1.42
		(0.013)			(0.014)			(0.025)		
Gumbel	$\hat{\theta}$	1.801	503.59	-1,005.18	1.675*	407.26	-812.52	1.012	0.67	0.66
		(0.040)			(0.036)			(0.010)		
Rot. Gumbel	$\hat{\theta}$	1.844	540.58	-1,079.16	1.706*	437.02	-872.04	1.016	1.09	-0.18
		(0.041)			(0.047)			(0.009)		
Clayton	$\hat{\theta}$	1.282	458.82	-915.64	1.065*	369.63	-737.26	0.039	1.44	-0.88
		(0.061)			(0.058)			(0.022)		
Rot. Clayton	$\hat{\theta}$	1.149	398.03	-794.06	0.971*	331.25	-660.50	0.003	0.01	1.98
		(0.068)			(0.055)			(0.024)		
Plackett	$\hat{\theta}$	9.809	532.08	-1,062.16	7.539*	456.95	-911.90	1.086	0.66	0.68
		(0.670)			(0.442)			(0.085)		
Frank	$\hat{\theta}$	5.355	511.43	-1,020.86	4.933*	472.05	-942.10	0.162	0.64	0.72
		(0.190)			(0.160)			(0.115)		
SJC	$\hat{\tau}^U$	0.433	546.57	-1,089.14	0.373*	426.14	-848.28	0.000	0.81	2.38
		(0.015)			(0.030)			(0.014)		
	$\hat{\tau}^L$	0.515			0.448*			0.001		
		(0.025)			(0.025)			(0.016)		
Student t	$\hat{\rho}$	0.677	**565.14**	**-1,126.28**	-0.640	476.58	-949.16	0.021	2.83	-1.66
		(0.015)			(0.017)			(0.025)		
	$\hat{\nu}$	5.671			379.173			20.704		
		(1.159)			(138.202)			(27.093)		

Table 3: Estimation results, where nine copula specifications are considered for each pair-dependence in the vine. The maximized value of the copula log-likelihood is denoted $\log \mathcal{L}_c$. The bold numbers in the first tree correspond to the optimal copulas, where selection is based on AIC. SJC is short for the Symmetrized Joe-Clayton copula introduced in Patton (2006). The Rot. Gumbel and Rot. Clayton refer to the rotated copulas. For the functional forms of the considered copulas and other characteristics, we refer to J. (1997) and Nelsen (1999). The asterisk refers to copulas fitted to rotated data, that is the pair U_{DK1}, $1 - U_W$ instead of the pair U_{DK1}, U_W. Standard errors are based on 999 simulations.

Mixed C-vine		Student t C-vine			3-dim. Student t		3-dim. Gaussian	
Param.	Final (s.e.)	Param.	Seq. (s.e.)	Final (s.e.)	Param.	Final (s.e.)	Param.	Final (s.e.)
$\rho_{DK1,DE}$	0.677 (0.015)	$\rho_{DK1,DE}$	0.677 (0.015)	0.677 (0.015)	$\rho_{DK1,DE}$	0.680 (0.013)	$\rho_{DK1,DE}$	0.667 (0.013)
$\nu_{DK1,DE}$	5.671 (1.159)	$\nu_{DK1,DE}$	5.671 (1.159)	5.634 (0.995)	$\rho_{DK1,W}$	-0.639 (0.013)	$\rho_{DK1,W}$	-0.640 (0.015)
$\rho_{DK1,W}$	-0.640 (0.014)	$\rho_{DK1,W}$	-0.640 (0.017)	-0.641 (0.014)	$\rho_{DE,W}$	-0.423 (0.020)	$\rho_{DE,W}$	-0.420 (0.019)
		$\nu_{DK1,W}$	379.173 (138.202)	379.173 (47.786)	ν	13.195 (3.735)		
		$\rho_{DE,W \mid DK1}$	0.021 (0.025)	0.021 (0.027)				
		$\nu_{DE,W \mid DK1}$	20.678 (27.125)	20.694 (24.754)				
log $\mathcal{L}_*$	**1,041.72**		1,044.55	1,044.57		1,028.65		1,008.43
AIC	**-2,077.44**		-2,077.10	-2,077.14		-2,065.30		-2,010.86
BIC	**-2,060.94**		-2,044.10	-2,044.14		-2,027.30		-1,994.36

Table 4: Comparison of the mixed C-vine copula with other competing models. Numbers in bold correspond to the preferred model, where selection is based on the AIC and the BIC. The maximized value of the full copula log-likelihood is denoted log $\mathcal{L}_*$. Standard errors are simulation-based (999 simulations), and take into account the estimation error coming from the marginal models.

To sum up, our proposed mixed C-vine allows for a lot of flexibility while only having three parameters, which coincides with the smallest alternative model. Furthermore, the model is easily estimated since a sequential procedure suffices.

Time-varying copula models

So far in our analysis, we have assumed that the dependence is static. However, since this is rarely the case in practice, a natural extension is to investigate whether or not time variation should be introduced in the copula model. Having established that the mixed C-vine copula is our preferred model cf. Table 4, we will restrict our attention to this model, and thus to introducing time variation in the Student t and Gaussian copulas, which are the selected specifications for the the unconditional dependence of the pair U_{DK1}, U_{DE} and the pair U_{DK1}, U_W, respectively.

Our approach to capturing time-varying dependence is through the Generalized Autoregressive Score (GAS) model of Creal et al. (2013). As the name suggests, the GAS model allows the time-varying copula parameter to evolve as a function of lagged values of the copula parameter and lagged values of the (scaled) score function of the copula log-likelihood. Assuming a copula with parameter ρ, the GAS specification we consider is:

$$h_{t+1} = \omega + \alpha h_t + \beta s_t I_t^{-1/2}, \tag{11}$$

where

$$h_t = g(\rho_t), \tag{12}$$

$$s_t = \frac{\partial}{\partial \rho} \log c((u_{1,t}, u_{2,t}); \rho_t), \tag{13}$$

$$I_t = \mathbb{E}_{t-1}[s_t^2]. \tag{14}$$

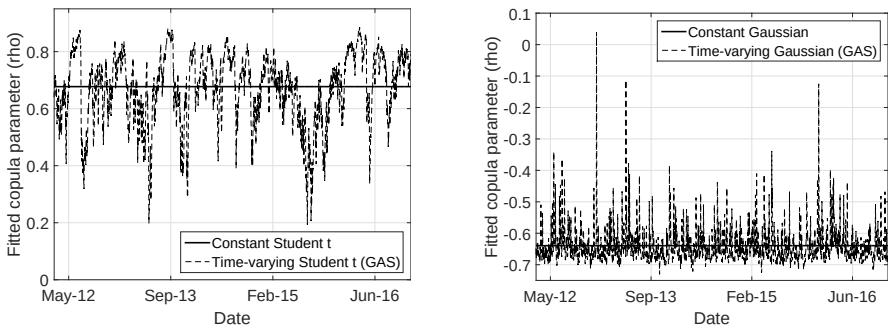
In Eqs. (11)–(14), h_t denotes the transformed copula parameter obtained by applying the tranformation $g(\cdot)$ to the copula parameter ρ_t, s_t denotes the score of the copula log-likelihood, and I_t is the Fisher information. For applications of the GAS model in the context of copulas, and also details on estimation and comparison with alternative models such as the ARMA-type processes employed in Patton (2006), we refer to e.g. Creal et al. (2013), Patton (2013), Avdulaj and Barunikl (2015) and Pircalabu et al. (2017).

Given our choice of a Student t and a Gaussian copula as elements of the vine, we model the correlation parameter of each of the two copulas according to the GAS equation. The degrees of freedom parameter of the Student t copula is kept constant. Further, since correlation is restricted to lie in the interval $(-1, 1)$, we let $h_t = \log(\rho_t + 1) - \log(1 - \rho_t)$ to ensure this.

The estimated parameters for the GAS models are reported in Table 5, pointing to evidence of time-varying dependence in both models. To provide a visual indication of how the correlations change with the time, we plot in Fig. 4 the dynamics of the correlation parameter estimates implied by the two GAS models.

	Tree 1				**Tree 2**	
	Student t GAS		Gaussian GAS		Independence	
	$c_{DK1,DE}$		$c_{DK1,W}$		$c_{DE,W\|DK1}$	
	Param.	s.e.	Param.	s.e.	Param.	s.e.
$\hat{\omega}$	0.057	(0.027)	-0.696	(0.301)	-	-
$\hat{\alpha}$	0.967	(0.016)	0.545	(0.196)	-	-
$\hat{\beta}$	0.120	(0.026)	0.142	(0.044)	-	-
$\hat{\nu}$	6.711	(2.274)	-	-	-	-
$\log \mathcal{L}_c^{tv}$	612.67		485.04		-	
AIC	-1,217.34		-964.08		-	

Table 5: Estimation results for time-varying copulas in the mixed C-vine, where the linear correlations evolve according to the GAS specification. Standard errors are simulation-based. The maximized value of the copula log-likelihood is denoted $\log \mathcal{L}_c^{tv}$.



(a) Pair $\hat{U}_{DK1}$, $\hat{U}_{DE}$ (b) Pair $\hat{U}_{DK1}$, $\hat{U}_{W}$

Fig. 4: Linear correlation implied by the constant copula models and the GAS models.

Having introduced time-varying copulas in the first tree of the vine, we note that this can affect the choice of copula in the second tree. To check that the independence copula remains a good specification for the conditional dependence in the second tree, we compute new variables $U_{DE|DK1}$ and $U_{W|DK1}$ based on the h-functions that now take as input the time-varying copula parameters ρ_t implied by the fitted GAS models. Performing the CvM test on the new conditional pair, we obtain a p-value of 0.239, confirming the goodness-of-fit of the independence copula. Basing the model selection on values of the AIC, both time-varying models yield lower values than the cor-

responding constant models, implying that the time-varying mixed C-vine copula model is preferred to the constant one. Hence, we will keep this time-varying model as our preferred specification and use it in the next section as a simulation tool based on which hedging decisions are made.

4 Hedging joint price and volumetric risk

In this section, we address the hedging of joint price and volumetric risk in the context of an energy trading company entering into longer-term contracts with wind power producers, where the future production is bought at a pre-determined fixed price; we shall refer to such contracts as fixed-for-fluctuating agreements.

First, let us introduce some notation and simplifying assumptions. We will assume that at time t_0, a hypothetical energy trading company enters into fixed-for-fluctuating agreements in DK1, corresponding to an installed capacity of c MW. The company offers a common fixed price $S_{t_0}^{fixed}$ per MWh to all wind power producers, and we consider a delivery period from T_1 to T_2. Furthermore, we will assume that the total wind power production in DK1 is perfectly correlated to the total production associated with the agreements entered into by the company. Since we model the total daily wind power production in DK1 relative to the total installed capacity, which we denote Q_t^{DK1}, we simply multiply Q_t^{DK1} by $24c$ to get the total daily production measured in MWh that the company has under management. Under these conditions, the total revenue R^f of the energy trading company over the period $[T_1, T_2]$ can be expressed as

$$R^f = 24c \sum_{t=T_1}^{T_2} Q_t^{DK1}(S_t^{DK1} - S_{t_0}^{fixed}), \quad \text{for } t_0 < T_1 < T_2. \tag{15}$$

where S_t^{DK1} denotes the average spot electricity price in DK1 at day t, and we assume that all days have 24 hours. Moreover, according to Eq. (15), we assume that the delivered wind power production is sold on the day-ahead market, and ignore possible balancing costs.

To construct a hedging portfolio, we first need to identify the available financial instruments in the market. Since we are attempting to hedge joint price and volumetric risk in the DK1 market, the selection of hedging instruments to choose from is small, and in addition, the market place is characterized as being rather illiquid. To overcome these limitations, we include hedging instruments with reference to the DE spot electricity price, and use them to construct a proxy hedge. This extension is believed to be advantageous due to the strong positive correlation between the spot electricity prices in DK1 and DE, which we have demonstrated earlier.

Here, we shall concentrate on a static hedge that is performed by the energy trading company at time t_0, i.e. the same time as entering the fixed-for-fluctuating agreements. Due to market incompleteness, we are clearly not attempting to replicate the payoff of fixed-for-fluctuating agreements exactly, and thus to attain a perfect hedge.

The hedging instruments we allow are as follows:

1. *Power forwards and futures*

   Power forwards and futures are the most widely used hedging instruments in electricity markets. Depending on the market place, they can have different characteristics, e.g. physical versus financial settlement, delivery during the base, peak or off-peak hours, and delivery periods corresponding to a year, season, quarter, month, etc. Here, we will allow trading these types of contracts with reference to both the DE and the DK1 spot electricity price. Also, we will assume a deterministic interest rate, thus equating forwards and futures. When referring to such contracts on the DE market, we will denote them DE futures, and when referring to them in DK1, we will use the term DK1 forwards. DE futures are among the most liquid contracts traded on the European Energy Exchange (EEX). Trading in DK1 forwards is somewhat more cumbersome, since two positions must be entered into: a futures contract on the Nordic spot system price and an electricity price area differential (EPAD), which is a forward contract with reference to the difference between the DK1 area price and the Nordic system price.

   For a contracted quantity of 1 MW, a long position in a DE futures contract yields the following revenue, denoted by $R^{F,DE}$:

   $$R^{F,DE} = 24 \sum_{t=T_1}^{T_2} S_t^{DE} - F_{t_0}^{DE}, \quad \text{for } t_0 < T_1 < T_2, \tag{16}$$

   where $F_{t_0}^{DE}$ denotes the DE futures price at time t_0, with delivery period from T_1 to T_2, and S_t^{DE} denotes the average DE spot electricity price at day t. Since we work on a daily basis, we implicitly assume base load delivery, and the contracted quantity of 1 MW entails delivering 24 MWh at each day $t \in [T_1, T_2]$. To obtain the revenue associated with a DK1 forward contract, which we denote $R^{F,DK1}$, DE is simply replaced by DK1 in Eq. (16), with the obvious changes in the interpretation of variables.

2. *Wind index futures*

   Wind index futures are at the present time the only standardized exchange-traded instruments with the underlying being the actual wind

index. For a contracted quantity of 1 lot, their revenue denoted by R^{WIF} can be expressed as

$$R^{WIF} = 24(T_2 - T_1 + 1)(I^{Actual} - I_{t_0}^{Locked}) \cdot 100 \text{ EUR}, \quad \text{for } t_0 < T_1 < T_2, \quad (17)$$

where I^{Actual} denotes the actual average wind index for the contracted period, which is obtained by dividing the total actual wind power production by the total available capacity and the number of hours in the delivery period, i.e. $0 \leq I^{Actual} \leq 1$. Furthermore, $I_{t_0}^{Locked}$ is the locked-in level set at time t_0. This level can be thought of as the market's expectation of the future average wind index corresponding to the delivery period.[2]

Wind index futures were launched late 2015 by NASDAQ OMX Commodities Europe and late 2016 by EEX. They can currently be traded on the German wind power index, and according to EEX (2014), the next step is to expand this product to the Danish and the UK market. Since we model the wind power production in DK1, we will allow for the inclusion of DK1 wind index futures in our hedging portfolio based on this outlook.

3. *Financial transmission rights (spread options)*

A financial transmission right (FTR) is an option written on the hourly spot electricity price difference in two price areas that are interconnected. It gives the holder the right to "transfer" electricity from e.g. DK1 to DE (or vice versa) whenever the price difference $S_h^{DE} - S_h^{DK1}$ (or $S_h^{DK1} - S_h^{DE}$) is positive for hour h. In reality, no physical delivery of electricity is made and only the price spread, if positive, is paid out to the holder of the option; hence the name FTR. Since we work with aggregated data on a daily basis, we shall consider the daily price spread instead of the hourly price spreads. If we denote by $P^{DK1 \to DE}$ the payoff associated with the right to transfer electricity from DK1 to DE each day t in the delivery period from T_1 to T_2, we have a sum of daily spread options represented by the payoff

$$P^{DK1 \to DE} = 24 \sum_{t=T_1}^{T_2} \left(S_t^{DE} - S_t^{DK1} \right)^+, \quad \text{for } t_0 < T_1 < T_2,$$

where we assume a contracted quantity of 1 MW. The revenue resulting from a 1 MW long position in such daily spread options is

$$R^{DK1 \to DE} = P^{DK1 \to DE} - V_{t_0}^{DK1 \to DE}, \quad (18)$$

[2]We note that our product description of wind index futures is based on NASDAQ OMX (2015) and EEX (2014).

where $V_{t_0}^{DK1\to DE}$ denotes the sum of fair option values at time t_0. Similarly, $R^{DE\to DK1}$ is obtained by considering the opposite spot price spread.

4. *Electric boilers (out-of-the-money put options)*

   In DK1, there are a number of combined heat and power (CHP) plants equipped with electric boilers. These plants can typically produce heat in one of three ways: 1) by the use of gas boilers, 2) by the use of electric boilers, where electricity is converted to heat, and 3) by using gas as an input to produce electricity, with heat being a by-product. We do not go into details with the optimization problem that CHP plants face every day in order to minimize their customers' heat expenses. We note however that when profitable, say when the price of electricity is at most K, plant owners will use electricity to produce heat. So, if we could sell a fixed amount of electricity to CHP plants at the price K each time the daily spot goes below K, we would be entering a string of daily put options. Such products can in fact be traded OTC in DK1, at a strike K which is usually significantly below the average electricity spot price in DK1, translating into the fact that the daily puts are out-of-the-money (OTM). The benefit of such instruments is clear, as they reduce the risk associated with high wind and low price scenarios by providing a price floor. Nonetheless, they are mostly available/useful during winter periods.[3] For a contracted quantity of 1 MW, their payoff P^{Put} is given by

$$P^{Put} = 24 \sum_{t=T_1}^{T_2} \left(K - S_t^{DK1}\right)^+, \quad \text{for } t_0 < T_1 < T_2,$$

   where the strike K is kept constant throughout the delivery period. Like in the case of FTRs, we consider here a sum of daily options. The revenue associated with a 1 MW long position is given by

$$R^{Put} = P^{Put} - V_{t_0}^{Put}, \tag{19}$$

   where $V_{t_0}^{Put}$ denotes the sum of fair option values at time t_0.

For a portfolio consisting of fixed-for-fluctuating agreements as well as all hedging instruments described above, the associated total revenue for the delivery period from T_1 to T_2 can be expressed as

$$R^{Total} = R^f + \sum_{n=1}^{N} \theta_n R^{(n)}, \tag{20}$$

[3] Due to lower demand for heat during summer, CHP plants have a lower capacity. Moreover, high wind / low price scenarios are less likely.

where the $R^{(n)}$'s represent the revenues corresponding to each of the hedging instruments (Eqs. (16) - (19)), and the θ_n's denote the corresponding contracted quantities. Specifically, $N = 6$ in Eq. (20), with the following correspondence: DK1 forwards ($n = 1$), DE futures ($n = 2$), wind index futures ($n = 3$), FTRs in the direction DE→DK1 ($n = 4$), FTRs in the direction DK1→DE ($n = 5$), and lastly OTM puts ($n = 6$).

To find the fair value of the financial contracts in the portfolio at t_0, we assume a zero interest rate, and let the risk neutral measure equal the physical measure, i.e. $\mathbb{Q} = \mathbb{P}$, as in Coulon et al. (2013). With these simplifying assumptions, all contracts in the portfolio can be priced by performing simulations from our proposed C-vine copula model. Specifically, the fair value of $S_{t_0}^{fixed}$ can be computed by the usual practice of setting the risk neutral expectation of the revenue R^f given in Eq. (15) equal to zero, which yields

$$S_{t_0}^{fixed} = \frac{\mathbb{E}_{t_0}^{\mathbb{Q}}\left[\sum_{t=T_1}^{T_2} Q_t^{DK1} S_t^{DK1}\right]}{\mathbb{E}_{t_0}^{\mathbb{Q}}\left[\sum_{t=T_1}^{T_2} Q_t^{DK1}\right]}. \tag{21}$$

The same approach is used to obtain forward and futures prices, which we assume are unbiased estimators of the future spot electricity prices. The expected wind index needed to obtain the locked in level in wind index futures is obtained by averaging across simulated paths of future wind power production corresponding to the delivery period. Lastly, fair option values are obtained by computing the risk neutral expectations of simulated payoffs. The assumption of $\mathbb{Q} = \mathbb{P}$ is of course unrealistic, but we do not go into details here and refer instead to Burger et al. (2004), Benth et al. (2008), and Kolos and Ronn (2008) for comprehensive discussions on the matter. It simplifies nonetheless our analysis a great deal, since the mean of the total revenue distribution is zero per construction, and is unaffected by varying hedge quantities. As a result, we are left with the variance aspect of the problem. To obtain the optimal static hedge, we consider the traditional variance-minimizing criterion

$$\min_{\boldsymbol{\theta}} \mathbb{V}\mathrm{ar}_{t_0}\left[R^f + \sum_{n=1}^{N} \theta_n R^{(n)}\right] \tag{22}$$

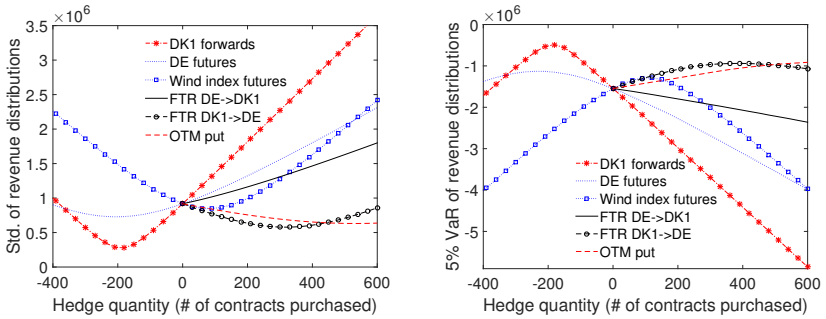
$$\text{s.t. } \theta_n \geq 0,\ n = 4, \ldots, 6, \tag{23}$$

where short-selling constraints are introduced regarding the options. This restriction is added due to practical reasons, since the hedger, i.e. the energy trading company, does not own the DK1–DE interconnector. Moreover, we assume that the hedger does not own CHP plants.

4.1 Hedging results

In order to exemplify the benefits of the proposed hedging instruments, we consider an energy trading company entering into monthly fixed-for-fluctuating agreements at $t_0 = 12$ Dec 2016, which corresponds to the last date in our sample. As delivery we consider the out-of-sample period from $T_1 = 1$ Jan 2017 to $T_2 = 31$ Jan 2017. Hedging is also performed at t_0, and all hedging instruments have the same delivery period as the fixed-for-fluctuating agreements. We let the installed capacity $c = 500$ MW and the daily strike of the OTM put options $K = 12$ EUR/MWh.

Before considering the minimization problem stated in Eqs. (22) – (23), we study the individual effectiveness of the proposed hedging instruments. First, we compute the fair value of the fixed-for-fluctuating agreements and all the hedging instruments at t_0, based on 100,000 simulations from the time-varying mixed C-vine copula model.[4] Then, we compute revenues for each simulated path, where we vary the hedge quantities θ_n for each instrument, $n = 1, \ldots, N$. The results are displayed in Fig. 5, where we consider as measures of risk the standard deviation of the revenue distributions in Fig. 5(a) and the 5% Value-at-Risk (VaR) of the revenue distributions in Fig. 5(b).



(a) Effect on the standard deviation

(b) Effect on the 5% VaR

Fig. 5: Isolated effect of different hedging strategies.

Not surprisingly, we observe a large variation in the hedging benefits of the proposed instruments. The DK1 forwards are the most effective, as they secure the revenue associated with a significant part of the production in advance. Then, only the remaining production - the difference between the actual and the hedged volume - will be exposed to spot price risk. The DE futures are less effective at reducing risk, which is expected since they are settled against the DE and not the DK1 spot prices. Somehow in contrast to

[4]Sampling from a copula vine is based on a recursive procedure and using the inverse of the h-functions, cf. Eq. 6. We refer to Aas et al. (2009) for a detailed description of the procedure.

DK1 forwards, the wind index futures fix a revenue associated with the difference between actual and expected production. Hence, this fixed revenue is associated with a much smaller volume than in the case of the DK1 forwards, which explains the inferiority of wind index futures at reducing risk.

When considering the options, we observe that the OTM puts are rather effective, owing to the fact that they provide a floor to the DK1 spot prices, thus reducing the risk related to a low price / high volume scenario. Lastly, the FTRs in the direction DK1 to DE are also effective, and are preferable to the opposite direction. The difference in hedging benefits between the FTRs relates mostly to the short selling constraint.

Next, we turn our attention to the optimal static hedge when allowing trading in all instruments simultaneously. In this context, we note that the classical parity argument (see Carmona and Durrleman (2003)) implies that an FTR in one particular direction can be written as a linear combination of an FTR in the opposite direction, a DK1 forward and a DE futures. Consequently, allowing trading in all these four instruments would result in the variance minimization problem in Eqs. (22) – (23) having many possible optimal solutions. To avoid this, we remove the FTRs in the direction DK1$\rightarrow$DE from the hedging portfolio, and present below one of many optimal solutions.

In Fig. 6(a) we plot three different simulated revenue distributions for comparison purposes. The blue dashed line displays the simulated monthly revenue distribution associated with the fixed-for-fluctuating agreements before performing a hedge. The black solid line represents the distribution after performing an optimal hedge using DK1 forwards only, and lastly the red dashed-dotted line corresponds to the distribution after performing an optimal hedge based on all proposed hedging instruments simultaneously. Fig. 6(b) provides a zoom of the left tails of the simulated distributions, and Fig. 6(c) specifies the optimal hedging strategy associated with the latter revenue distribution, and is obtained by solving Eqs. (22) – (23).

Figs. 6(a) and 6(b) clearly show that significant benefits are to be drawn by using all proposed hedging instruments. In Table 6, we provide the reduction in the variance and the 5% VaR of the monthly revenue distribution achieved by 1) using DK1 forwards as opposed to no hedge, and 2) using all instruments as opposed to using only DK1 forwards in the hedging portfolio. In both cases, significant reductions are attained. However, when examining the isolated effect of each hedging instrument in combination with the DK1 forwards, we see that it is mostly the OTM puts that reduce risk additionally.

Our results so far indicate that DK1 forwards and OTM puts are the most powerful, and that including additional instruments to the hedging portfolio does not contribute substantially to the variance reduction of the revenue distribution. Since neither of these two instruments depends on the DE price, the benefit of including this variable in our copula vine seems minimal. To

Initial hedging portfolio No hedge						
			Isolated effect			**Combined effect**
			DK1 forwards			
Reduction in variance of revenue dist.			90.36%			-
Reduction in 5% VaR of revenue dist.			68.23%			-
Initial hedging portfolio DK1 forwards only						
			Isolated effect			**Combined effect**
	DE futures	Wind index futures	FTR DE→DK1	FTR DK1→DE	OTM put	
Reduction in variance of revenue dist.	0.05%	1.40%	14.15%	11.62%	43.24%	46.72%
Reduction in 5% VaR of revenue dist.	0.20%	0.42%	9.60%	7.22%	30.42%	31.88%

Table 6: The upper panel displays the effects of hedging the joint price and volumetric risk implied by fixed-for-fluctuating agreements with DK1 forwards only, as opposed to no hedge at all. The lower panel displays the effects of adding other instruments to a hedging portfolio initially consisting of DK1 forwards.

highlight the importance of hedging instruments written on the DE price, we regard the optimal hedge quantity of DK1 forwards, which is approx. -190 MW cf. Fig. 6(c). As mentioned earlier, the DK1 market place is characterized by illiquidity, and so being able to enter a short position corresponding to e.g. -190 MW DK1 forwards at t_0 would imply a significant cost due to the rather large difference between the bid price and the mid-point of the bid-ask spread. We consider therefore including such a cost in our revenue equation, and in consequence minimize the total revenue variance and the expected cost of the replicating portfolio subject to an arbitrage free constraint and the same short-selling constraints as before, i.e.

$$\min_{S_{t_0}^{fixed},\boldsymbol{\theta}} \left\{ \mathbb{V}\mathrm{ar}_{t_0}\left[R^f + \sum_{n=1}^{N} \theta_n R^{(n)} - 24\xi \left|\theta_{n=1}\right| \right] + \mathbb{E}_{t_0}\left[\sum_{n=1}^{N} \theta_n R^{(n)} + 24\xi \left|\theta_{n=1}\right| \right] \right\} \quad (24)$$

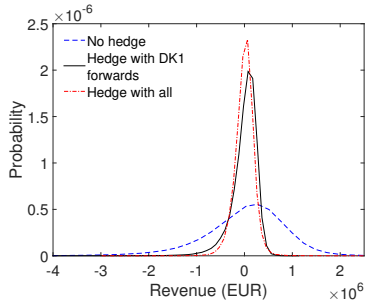
$$\text{s.t. } \mathbb{E}_{t_0}\left[R^f + \sum_{n=1}^{N} \theta_n R^{(n)} - 24\xi \left|\theta_{n=1}\right| \right] = 0 \quad (25)$$

$$\theta_n \geq 0,\ n = 4, \ldots, 6, \quad (26)$$

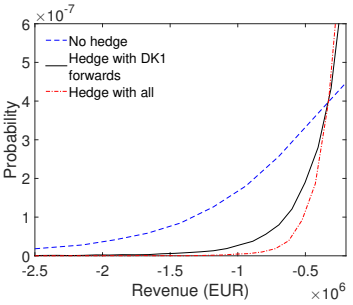
where ξ denotes the cost in EUR/MWh associated with a position of $\theta_{n=1}$ MW in DK1 forwards. We fix $\xi = 1$ EUR/MWh, which is a realistic choice considering recent numbers for volumes traded and bid-ask spreads associated with DK1 forwards, and solve the constrained minimization problem posed in Eqs. (24) – (26). We include FTRs in both directions here, since the indifference implied by the parity argument from before no longer holds. The solution is displayed in Fig. 6(d), and not surprisingly, the hedging strategy shifts towards less DK1 forwards. Still, a significant amount of DK1 forwards is required, which is due to the inability to short FTRs in the direction DE→DK1. The same degree of variance reduction is obtained with this alternative hedging strategy; we note however that in order to satisfy the no-arbitrage constraint and thus keep the expected revenue at zero, the price $S_{t_0}^{fixed}$ is re-estimated, with the solution yielding a lower price than under the assumption of a frictionless market.

4.2 An example highlighting the benefits of wind index futures

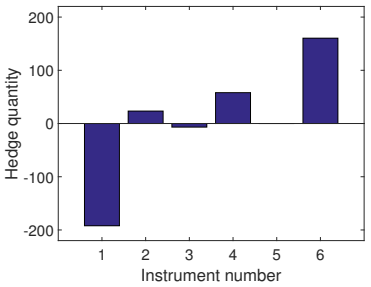
Since the examples in section 4.1 have to a limited extent brought out the usefulness of the wind index futures, one might think of them as inferior products compared to all other proposed hedging instruments. Although this is true in the case where a simultaneous hedge is performed at t_0, when also entering into fixed-for-fluctuating agreements, there are other scenarios that can better highlight their benefits. Let us consider the following example:



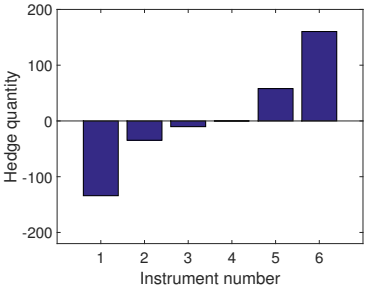
(a) Simulated monthly revenue distributions



(b) Zoom of left tail in figure 6(a)



(c) Optimal hedging strategy



(d) Optimal hedging strategy when including costs associated with trading in DK1 forwards

Fig. 6: Hedging results under different assumptions. The hedging instruments are numbered as 1) DK1 forwards, 2) DE futures, 3) Wind index futures, 4) FTRs DE $\rightarrow$ DK1, 5) FTRs DK1 $\rightarrow$ DE, and 6) OTM put.

- At some time t^* before t_0 = 12 Dec 2016, an energy trading company enters into fixed-for-fluctuating agreements with wind power producers.
- Delivery period for the agreements is the same as before, i.e. T_1 = 1 Jan 2017 and T_2 = 31 Jan 2017, and the installed capacity c remains fixed to 500 MW. We again here assume perfect correlation between the wind power production that the company buys and the total wind power production in the price area.
- Also at time t^*, the company performs a hedge corresponding to a short position in DK1 forwards, where the contracted volume equals the company's expectation of the total future production for January 2017. We assume that the expected wind index is 0.38, and thus the hedge quantity equals $-0.38 \cdot 500$ MW.

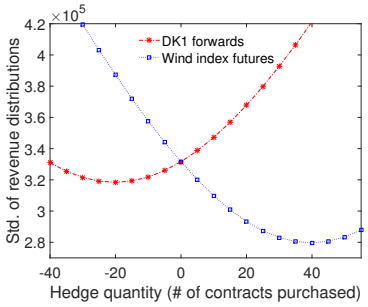
- We fix $S_{t^*}^{fixed} = 33$ EUR/MWh and $F_{t^*}^{DK1} = 37$ EUR/MWh.
- As time goes by and we reach t_0, the price level has dropped. This is indeed reflected in our simulations, where we estimate lower values for S^{fixed} and F^{DK1} at t_0.
- At t_0, the trading company wishes to rebalance their hedging portfolio, and they consider whether to use additional DK1 forwards and/or wind index futures.

By performing similar calculations as in the previous examples, we obtain the standard deviations of revenue distributions under different rebalancing strategies consisting of trading in DK1 forwards and wind index futures, respectively. The results are summarized in Fig. 7(a), showing that wind index futures are more effective than DK1 forwards. This can be explained by considering e.g. a high wind / low price scenario, which causes the realized production to exceed the company's expectation. The excess production will have to be sold on the day-ahead market, and since prices have fallen substantially since t^*, the spot price will be below the fixed price that the company promised to pay wind power producers. In this situation, a long position in wind index futures will yield a premium corresponding to precisely the excess production, independent of the current price level cf. Eq. (17).

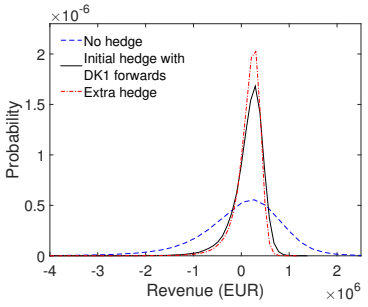
We complement these results with Figs. 7(b) and 7(c), where we illustrate the revenue distribution and a zoom of the left tail before the hedge (blue dashed line), after an optimal hedge in DK1 forwards at t^* (black solid line), and after rebalancing with DK1 forwards and wind index futures at t_0 (red dashed-doted line). As expected, the revenue distribution is shrunk after the extra hedge, compared to the revenue distribution obtained after the initial hedge in DK1 forwards. According to Table 7, the benefit is measured to be a reduction in the variance of the revenue distribution of approx. 29%. When considering the isolated effects, Table 7 reveals that an extra hedge based on wind index futures alone achieves almost the same risk reduction. So given wind index futures in the extra hedge, adding DK1 forwards yields a minimal benefit in this example.

	Isolated effect		**Combined effect**
	Wind index futures	DK1 forwards	
Reduction in variance of revenue dist.	28.28%	7.80%	28.92%
Reduction in 5% VaR of revenue dist.	22.25%	8.67%	22.31%

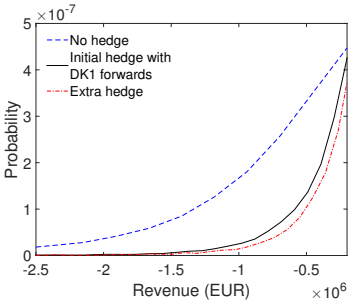
Table 7: Effects of rebalancing the hedging portfolio due to decreasing prices.



(a) Effect on the standard deviation



(b) Simulated monthly revenue distributions



(c) Zoom of left tail in figure 7(b)

Fig. 7: Illustrating the usefulness of wind index futures.

4.3 Comparison with alternative models

So far in our hedging analysis, we have concentrated on applications where the copula model remains fixed to our preferred specification, which is the time-varying mixed C-vine copula model. In this section, we wish to compare this model against "suboptimal" alternatives, and highlight the importance of capturing characteristics such as heavy tails in the marginal distributions, tail dependence, and time variation in the copula. To this end, we consider four models, which in the interest of clarity are stated in Table 8.

	Model 1	Model 2	Model 3	Model 4
Marginal distributions	Normal	Skew t (skew n.)	Skew t (skew n.)	Skew t (skew n.)
Copula model	Trivar. Gaussian	Trivar. Gaussian	Mixed C-vine	TV mixed C-vine
Heavy tails (margins)	×	✓	✓	✓
Tail dependence	×	×	✓	✓
TV dependence	×	×	×	✓

Table 8: Alternative models for comparison study, and their characteristics. TV is short for time-varying.

Given the same contractual specifications as in section 4.1, we perform out-of-sample simulations from all four models in Table 8. Then, we solve for the optimal hedging strategy when 1) allowing for only DK1 forwards in the hedging portfolio, and 2) allowing for all proposed hedging instruments in the hedging portfolio. Finally, we compute reductions in the variance and 5% VaR of the revenue distributions obtained when hedging with all proposed hedging instruments as opposed to using only DK1 forwards. Also, we compute the values in EUR corresponding to the 5% VaR and the 5% Expected Shortfall (ES) of the revenue distributions obtained by including all hedging instruments. To avoid the effect of simulation errors, the same random seed is used for simulations from the four models. The results are given in Table 9.

Let us start by considering the first two rows of Table 9 and discuss the differences in the reductions. When regarding the difference between Model 1 and 2, we see that the introduction of heavy tails implies an increase in the hedging benefits by yielding higher values for the reduction in variance and 5% VaR. The introduction of tail dependence between the DK1 and DE prices (i.e. the difference between Model 2 and 3) lowers the hedging benefits, which is expected since tail dependence lowers the benefits of the FTRs in the sense that a positive probability of extreme prices occurring simultaneously lowers the overall potential payout. Lastly, the introduction of time-varying dependence (i.e. the difference between Model 3 and 4) lowers further the hedging benefits. While tail dependence is in this case always expected to entail lower hedging benefits, the effects of time-varying vs. static dependence can go both ways, depending on the difference between the levels of dependence implied by the two models. In our example, the lower hedging benefits produced with Model 4 are caused by the stronger dependence between the DK1 and DE prices implied by the GAS model. This stronger dependence weakens once again the usefulness of the FTRs, thereby contributing negatively to the overall benefits.

Moving on to the last two rows of Table 9, we notice that the values of the 5% VaR and 5% ES increase when considering the models in decreasing order: Our preferred model (Model 4) implies most risk, followed by Model 3 and so forth.

Since the prices of the fixed-for-fluctuating agreements and all hedging instruments have already been computed to perform our hedging exercise above, we also quantify in Table 10 the effects of the different models from a pricing perspective. Cf. Table 10, the introduction of heavy tails in the margins (measured by Δ_1) has a mixed but noteworthy effect on most instrument prices. The largest absolute impact is on the price of the out-of-the-money puts, which is not surprising, since out-of-the-money options are much more affected by tail behavior than at-the-money or in-the-money options.

Regarding Δ_2 and Δ_3, most interesting to consider are the FTRs and the

Initial hedging portfolio DK1 forwards only	Model 1	Model 2	Model 3	Model 4
Reduction in variance of revenue dist.	40.49%	49.25%	48.12%	46.72%
Reduction in 5% VaR of revenue dist.	29.03%	33.15%	32.70%	31.88%
Value of the 5% VaR (EUR)	-321,440	-326,840	-328,410	-334,890
Value of the 5% ES (EUR)	-440,690	-462,200	-465,720	-476,940

Table 9: Combined effects of adding the other proposed instruments to a hedging portfolio initially consisting of DK1 forwards. The results are obtained with four different models. The values in EUR of the 5% VaR and 5% ES correspond to the revenue distributions obtained by hedging with all proposed instruments.

fixed-for-fluctuating agreements. This is because their payoffs depend on two underlyings, and hence the choice of copula can influence the prices of these instruments to a large extent, as opposed to all other instruments. For the FTRs, the introduction of tail dependence, followed by time-varying dependence, have both a negative and significant effect on prices. The overall reduction in prices from Model 1 to Model 4 reaches around 11% for the direction DE→DK1, and around 9% for the direction DK1→DE. For the fixed-for-fluctuating agreements, the fact that Δ_2 is zero is not surprising, since tail dependence is introduced only between the DK1 and DE prices, and not the DK1 price and wind power production. As a result, the transition from a trivariate Gaussian to a mixed C-vine copula model has no implications for the price of this instrument. Δ_3 being zero for S^{fixed} can be explained by the fact that the mean value of the time-varying dependence implied by the GAS model for the DK1 price and wind power production is very close to the dependence implied by the static model.

Estimated prices
Pricing date (t_0): 12 December 2016
Delivery period: January 2017

	Model 1	Model 2	Model 3	Model 4	Δ_1	Δ_2	Δ_3
S^{fixed}	23.661	23.736	23.736	23.732	**0.32%**	0.00%	0.00%
F^{DK1}	26.322	26.324	26.324	26.324	0.00%	0.00%	0.00%
F^{DE}	26.931	26.935	26.936	26.937	0.00%	0.00%	0.00%
I^{Locked}	0.378	0.382	0.382	0.382	**1.06%**	0.00%	0.00%
$V^{DE\to DK1}$	2.919	2.859	2.775	2.585	**-2.06%**	**-2.94%**	**-6.85%**
$V^{DK1\to DE}$	3.529	3.471	3.386	3.198	**-1.64%**	**-2.45%**	**-5.88%**
V^{Put}	0.519	0.548	0.548	0.548	**5.59%**	0.00%	0.00%

Table 10: Prices computed with different models. All prices are given in EUR/MWh with the exception of I^{Locked}, which represents an index. Δ_1, Δ_2 and Δ_3 refer to the relative differences between Models 1 and 2, Models 2 and 3, and Models 3 and 4, respectively, and are highlighted if different from zero.

As a last remark, we investigate the effects of using "suboptimal" models for hedging. Specifically, we let pricing and hedging decisions be based on one of Model 1, 2, or 3, while afterward assuming that the *real* prices

and wind power production evolve according to Model 4. That is, based on the prices and the hedging strategies obtained earlier with the more naive models, we construct revenue distributions where the simulated paths of our three variables come from our preferred model. Then, we ask the following question: By how much do we under- or overestimate risk?

The results are given in Table 11, and we report the 5% VaR, the 5% ES and the expected value of the revenue distribution in two cases: (a) using the same model to compute instrument prices and an optimal hedging strategy, and also to construct a revenue distribution (as in Table 9); (b) using Models 1–3 to compute instrument prices and an optimal hedging strategy, while constructing the revenue distribution by using simulations from Model 4. The columns denoted Δ state the relative differences between results obtained in (a) and (b), showing that basing decisions on the "wrong" model leads to an underestimation of risk in all three cases. Furthermore, we note that the expected value of the revenue distribution is no longer zero in the case of columns (b). If we based pricing and hedging decisions on Model 1, the expected revenue would actually be positive, implying a gain on average. Despite the suboptimal hedging strategy and the overestimated FTR prices, which all push the revenue distribution to the left, the decisive factor that ends up shifting the revenue distribution to the right is the price of fixed-for-fluctuating agreements, which is too low under Model 1. If we based pricing and hedging decisions on Model 2 or 3, we would expect a loss on average. This is due to the hedging strategy implied by Model 2 and 3, respectively, which pulls the revenue distribution to the left. Adding to this effect are the FTR prices, which continue to be overestimated in both Model 2 and 3.

	Model 1			Model 2		
	(a)	(b)	Δ	(a)	(b)	Δ
5% VaR	-321,440	-325,150	-1.14%	-326,840	-351,040	-6.89%
5% ES	-440,690	-458,890	-3.97%	-462,200	-494,050	-6.45%
Mean	0	586	-	0	-14,508	-

	Model 3		
	(a)	(b)	Δ
5% VaR	-328,410	-345,910	-5.06%
5% ES	-465,720	-488,200	-4.60%
Mean	0	-10,221	-

Table 11: Values in EUR of the mean, 5% VaR and 5% ES of revenue distributions obtained under different assumptions: (a) pricing and hedging decisions are performed with a suboptimal model while the simulated data used to construct revenue distributions also evolves according to the suboptimal model, and (b) pricing and hedging decisions are performed with a suboptimal model while the simulated data used to construct revenue distributions evolves according to Model 4, the time-varying mixed C-vine copula.

5 Conclusion

In this paper we have considered the joint price and volumetric risk in wind power trading, and proposed a flexible model that has practical applicability and can be used to determine optimal hedging decisions. The problem has been addressed from an energy company's perspective, who pays wind power producers a fixed price in return for their fluctuating wind power production. Our study concentrates on the joint price and volumetric risk in Denmark, a market place characterized by illiquidity and a restricted range of available hedging instruments. As a consequence, we have extended the modeling framework to include the German spot electricity price, thereby exploiting the strong positive relation between the two price series to increase the hedging possibilities.

We have proposed and fitted a three-dimensional time-varying mixed C-vine copula model to our variables of interest, capturing rather well the marginal behavior of each variable and also the dependence between the variables. Based on simulations from the proposed model, we have shown through different examples that significant benefits are to be drawn from including other than the standard and usually employed power forwards in a hedging portfolio. We find that derivative instruments associated with the right to convert electricity to heat are especially effective at reducing risk. Moreover, instruments with reference to the German spot electricity price are especially beneficial when accounting for illiquidity on the Danish power market, in the sense that they contribute to lowering the cost of the hedging portfolio. Lastly, wind index futures have risk reducing benefits when rebalancing the hedging portfolio under certain market conditions.

To highlight the importance of capturing heavy tails in the marginal distributions, tail dependence and time variation in the copula model - all characteristics that are captured by the proposed empirical model - similar calculations were performed with alternative and more naive models. We find that the inability to capture these characteristics leads to an underestimation of risk. Also, some instrument prices are affected to a large extent.

Power markets can differ greatly from region to region, and our framework is *tailored* to the case of Denmark. Depending on the availability of hedging instruments in other power markets, one might wish to include additional variables to the model. In e.g. Coulon et al. (2013), the authors propose gas-fired plants to hedge joint price and volumetric risk associated with load-serving obligations. We acknowledge that gas-fired (or coal-fired) plants could potentially be beneficial in our application as well. In this connection, the evolution of gas (or coal) prices would have to be taken into account. What is particularly useful about our model framework is that other variables can easily be incorporated in the vine copula, and hence different

situations can be accommodated without much difficulty.

Acknowledgments

The authors would like to thank Esben Høg, Thomas Aalund Fredholm, Christian Sønderup and Jakob Vive Munk for providing valuable ideas and comments. The authors also thank two anonymous referees for providing constructive criticism and suggestions that enhanced the quality of this paper.

Funding

Anca Pircalabu acknowledges support from Innovation Fund Denmark, grant 4135-00082B.

References

K. Aas, C. Czado, A. Frigessi, and H. Bakken. Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics*, 44(2):182 – 198, 2009.

K. Avdulaj and J. Barunikl. Are benefits from oil–stocks diversification gone? New evidence from a dynamic copula and high frequency data. *Energy Economics*, 51:31 – 44, 2015.

T. Bedford and R. M. Cooke. Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematical and Artificial Intelligence*, 32(1):245 – 268, 2001.

T. Bedford and R. M. Cooke. Vines - a new graphical model for dependent random variables. *Annals of Statistics*, 30(4):1031 – 1068, 2002.

F. E. Benth and P. C. Kettler. Dynamic copula models for the spark spread. *Quantitative Finance*, 11(3):407 – 421, 2011.

F. E. Benth, J. Šaltytė Benth, and S. Koekebakker. *Stochastic Modelling of Electricity and Related Markets*. World Scientific, 2008.

D. Berg. Copula goodness-of-fit testing: An overview and power comparison. *European Journal of Finance*, 15(7-8):675 – 701, 2009.

E. C. Brechmann and C. Czado. Risk management with high-dimensional vine copulas: An analysis of the Euro Stoxx 50. *Statistics & Risk Modeling*, 30(4):307 – 342, 2013.

M. Burger, B. Klar, A. Muller, and G. Schindlmayr. A spot market model for pricing derivatives in electricity markets. *Quantitative Finance*, 4(1):109 – 122, 2004.

R. Carmona and V. Durrleman. Pricing and hedging spread options. *SIAM Review*, 45(4):627 – 685, 2003.

U. Cherubini and E. Luciano. Bivariate option pricing with copulas. *Applied Mathematical Finance*, 9(2):69 – 85, 2002.

M. Coulon, W. B. Powell, and R. Sircar. A model for hedging load and price risk in the Texas electricity market. *Energy Economics*, 40:976 – 988, 2013.

D. Creal, S. J. Koopman, and A. Lucas. Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, 28(5):777 – 795, 2013.

C. Czado, U. Schepsmeier, and A. Min. Maximum likelihood estimation of mixed C-vines with application to exchange rates. *Statistical Modelling*, 12 (3):229 – 255, 2012.

EEX. EEX group workshop: European power. EEX.com, 2014.

C. Elberg and S. Hagspiel. Spatial dependencies of wind power and interrelations with spot price dynamics. *European Journal of Operational Research*, 241(1):260 – 272, 2015.

H. Gatfaoui. Linking the gas and oil markets with the stock market: Investigating the U.S. relationship. *Energy Economics*, 53:5 – 16, 2016.

C. Genest, B. Remillard, and D. Beaudoin. Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and Economics*, 44(2): 199 – 213, 2009.

O. Grothe and J. Schnieders. Spatial dependence in wind and optimal wind power allocation: A copula-based analysis. *Energy Policy*, 39(9):4742 – 4754, 2011.

Harry J. *Multivariate Models and Dependence Concepts*. Chapman and Hall, 1 edition, 1997.

S. P. Kolos and E. I. Ronn. Estimating the commodity market price of risk for energy prices. *Energy Economics*, 30(2):621 – 641, 2008.

R. I. McKinnon. Future markets, buffer stocks, and income stabiliy for primary producers. *Journal of Political Economy*, 75(6):844 – 861, 1967.

NASDAQ OMX. Trading appendix 2: Contract specifications. nasdaqomx.com, 2015.

R. B. Nelsen. *An Introduction to Copulas*. Springer, 1999.

Y. Oum and S. S. Oren. VaR constrained hedging of fixed price load-following obligations in competitive electricity markets. *Risk and Decision Anaysis*, 1 (1):43 – 56, 2009.

Y. Oum and S. S. Oren. Optimal static hedging of volumetric risk in a competitive wholesale electricity market. *Decision Analysis*, 7(1):107 – 122, 2010.

Y. Oum, S. S. Oren, and S. Deng. Hedging quantity risks with standard power options in a competitive wholesale electricity market. *Naval Research Logistics*, 53(7):697 – 712, 2006.

A. J. Patton. Modelling asymmetric exchange rate dependence. *International Economic Review*, 47(2):527 – 556, 2006.

A. J. Patton. Copula methods for forecasting multivariate time series. In Graham Elliott and Allan Timmermann, editors, *Handbook of Economic Forecasting*, volume 2B, pages 899 – 960. Elsevier B.V., 2013.

A. Pircalabu, T. Hvolby, J. Jung, and E. Høg. Joint price and volumetric risk in wind power trading: A copula approach. *Energy Economics*, 62:139 – 154, 2017.

J. C. Reboredo and A. Ugolini. A vine-copula conditional value-at-risk approach to systemic sovereign debt risk for the financial sector. *The North American Journal of Economics and Finance*, 32:98 – 123, 2015.

A. Sklar. Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de L'Université de Paris*, 8:229 – 231, 1959.

Paper III

A regime-switching copula approach to modeling day-ahead prices in coupled electricity markets

List of authors: Anca Pircalabu[*,†] and Fred Espen Benth[‡]
[*] Department of Mathematical Sciences, Aalborg University
[†] Quantitative Analytics, Neas Energy
[‡] Department of Mathematics, University of Oslo

Published in
Energy Economics Vol. 68, pp. 283–302, 2017.
DOI: https://doi.org/10.1016/j.eneco.2017.10.008

The layout has been revised.

ABSTRACT

The recent price coupling of many European electricity markets has triggered a fundamental change in the interaction of day-ahead prices, challenging additionally the modeling of the joint behavior of prices in interconnected markets. In this paper we propose a regime-switching AR–GARCH copula to model pairs of day-ahead electricity prices in coupled European markets. While capturing key stylized facts empirically substantiated in the literature, this model easily allows us to 1) deviate from the assumption of normal margins and 2) include a more detailed description of the dependence between prices. We base our empirical study on four pairs of prices, namely Germany–France, Germany–Netherlands, Netherlands–Belgium and Germany–Western Denmark. We find that the marginal dynamics are better described by the flexible skew t distribution than the benchmark normal distribution. Also, we find significant evidence of tail dependence in all pairs of interconnected areas we consider. As a first application of the proposed model, we consider the pricing of financial transmission rights, and highlight how the choice of marginal distributions and copula impacts prices. As a second application we consider the forecasting of tail quantiles, and evaluate the out-of-sample performance of competing models.

1 Introduction

Since the many projects launched over the past decade to achieve increased market integration across day-ahead electricity markets in Europe, the day-ahead electricity price convergence between market areas has increased significantly, adding to the complexity of modeling the joint behavior of day-ahead electricity prices.

The first significant step towards market integration in Europe was taken with the introduction of the Nord Pool market, which is the main platform for trading power in most Nordic and Baltic countries. At Nord Pool, the bidding areas are price coupled, meaning that the transmission capacity is auctioned as an implicit part of the day-ahead auction of electrical energy. This achieves - as opposed to the explicit auctioning, where the transmission capacity auction and the day-ahead auction of electrical energy are separate actions - the efficient utilization of interconnectors and results in smaller price differences between the bidding areas. Since Nord Pool's creation in the mid 90's, many countries have joined and Nord Pool comprises today Norway, Sweden, Denmark, Finland, Estonia, Latvia and Lithuania. For the sake of completion, we also mention that Nord Pool has taken sole ownership of the UK market as of 2014.

Although Nord Pool has a long-standing history of price coupling, such initiatives are more recent for other European countries. In 2006, the so-

called Tri-Lateral Market Coupling (TLC) project comprising France, Belgium and the Netherlands was initiated. Later in 2010, the price coupling of Central West Europe (CWE) was achieved, integrating Germany, France, Austria, Belgium, the Netherlands and Luxembourg. In 2014, a crucial milestone was reached with the price coupling of the North Western European (NWE) region, covering the CWE region, the UK, the Nordic and Baltic countries - and as of 2016, the price coupling region has grown to include 19 European countries. Also, a new algorithm regarding how cross-border capacities are included in the day-ahead price calculation was introduced on 20 May 2015 in the CWE region, with the transition from the Available Transmission Capacity (ATC) methodology to the more efficient Flow-Based (FB) methodology.

For the joint modeling of day-ahead electricity prices in two interconnected areas, price coupling has induced a fundamental change that cannot be ignored. In the case of the CWE region for example, the transition from explicit to implicit auctioning of transmission capacity has meant that equal prices in two adjacent markets are very often observed now. As already accentuated in the existing literature (see e.g. Füss et al. (2017)), this has generally resulted in the non-feasibility of previously proposed models. For one, the pricing of spread options such as the financial transmission rights cannot be achieved using spread option formulas á la Margrabe. Also in risk management applications, the classical reduced-form models are no longer suitable for modeling the joint behavior of prices. To exemplify some events and their influence on the day-ahead electricity prices, we plot in Fig. 1 the evolution of the German–French price spread at hour 8, for the period 1 January 2009 to 25 September 2016. According to Fig. 1, the most visible change in the spread dynamics was caused by the price coupling of Germany and France in 2010, leading to a large amount of exact price convergence in the following period.

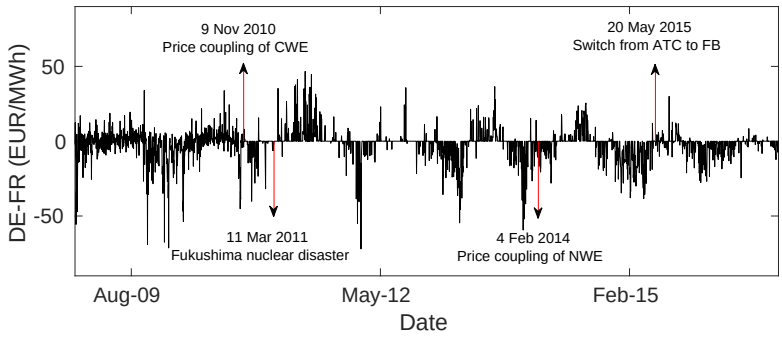


Fig. 1: Day-ahead electricity price spread for Germany–France (DE-FR) at hour 8, from 1 January 2009 to 25 September 2016.

In this paper, we propose a regime-switching AR–GARCH copula model for the joint behavior of pairs of day-ahead prices in coupled electricity markets, inspired by ideas from Haldrup and Nielsen (2006) and Haldrup et al. (2010). The regime-switching part is essential for distinguishing the case of identical prices from the case of non-identical prices. The AR–GARCH functions as a filter to remove the serial dependence in the conditional means and conditional variances. The use of copula models facilitates the analysis in the sense that marginal models for day-ahead prices are treated separately from the dependence structure. Furthermore, copulas allow for the easy transition to more realistic distributional assumptions for the residuals, and also allow for the inclusion of a more detailed description of the dependence structure when the prices do not coincide. Our empirical study concentrates on four coupled markets: Germany–France, Germany–Netherlands, Netherlands–Belgium and Germany–Western Denmark.

The proposed model has a wide range of applications in the context of derivatives pricing, hedging and risk management. Here, we consider two applications: The pricing of financial transmission rights, and the out-of-sample forecasting of tail quantiles. In both applications, we study how the results are impacted by the choice of marginal distributions and the choice of copula. Furthermore in the forecasting exercise, we demonstrate the usefulness of the proposed modeling framework by comparing the forecast performance of copula-based models against a simple alternative that is not copula-based.

The existing literature on bivariate or multivariate models for day-ahead prices in coupled electricity markets is scarce. Due to the aforementioned increasing complexity in the joint price dynamics, some authors employ fundamental models, with examples being Kiesel and Kustermann (2016) and Füss et al. (2017). Kiesel and Kustermann (2016) propose a fundamental model in the spirit of Carmona et al. (2013), and derive closed-form pricing formulas for futures and plain vanilla options under market coupling. Moreover, they provide an empirical application using data from the German and French markets. In Füss et al. (2017), the authors also employ a fundamental model, and consider the pricing of futures and transmission rights under different allocation schemes. Aside from providing closed-form pricing formulas, sensitivity analyses are also carried out.

However appealing fundamental models are when it comes to their adaptability to changes in the market structure - which indeed happen quite often - they can be extremely data-heavy and thus both theoretically and empirically challenging, as also acknowledged by Kiesel and Kustermann (2016). In the empirical analysis of Kiesel and Kustermann (2016) regarding Germany–France, the data input amounts to many different time series: Day-ahead electricity price data, coal, gas and oil data to obtain a proxy for the marginal fuel, residual expected demand data for both countries (which is in turn

obtained by subtracting renewable electricity generation from the expected demand), and also data on installed transmission capacities. Our proposed model on the other hand, which belongs to the class of reduced-form models, is not nearly as data intensive, while still being able to comply with the institutional framework.

Examples on applications of time series models in the context of price coupled electricity markets are Haldrup and Nielsen (2006), where a regime-switching multiplicative SARFIMA model is proposed for the modeling of price spreads in the Nord Pool area, and Haldrup et al. (2010), where a vector autoregressive model with long memory and regime switches is proposed and applied again to Nord Pool data.

In the context of copula models and their applications in energy markets, copulas have been gaining more interest recently. Some examples are Börger et al. (2009), who use copulas in the context of cross-commodity risk management, Benth and Kettler (2011), who employ copulas to model the spark spread, Avdulaj and Barunikl (2015), who use dynamic copulas to investigate oil-stock diversification, and Grothe and Schnieders (2011) and Elberg and Hagspiel (2015), who consider copulas in the context of wind power in Germany.

The remaining of this paper is organized as follows: Section 2 describes the data that our analysis is based on. Section 3 introduces the regime-switching AR–GARCH copula model. In Section 4 the proposed model is fitted to the data. Sections 5 and 6 present two empirical applications of the model. Section 5 concludes.

2 Data

Our empirical study relies on day-ahead electricity price data from four price coupled markets: Germany–France (DE–FR), Germany–Netherlands (DE–NL), Netherlands–Belgium (NL–BE), and Germany–Western Denmark (DE–DK1). After careful considerations, we choose to fix the sample window to the period 1 May 2011 to 20 May 2015. Since areas of CWE are mostly represented in our analysis, and we wish to analyze the period following price coupling, a natural starting point for the sample is 10 November 2010, which marks the first day of price coupling in the CWE region[1]. Shortly after however, the Fukushima nuclear disaster followed, leading to a sudden change in German policy, with the closure of nuclear power plants and the significant focus on the development of renewable generation sources. To exclude the immediate perturbation caused by the Fukushima accident, we thus let the start of our sample be 1 May 2011. Regarding the end point, we choose the

[1]We note that price coupling in CWE was launched 9 November 2010, for delivery day 10 November 2010.

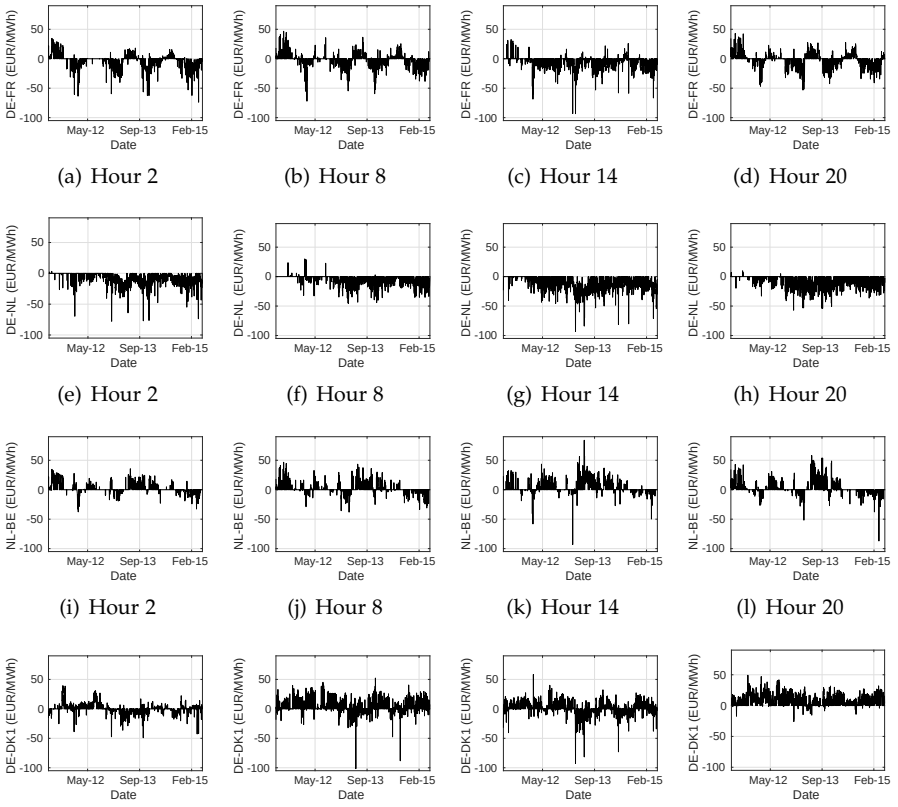
date marking the switch to the FB methodology. According to EPEX (2015), the FB coupling algorithm allows for more sophisticated grid modeling to optimize the flows on interconnectors in comparison to the ATC methodology. Effects of the change regarding how flows are included in the calculation of day-ahead prices are not evident from just looking at time series plots of the data. However, we do suspect some implications on the joint price dynamics, and let the sample end at this date as a result. We will get back to this issue later in the paper.

Having fixed the sample window to the period 1 May 2011 - 20 May 2015, we avoid the most significant structural changes in the institutional framework for the pairs belonging to the CWE region. Careful attention is however needed in the case of the DE–DK1 pair, where price coupling was first introduced on 4 February 2014. Before then, Western Denmark was volume coupled to Germany, using explicit auctioning. This is not necessarily as worrying as the case of price coupling in the CWE region, since exact price convergence between Germany and Western Denmark was achieved frequently also before their price coupling. Nevertheless, we shall treat the pair DE–DK1 with special consideration before applying the same modeling framework to all pairs of prices. Although the pair DE–DK1 complicates our analysis somewhat, its inclusion is motivated by having one of Nord Pool's bidding areas represented in our analysis, and also by having a pair where the in-feed of renewable energy is very significant in both markets.

We consider hourly price data separately for each hour of the day - so for all four pairs, we will have 24 hourly time series, consisting of 1,481 observations each. Fig. 2 shows time series plots of the four pairs of spreads at four different hours of the day, and Table 1 provides summary statistics in order to get a better sense of the data. To avoid distortions, few extreme prices have been truncated: Specifically, prices below -50 EUR/MWh were set to -50, and prices above 150 EUR/MWh were set to 150.[2]

The spreads show a rather different behavior across pairs and also across the hours of the day. For DE–FR, we observe a changing seasonal pattern during (or near) off-peak. During summer off-peak, Germany usually imports electricity from France due to low renewables generation in Germany (less wind during summer periods, and very little solar during off-peak) and excess nuclear production in France due to the lower demand. During winter peak and off-peak, France usually imports from Germany due to increased demand and higher production in Germany caused by increased renewables

[2]For most hourly price time series no truncation is necessary, since prices below -50 or above 150 never occur. The maximum number of observations below -50 that are truncated in one single hourly price series is 7, and corresponds to an hourly price series for DE. The maximum number of observations above 150 that are truncated in one single hourly price series is 6, and corresponds to an hourly price series for FR. The truncation rule is mostly introduced to deal with cases like e.g. the extreme spikes we observe in DK1 on 7 June 2013, where prices reached approx. 2,000 EUR/MWh for five consecutive morning hours.



(a) Hour 2 (b) Hour 8 (c) Hour 14 (d) Hour 20

(e) Hour 2 (f) Hour 8 (g) Hour 14 (h) Hour 20

(i) Hour 2 (j) Hour 8 (k) Hour 14 (l) Hour 20

(m) Hour 2 (n) Hour 8 (o) Hour 14 (p) Hour 20

Fig. 2: Historical hourly spreads between the pairs DE & FR, DE & NL, NL & BE and DE & DK1 at four different hours of the day.

generation.

In the case of the DE–NL pair, the German price is almost always below the Dutch price, signaling Netherlands to be a significant importer of German power. This relation can be explained by the electricity generation mix in the Netherlands, which relies heavily on more expensive natural gas.

For NL–BE we see that the Dutch price is more often observed to be above the Belgian price, which again is due to the more expensive generation sources in the Netherlands compared to Belgium. However, both countries are mainly importers, and their day-ahead electricity prices are heavily influenced by conditions in Germany and France. A seasonal pattern is also observed here, but is not as persistent as in the case of DE–FR. According to TenneT (2013), this is because higher demand for electrical heating in France during winter is somehow reflected in the Belgian price, which in such situ-

ations would typically increase above the level of Dutch prices.

Regarding the DE–DK1 pair, prices in Western Denmark are usually below Germany. Renewable power is significant in both countries - DK1 has a lot of wind and DE has both wind and solar - however Denmark is connected to both Sweden and Norway, from where it can import cheap hydropower. A generation deficit typically takes place during summer periods in DK1 due to less wind. Depending on transmission capacity limitations, weather conditions, etc., imports come from Germany, Sweden and/or Norway.

	Spread $= 0$	Spread > 0	Spread < 0	Mean	Std. dev.	Lin. corr.	Lin. corr. (Spread$\neq 0$)
Germany and France (DE - FR)							
Hour 2	51.11%	14.32%	34.57%	-3.41	11.53	0.53	0.39
Hour 8	57.60%	15.67%	26.73%	-1.78	10.27	0.85	0.70
Hour 14	49.90%	4.25%	45.85%	-5.28	10.22	0.80	0.68
Hour 20	49.29%	20.05%	30.66%	-1.96	10.86	0.70	0.55
Germany and Netherlands (DE - NL)							
Hour 2	35.11%	0.14%	64.75%	-7.96	10.38	0.48	0.43
Hour 8	58.41%	1.08%	40.51%	-4.69	8.29	0.88	0.81
Hour 14	34.77%	0.00%	65.23%	-10.79	11.90	0.67	0.55
Hour 20	50.17%	0.54%	49.29%	-6.96	9.90	0.63	0.49
Netherlands and Belgium (NL - BE)							
Hour 2	65.83%	21.00%	13.17%	1.50	7.71	0.72	0.57
Hour 8	71.10%	17.89%	11.01%	1.17	7.59	0.90	0.79
Hour 14	69.82%	21.54%	8.64%	2.75	9.60	0.75	0.58
Hour 20	66.51%	21.81%	11.68%	1.92	9.33	0.70	0.36
Germany and Western Denmark (DE - DK1)							
Hour 2	36.60%	35.72%	27.68%	0.09	7.90	0.76	0.57
Hour 8	30.11%	55.10%	14.79%	4.75	10.99	0.79	0.68
Hour 14	28.83%	46.58%	24.58%	1.47	10.11	0.78	0.70
Hour 20	23.63%	71.51%	4.86%	7.76	9.24	0.71	0.69

Table 1: Summary statistics for the spread between day-ahead electricity prices for four selected periods. The hour denoted "Hour 2" represents the hour starting at 2 and ending at 3, and the same holds for all other hours. Furthermore, the time zone is CET. The sample covers the period 1-May-2011 to 20-May-2015, a total of 1,481 observations in each of the four time series per spread we consider here. All spreads are measured in EUR/MWh.

Considering in more detail the summary statistics presented in Table 1, we see that exact price convergence or equivalently a price spread of zero occurs very often for all pairs and hours of the day. Moreover, the exact price coupling percentage can vary quite a lot depending on the hour of the day, and so can the mean and the standard deviation. We also observe a strong positive relation between the prices conditional on the spread being nonzero. This is of course expected due to the limited transmission capacity which impedes exact price convergence at all times - note that only a utilization of the available interconnector capacity below 100% would allow exact price convergence.

In Fig. 3, we consider in more detail the behavior of exact price convergence across the hours of the day. For the CWE, there is generally less

exact price convergence during the night hours, and also during midday. For DE–DK1, the situation is somewhat reversed, with more coupling during off-peak.

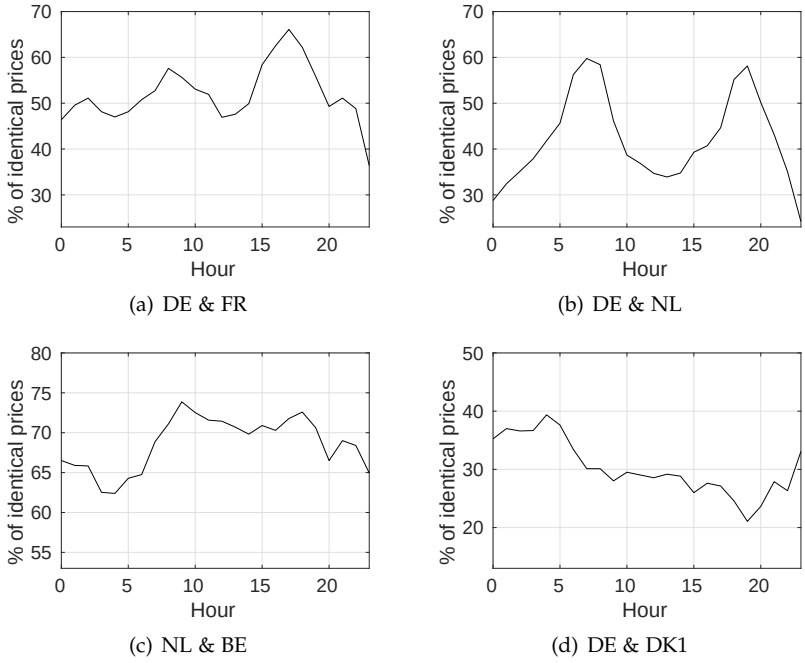


Fig. 3: Percentages of absolute price convergence grouped by hour of the day.

The shape of the plots for the CWE areas, especially the plot for DE–NL, where the shape is most pronounced, reminds us of the "two hump" behavior of German day-ahead prices. Typically, the German price is low during the night because of lower demand and high wind in-feed (especially during winter). Prices are also lowered during midday, due to the high solar in-feed, which peaks at noon. Why the shape of German prices during the day is reflected in the absolute coupling percentage can, to a high extent, be explained by the mix of generation assets across markets.

Let us consider the example of DE–NL. The shape of the Dutch price during the peak hours is rather flat compared to Germany, and the level of prices is higher due to the more expensive marginal cost of natural gas-fired power plants, which play a major role in the daily fuel mix. A smaller price difference between DE–NL is thus more likely to occur during morning and evening, when the German prices reach their highest level. This implies that the remaining gap that needs to be closed in order to obtain exact price convergence is smaller during these periods, and more likely to be achieved

with less than a 100% utilization of the available transmission capacity.

In the case of DE–DK1, the increased exact price convergence during off-peak can be explained by the combination of lower demand and high in-feed of wind generation in both countries, which leads to smaller price differences during off-peak.

3 Model specification

To model the evolution of prices in two interconnected markets, we propose a regime-switching AR–GARCH copula model for each hour of the day separately. We allow for two regimes, which are in fact observable and can be categorized as 1) identical and 2) non-identical prices. Making this distinction is of uttermost importance since it allows us to replicate periods of spreads equaling zero (see Fig. 2 and Table 1).[3]

Our modeling procedure can be divided into three steps: In a first step, we deal with the seasonality aspect. For all price areas we consider in this paper, that is BE, DK1, DE, FR and NL, the mechanism of day-ahead price formation is based on matching supply and demand curves on an hourly basis. Hence, the strong seasonal variation that characterizes both the demand and the supply is reflected in the behavior of the electricity prices. To correct the data for seasonality, suitable seasonal functions are applied to each pair of prices, for each hour of the day. The seasonal functions are such that the identical/non-identical price patterns are preserved in the deseasonalized prices.

In a second step, regime-switching AR–GARCH filters are applied to each pair of deseasonalized prices. Again in this case, the resulting residuals will maintain the same pattern of identical and non-identical prices that we observe in the actual price data. Hence, the residuals can be split according to 1) a case of perfect positive dependence, where the residuals from the regime-switching AR–GARCH will coincide, and 2) a case of *non-perfect* dependence.

In a third step, we study the dependence of non-equal residuals by applying the flexible copula models. Below, we elaborate on the second and third steps in our modeling framework.

3.1 A regime-switching AR–GARCH copula model

Let us denote by $y_{n,t,i}$ the price in area i, at hour n and day t, which has been corrected for deterministic seasonality. Since we only consider pairs of

[3]We mention in passing that univariate time series models of the ARMA–GARCH type have been successfully applied to model day-ahead electricity prices previously (see e.g. Keles et al. (2012) and Paraschiv (2013)).

day-ahead prices, $i = 1, 2$. To simplify the notation in what follows, we suppress the subscript n. Whenever we refer to $y_{t,i}$, we mean the deseasonalized day-ahead price for some given hour of the day. We consider the following dynamics for the prices: If $y_{t,1} = y_{t,2}$, then we are in the regime of equal prices denoted by $s_t = 0$, and for $i = 1, 2$,

$$y_{t,i} = \sum_{p=1}^{P} \phi_p^{(0)} \frac{1}{2}(y_{t-p,1} + y_{t-p,2}) + \epsilon_{t,i} \tag{1}$$

$$\epsilon_{t,i} = \sigma_{t,i}\eta_{t,i} \tag{2}$$

$$\sigma_{t,i}^2 = \omega^{(0)} + \alpha^{(0)} \frac{1}{2}(\epsilon_{t-1,1}^2 + \epsilon_{t-1,2}^2) + \beta^{(0)} \frac{1}{2}(\sigma_{t-1,1}^2 + \sigma_{t-1,2}^2). \tag{3}$$

Note that under $s_t = 0$, it follows from the above that $\epsilon_{t,1} = \epsilon_{t,2}$ and $\sigma_{t,1}^2 = \sigma_{t,2}^2$. If $y_{t,1} \neq y_{t,2}$, then we are in the regime of non-equal prices denoted by $s_t = 1$, and for $i = 1, 2$,

$$y_{t,i} = \sum_{p=1}^{P} \phi_p^{(i)} y_{t-p,1} + \sum_{q=1}^{Q} \xi_q^{(i)} y_{t-q,2} + \epsilon_{t,i} \tag{4}$$

$$\epsilon_{t,i} = \sigma_{t,i}\eta_{t,i} \tag{5}$$

$$\sigma_{t,i}^2 = \omega^{(i)} + \alpha^{(i)}\epsilon_{t-1,i}^2 + \beta^{(i)}\sigma_{t-1,i}^2. \tag{6}$$

According to Eqs. (1) and (4) for the conditional mean, and Eqs. (3) and (6) for the conditional variance, all parameters vary depending on the regime. We shall thus obtain three sets of parameter estimates: One set corresponding to the regime where prices are identical, denoted with superscript (0), and two sets of parameters corresponding to the regime where prices are not identical, denoted with superscripts (i), for $i = 1, 2$.

To provide a parsimonious representation, the structure of the models is slightly different under the two regimes. In the model for the conditional mean, identical prices are explained by lagged average prices from the two areas, whereas no averaging appears in the case of non-identical prices. We note that aside from autoregressive terms, we allow for cross-equation effects by including lagged values of "the other price series", in the spirit of vector autoregressions. Nevertheless, we will refer to the model for the conditional mean as an AR. The model for the conditional variance resembles the GARCH(1,1) specification: In times of identical prices, the model contains averages of lagged values, whereas the non-identical price regime is characterized by the classical GARCH(1,1) dynamics. While accounting for serial dependence and heteroskedastisity, we stress that the construction of the model in Eqs. (1)–(6) ensures that the regime switching nature of prices is maintained in the standardized residuals.

For the distribution of the standardized residuals denoted by η in

Eqs. (2) and (5), we consider different possibilities. First, we consider the standard choice of a normal distribution:

$$\eta_{t,i} \sim \mathrm{N}(0,1) \qquad \text{for } i = 1,2. \tag{7}$$

To allow for more flexibility, we also consider

$$\eta_{t,i} \sim \begin{cases} \text{Skew } t(0,1,\nu^{(0)},\lambda^{(0)}) & \text{if } s_t = 0 \\ \text{Skew } t(0,1,\nu^{(i)},\lambda^{(i)}) & \text{if } s_t = 1, \text{for } i = 1,2, \end{cases} \tag{8}$$

where ν and λ denote the degrees of freedom and skewness parameters of the skew t distribution, respectively, and are allowed to vary with the regimes.

Next, we introduce copulas to model the dependence structure of standardized residuals belonging to the non-equal price regime. Copulas are flexible tools that enable the study of dependence structures beyond the linear correlation. For the two-dimensional case we consider here, a copula is essentially a bivariate distribution function C defined on the unit cube, with standard uniform margins. The central result concerning copulas is due to Sklar's theorem [Sklar (1959)], which states that if we assume (η_1, η_2) to have bivariate distribution function F and univariate marginal distribution functions F_1, F_2, then there exists a bivariate copula C, such that

$$F(\eta_1, \eta_2) = C(F_1(\eta_1), F_2(\eta_2)). \tag{9}$$

If the marginal distribution functions are continuous, we have furthermore that the copula is unique, and can be defined through

$$C(u_1, u_2) = F(F_1^{-1}(u_1), F_2^{-1}(u_2)), \tag{10}$$

where the u's represent standard uniform variables and the F^{-1}'s represent the inverse of the marginal distribution functions. Adding to the flexibility of copulas is that the converse of Sklar's theorem also holds, meaning that we can combine two marginal distributions, be it from different families even, to form a joint distribution.

Since we condition on the information generated by past observations of the variables (by using AR–GARCH filters), we are in fact considering here the conditional copula. Also in this case, Sklar's theorem applies, as shown in Patton (2001).

3.2 Estimation procedure

Estimation of the model parameters is done by maximum likelihood. The joint conditional density function is obtained by differentiating Eq. (9), and thus the full-sample log-likelihood takes the form

$$\log \mathcal{L} = \sum_{i=1}^{2} \sum_{t=1}^{T} \log f_i(y_{t,i} \mid \mathcal{F}_{t-1}; \Theta_i) + \mathbb{I}_{s_i=1} \sum_{t=1}^{T} \log c((u_{t,1}, u_{t,2}) \mid \mathcal{F}_{t-1}; \Theta_3), \tag{11}$$

where

$$\sum_{t=1}^{T} \log f_i(y_{t,i} \mid \mathcal{F}_{t-1}; \Theta_i) = \mathbb{I}_{s_t=0} \sum_{t=1}^{T} \log f_i(y_{t,i} \mid \mathcal{F}_{t-1}; \Theta_i^{(0)}) + \mathbb{I}_{s_t=1} \sum_{t=1}^{T} \log f_i(y_{t,i} \mid \mathcal{F}_{t-1}; \Theta_i^{(i)}). \quad (12)$$

In Eqs. (6) and (12), the functions f_i denote the conditional marginal density functions for area 1 and 2, and c is the conditional copula density representing the non-equal price regime. The variable $\mathbb{I}_\Omega$ denotes the indicator function of the event Ω. Furthermore, $\mathcal{F}_{t-1}$ denotes the filtration, and $(\Theta_1, \Theta_2, \Theta_3)$ refer to the parameters for the full model. Specifically, (Θ_1, Θ_2) refer to the parameters of the regime-switching AR–GARCH model, and Θ_3 refers to the copula parameters.

Recall that by construction of the model, $u_{t,1} = u_{t,2}$ if $y_{t,1} = y_{t,2}$, and $u_{t,1} \neq u_{t,2}$ otherwise. In the equal price regime, the residuals thus coincide and we have a case of perfect positive dependence (also referred to as the Fréchet-Hoeffding upper bound), where there is no copula parameter to be estimated. Hence, the only copula parameter to be estimated corresponds to the pair of standardized residuals in the non-equal state.

Usually, the model parameters are obtained by performing a multi-stage maximum likelihood estimation, where the marginal models and the copula are considered separately, see e.g. Dias and Embrechts (2009) and Patton (2013). Since the parameter vectors of the marginal models have common elements in our case, these cannot be estimated separately. Hence, in a first step, a joint numerical maximization of the the first term in Eq. (6) is performed to obtain estimates for (Θ_1, Θ_2). Then, standardized residuals are obtained, and the uniforms entering the copula log-likelihood follow immediately from applying the probability integral transform. In a second step, the copula parameters are estimated by maximizing the copula log-likelihood numerically.

4 Model fitting

4.1 Seasonal function

As mentioned in Section 3, the first step in our data analysis is concerned with the seasonality correction. Following the related literature (see e.g. Lucia and Schwartz (2002), Haldrup et al. (2010) and Janczura et al. (2013)), we consider the following seasonal function Λ for each area in a price pair:

$$\Lambda_t^{s_t} = a^{s_t} + b^{s_t} t + c^{s_t} \sin(2\pi t/365) + d^{s_t} \cos(2\pi t/365) + \sum_{j=1}^{4} w_j^{s_t} W_t^j + h^{s_t} H_t, \quad (13)$$

for $s_t = 0, 1$. In Eq. (13), a is a constant, b denotes the trend coefficient and c and d are coefficients for the yearly cycle. To capture the weekly seasonality, we use day-of-week dummies denoted by W^j, with their corresponding coefficients denoted by w_j. We group Tuesday, Wednesday and Thursday in one dummy variable, and have therefore limited ourselves to only four day-of-week coefficients. Lastly, the dummy variable H is included to capture holiday effects, and its coefficient is denoted by h.[4]

The seasonal functions are fitted to the data by ordinary least squares. To ensure that the deseasonalized price data corresponding to a given pair retains the same regime-switching structure, note that we let all parameters vary with the regime s_t. For a given pair, the parameter estimates corresponding to the state $s_t = 0$, i.e., state of equal prices, will by construction be the same for area $i = 1$ and area $i = 2$. For the state $s_t = 1$, i.e., state of non-equal prices, the parameter estimates will differ for the two areas.

4.2 Marginal models

After removing the deterministic seasonal component, the regime-switching AR–GARCH model introduced in Eqs. (1) – (6) is fitted to the data. Specifically, the model is fitted to all four pairs of prices and all hours of the day, which corresponds to a total of 96 models. For the order of the autoregression (P) and cross-equation effects (Q) in the equations for the conditional mean, we consider $P = 1, \ldots, 7$ and $Q = 1, \ldots, 7$. The optimal order of the models is chosen based on the Bayesian Information Criterion (BIC). In the interest of brevity, we provide the detailed results in Appendix A, Table 14. We do however mention that the order of Q is zero for the majority of the models, revealing that the cross-equation effects (cf. Eq. (4)) can generally be omitted.

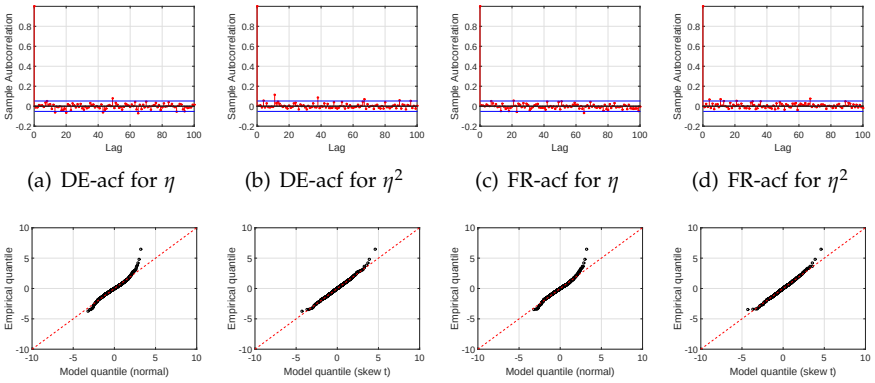
Model specifications other than those stated in Eqs. (1) – (6) were experimented with, and we note that our proposed regime-switching AR–GARCH model is one of many possible specifications. We found however our specification to be superior in terms of parsimony and its fit to the data.

To provide some evidence for the fit of the regime-switching AR–GARCH, we consider as an example the DE–FR pair at hour 8. Table 2 presents the estimation results, and Fig. 4 displays sample autocorrelation and quantile plots. Sample autocorrelation plots are provided for both the standardized residuals and the squared standardized residuals, and give indication of almost no serial correlation left. Regarding the quantile plots, we provide both the fit with the normal and the skew t distribution, illustrating the superiority of the latter. Similar results to those in Fig. 4 were obtained for the

[4] In the case of identical prices, a holiday might occur in one area but not the other. To avoid this, the holiday dummy returns a 1 if a holiday occurs in one of the two areas, and 0 otherwise.

Model for the pair DE–FR hour 8								
Equal price regime ($s_t = 0$)			**Non-equal price regime** ($s_t = 1$)					
			Area 1: DE			**Area 2: FR**		
Conditional mean (optimal order $P = 6$, $Q = 0$)								
$\hat{\phi}_1^{(0)}$	0.3458	(0.0381)	$\hat{\phi}_1^{(1)}$	0.4048	(0.0507)	$\hat{\phi}_1^{(2)}$	0.5036	(0.0481)
$\hat{\phi}_2^{(0)}$	0.1648	(0.0411)	$\hat{\phi}_2^{(1)}$	0.0810	(0.0508)	$\hat{\phi}_2^{(2)}$	0.1192	(0.0500)
$\hat{\phi}_3^{(0)}$	0.0134	(0.0396)	$\hat{\phi}_3^{(1)}$	0.0203	(0.0496)	$\hat{\phi}_3^{(2)}$	0.0611	(0.0502)
$\hat{\phi}_4^{(0)}$	0.0666	(0.0392)	$\hat{\phi}_4^{(1)}$	0.0329	(0.0483)	$\hat{\phi}_4^{(2)}$	0.0644	(0.0519)
$\hat{\phi}_5^{(0)}$	0.0420	(0.0377)	$\hat{\phi}_5^{(1)}$	0.0111	(0.0458)	$\hat{\phi}_5^{(2)}$	0.0043	(0.0491)
$\hat{\phi}_6^{(0)}$	0.0949	(0.0361)	$\hat{\phi}_6^{(1)}$	0.1910	(0.0431)	$\hat{\phi}_6^{(2)}$	0.1987	(0.0450)
Conditional variance								
$\hat{\omega}^{(0)}$	6.7896	(5.7732)	$\hat{\omega}^{(1)}$	8.2661	(8.1208)	$\hat{\omega}^{(2)}$	15.6209	(11.2384)
$\hat{\alpha}^{(0)}$	0.0487	(0.0410)	$\hat{\alpha}^{(1)}$	0.2058	(0.0880)	$\hat{\alpha}^{(2)}$	0.1218	(0.0787)
$\hat{\beta}^{(0)}$	0.7641	(0.1260)	$\hat{\beta}^{(1)}$	0.7359	(0.1681)	$\hat{\beta}^{(2)}$	0.7693	(0.1935)
Marginal distribution (skew t)								
$\hat{\nu}^{(0)}$	4.8602	(1.0360)	$\hat{\nu}^{(1)}$	8.0736	(3.0836)	$\hat{\nu}^{(2)}$	11.5744	(4.9559)
$\hat{\lambda}^{(0)}$	0.0360	(0.0481)	$\hat{\lambda}^{(1)}$	0.0482	(0.0562)	$\hat{\lambda}^{(2)}$	0.0839	(0.0582)

Table 2: Parameter estimates for the DE–FR hour 8 model. Simulation based standard errors are given in parenthesis, and are based on 999 simulations following the procedure in Section 4.5.



(a) DE-acf for η (b) DE-acf for η^2 (c) FR-acf for η (d) FR-acf for η^2

(e) DE-qq plot normal (f) DE-qq plot skew t (g) FR-qq plot normal (h) FR-qq plot skew t

Fig. 4: Sample autocorrelation and quantile plots of standardized residuals resulting from fitting regime-switching AR–GARCH models to the DE–FR pair for hour 8.

remaining models. In Appendix A, we provide some additional results for the remaining pairs at hour 8.

4.3 The pair DE–DK1

Before proceeding to copula modeling, we recall that the pair DE–DK1 behaved differently compared to the other pairs in terms of initiatives that impacted the market structure. With price coupling first being introduced on 4 February 2014 as opposed to 10 November 2010, we wish to consider the pair DE–DK1 more closely. As mentioned earlier, the dynamics of the DE–DK1 spread has not changed visibly with the transition to price coupling, with equal prices having been observed prior as well as post price coupling. Therefore, we applied the same regime-switching AR–GARCH framework to this pair cf. Section 4.2. However, we wish to investigate whether a break in the dependence structure has occurred at the time of transitioning from volume coupling (VC) to price coupling (PC). If this were the case, the subsequent copula modeling would need to account for such a break.

We consider searching for a break in the rank correlation of the filtered prices (i.e., the standardized residuals obtained after applying the regime-switching AR–GARCH filter) corresponding to the non-equal price regime, at $t^* = 4$ February 2014. Specifically, we wish to test

$$H_0 : \rho^{VC} = \rho^{PC} \quad \text{vs.} \quad H_1 : \rho^{VC} \neq \rho^{PC},$$

where ρ^{VC} is the rank correlation (Spearman's ρ) of filtered prices corresponding to the non-equal price regime in the interval $[1, t^*]$, and ρ^{PC} denotes the rank correlation (Spearman's ρ) in the interval $(t^*, T]$. A critical value for $\rho^{VC} - \rho^{PC}$ is obtained by using an *i.i.d.* bootstrap as in Patton (2013).[5]

We perform the test for all 24 DE–DK1 models, and only find evidence against the null of equal rank correlation in 2 out of 24 cases (we consider a 5% significance level). Since we do not find stronger evidence for a break in the dependence, we decide to treat the pair DE–DK1 like all others in the subsequent analysis.

4.4 Copula models

Next, we turn to the study of the dependence between residual pairs obtained from the regime-switching AR–GARCH models. As mentioned in Section 3, only the pairs of residuals belonging to the non-equal price regime are relevant in this context, and hence copula models are fitted to this data. We restrict our attention to five copulas, namely the Gaussian, Student t, Gum-

[5]Note that by performing an *i.i.d.* bootstrap, we impose H_0.

bel, Rotated Gumbel[6] and Symmetrized Joe-Clayton (SJC)[7]. The choice of these copulas is motivated by having models that can capture different types of dependence structures, and also by their extensive use in the economic and financial literature. In the interest of clarity, we summarize in Table 3 the key features of the copula models considered in this paper.

Properties of selected copulas				
Copula	Parameters	Asymmetric	Upper tail dep.	Lower tail dep.
Gaussian	$\rho \in [-1,1]$	No	0	0
Student t	$\rho \in [-1,1], \nu > 2$	No	$2t_{\nu+1}\left(-\sqrt{\nu+1}\sqrt{\frac{1-\rho}{1+\rho}}\right)$	$2t_{\nu+1}\left(-\sqrt{\nu+1}\sqrt{\frac{1-\rho}{1+\rho}}\right)$
Gumbel	$\theta \geq 1$	Yes	$2-2^{1/\theta}$	0
Rot. Gumbel	$\theta \geq 1$	Yes	0	$2-2^{1/\theta}$
SJC	$\tau^{U,L} \in [0,1]$	Yes/No	τ^U	τ^L

Table 3: Key features of the five selected copula models. In the tail dependence expression for the Student t copula, $t_{\nu+1}$ refers to the probability density function of the Student t distribution, with $\nu+1$ degrees of freedom. For further details on copulas and their properties, we refer to Joe (1997), Nelsen (1999) and Patton (2013).

The tail dependence terms appearing in Table 3 refer to measures of dependence between extreme events. For the case of positive dependence, the lower and upper tail dependence are defined as

$$\tau^L = \lim_{q\to 0^+} \mathbb{P}(u_{t,1} \leq q \,|\, u_{t,2} \leq q) = \lim_{q\to 0^+} \frac{C(q,q)}{q},$$
$$\tau^U = \lim_{q\to 1^-} \mathbb{P}(u_{t,1} > q \,|\, u_{t,2} > q) = \lim_{q\to 1^-} \frac{1-2q+C(q,q)}{1-q},$$

where q denotes the quantile and C denotes the bivariate copula distribution function. As it turns out, the expressions above have simple closed forms for many copulas. Specifically for the five copulas considered in this paper, tail dependence coefficients are easily computed using the expressions given in Table 3.

Turning to the estimation results, the five copula models are fitted to each residual pair; basing the model selection on the commonly used Akaike Information Criterion (AIC), we report the preferred copula for each model in Table 4. The results reveal that the Gaussian copula is only chosen for 27 out of 96 models, confirming that linear correlation is, in most cases, not enough to describe the dependence between prices in interconnected markets. The Student t copula, which allows for symmetric upper and lower tail dependence, is preferred for 29 of the models. The even more flexible SJC copula, which allows for asymmetric upper and lower tail dependence (nesting the

[6] A Rotated Gumbel (also called Survival Gumbel) is a Gumbel copula fitted to the rotated data (180 degrees rotation).

[7] For details regarding the SJC copula, we refer to Patton (2006).

case of symmetric dependence), is chosen 26 times. Lastly, the Rotated Gumbel copula and the Gumbel copula, which allow for tail dependence in only one of the tails, are preferred in 8 and 6 cases, respectively.

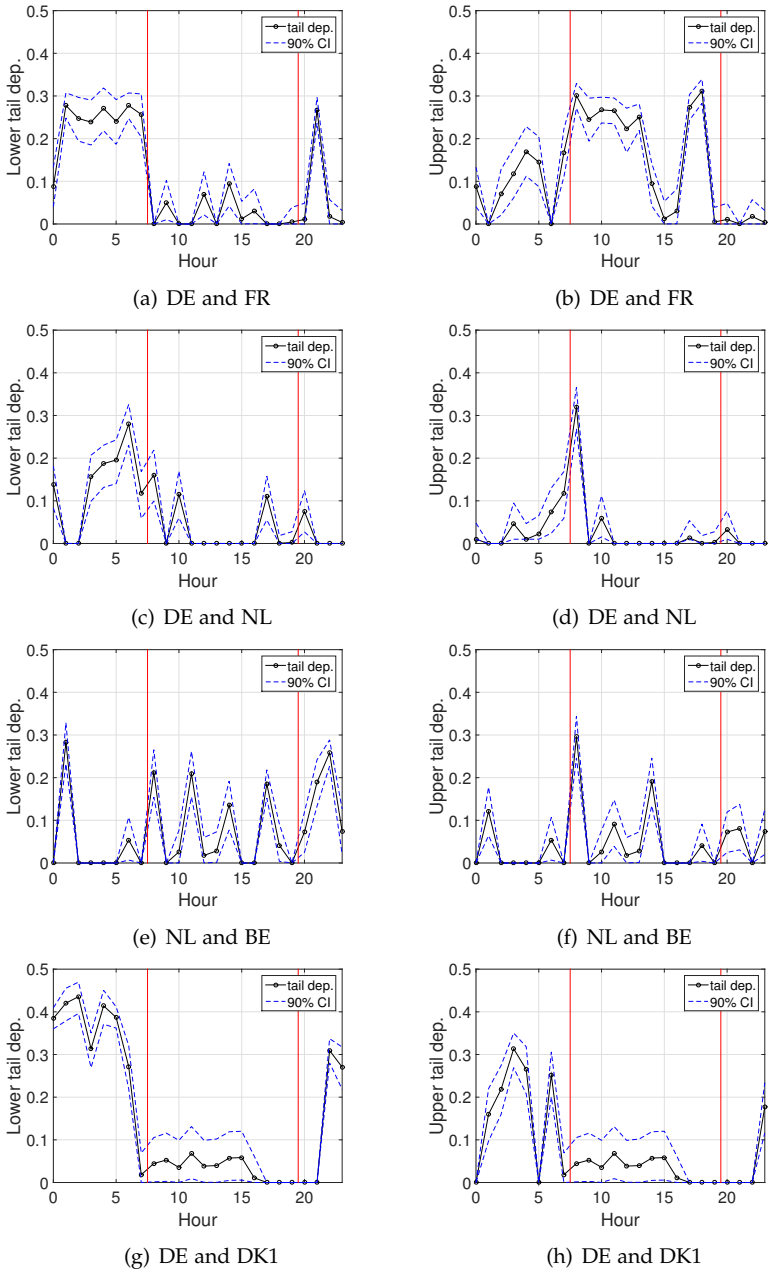
	Optimal copula			
Hour	DE–FR	DE–NL	NL–BE	DE–DK1
0	Student t	SJC	Gaussian	Rot. Gum.
1	Rot. Gum.	Gaussian	SJC	SJC
2	SJC	Gaussian	Gaussian	SJC
3	SJC	SJC	Gaussian	Student t
4	SJC	SJC	Gaussian	SJC
5	SJC	SJC	Gaussian	Rot. Gum.
6	Rot. Gum.	SJC	Student t	SJC
7	SJC	Student t	Gaussian	Student t
8	Gumbel	SJC	SJC	Student t
9	SJC	Gaussian	Gaussian	Student t
10	Gumbel	SJC	Student t	Student t
11	Gumbel	Gaussian	SJC	Student t
12	SJC	Gaussian	Student t	Student t
13	Gumbel	Gaussian	Student t	Student t
14	Student t	Gaussian	SJC	Student t
15	Student t	Gaussian	Gaussian	Student t
16	Student t	Gaussian	Gaussian	Student t
17	Gumbel	SJC	Rot. Gum.	Gaussian
18	Gumbel	Student t	Student t	Gaussian
19	Student t	Student t	Gaussian	Gaussian
20	Student t	SJC	Student t	Gaussian
21	Rot. Gum.	Gaussian	SJC	Gaussian
22	Student t	Gaussian	Rot. Gum.	Rot. Gum.
23	Student t	Gaussian	Student t	SJC

Table 4: Optimal copulas for each hourly model based on AIC.

Tail dependence implied by the optimal copulas

To gain more insight regarding the optimal copulas listed in Table 4, we consider next their implied tail dependence. In Fig. 5, we plot the implied lower and upper tail dependence coefficients for all pairs of prices and all hours of the day. An *i.i.d.* bootstrap procedure was used to obtain the confidence intervals for the estimates (see e.g. Patton (2013) for more details). Moreover, the estimated lower and upper tail dependence coefficients grouped by peak (hours 8-19) and off-peak (hours 20-7) are reported in Table 5.

Recall that the analysis we perform in this section is based on filtered prices belonging to the non-equal price regime. Recall further that non-equal prices occur when there is a need to move more electricity across the border than what the available transmission capacity allows. Hence, we can interpret the lower (upper) tail dependence as being the probability of observing



(a) DE and FR (b) DE and FR

(c) DE and NL (d) DE and NL

(e) NL and BE (f) NL and BE

(g) DE and DK1 (h) DE and DK1

Fig. 5: Estimated lower and upper tail dependence implied by the optimal copulas listed in Table 4. The tail dependence coefficients are obtained using the formulas given in Table 3. The red vertical lines indicate (from left to right) start and end of peak hours.

	Lower tail dependence			
	DE–FR	DE–NL	NL–BE	DE–DK1
Mean peak	0.0217	0.0324	0.0714	0.0338
Mean off peak	0.1830	0.1008	0.0776	0.2687
	Upper tail dependence			
	DE–FR	DE–NL	NL–BE	DE–DK1
Mean peak	0.1899	0.0328	0.0576	0.0338
Mean off peak	0.0658	0.0261	0.0335	0.1170

Table 5: Tail dependence implied by the optimal copulas listed in Table 4, grouped by peak and off-peak hours.

very low (high) prices in two interconnected markets conditional on the full utilization of the interconnector.

According to Fig. 5, the estimates of the preferred copulas translate into tail dependence coefficients that are often significantly different from zero – and this holds for selected hours in all price pairs. Further, a general tendency is that the lower tail dependence is stronger than the upper tail dependence during the off-peak hours (cf. Fig. 5 and Table 5). One explanation for this is that the probability of very low prices is higher during the off-peak hours, where the demand is lower. In a situation where two interconnected areas exhibit a low demand and at the same time a high generation from renewable sources, e.g. wind energy generation, extremely low prices happening simultaneously is a likely outcome.

Another general feature we observe in Fig. 5 and Table 5 is that the lower tail dependence is usually higher during off-peak than during peak. Again here, we expect the lower demand during off-peak to be the main driver of this behavior. The most pronounced difference between peak and off-peak lower tail dependence is found for the DE–DK1 pair. This can be explained by the high wind infeed in both countries as well as similar weather conditions. Most wind turbines in Germany are placed in the northern part, that is, rather close to the Danish border, implying that high wind scenarios are likely to happen simultaneously in both countries.

While low prices happening simultaneously in two neighboring countries can be explained by e.g. low demand and/or high levels of renewables production, high prices occurring simultaneously can be explained by combinations of e.g. high demand, low renewables production, low temperatures, increasing marginal prices of the price setting unit (coal, gas), and plant outages. To provide an economic interpretation of particular events however, like the high upper tail dependence estimate for DE–NL for hour 8, can be very difficult, since a set of very specific combinations of events might have been the cause.

4.5 Simulation

Performing simulations from the proposed regime-switching AR–GARCH copula model, which is a central point in almost all applications, requires a last modeling element: A model for the probabilities of transitions between the two regimes.

Transition matrices

We let the regimes be determined by a Markov chain of order one, with transition probability matrix Π:

$$\Pi = \begin{bmatrix} \pi_{00} & 1-\pi_{00} \\ 1-\pi_{11} & \pi_{11} \end{bmatrix}. \tag{14}$$

The probability of prices being equal conditional on prices being equal at the previous time step is denoted by π_{00}. Similarly, π_{11} denotes the probability of prices not being equal given that they were not equal at the previous time step. Recalling that the regimes are observable, the transition probabilities are simply estimated from the data as

$$\hat{\pi}_{00} = \frac{N_{00}}{N_{00}+N_{01}},$$
$$\hat{\pi}_{11} = \frac{N_{11}}{N_{11}+N_{10}},$$

where N_{jk} denotes the number of times regime j was followed by regime k, for $j = 0, 1$ and $k = 0, 1$.

Simulation procedure

Performing simulations from the proposed model is straightforward, but there are many steps involved. In order to create an overview, we briefly state the necessary steps below.

To generate a pair of (deseasonalized) prices $(y_{t,1}, y_{t,2})$ at time t, we first generate the regime s_t based on the transition probability matrix in Eq. (14).

If **a)** $s_t = 0$, we are in the equal price regime, i.e., $y_{t,1} = y_{t,2}$, and a standardized residual η_t is to be drawn from the univariate marginal distribution, which can be either $N(0,1)$ or Skew $t(0, 1, \nu^{(0)}, \lambda^{(0)})$ depending on the model specification. By applying Eqs. (2) and (1), respectively, a pair $(y_{t,1}, y_{t,2})$ is subsequently obtained.

If **b)** $s_t = 1$, we have that $y_{t,1} \neq y_{t,2}$, and a pair $(u_{t,1}, u_{t,2})$ is generated from the copula[8]. A standardized residual pair $(\eta_{t,1}, \eta_{t,2})$ is obtained by applying the inverse probability integral transform to the generated uniform

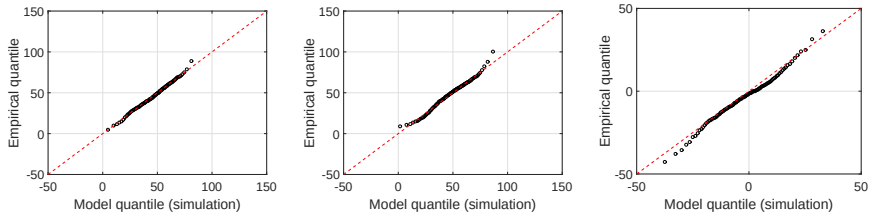
[8]Generating random draws from a copula is based on the inverse of the conditional copula functions (often referred to as h-functions). See e.g. Aas et al. (2009) for more details.

pair $(u_{t,1}, u_{t,2})$. This is done for each uniform individually, based on the univariate marginal distributions, i.e, $\mathrm{N}(0,1)$ or Skew $t(0,1,\nu^{(i)},\lambda^{(i)},)$, for $i = 1,2$. Finally, Eqs. (5) and (4) are employed, respectively, to generate a pair $(y_{t,1}, y_{t,2})$.

In-sample simulations

To provide more evidence for the fit of the model, we perform in-sample simulations. Revisiting the example of DE–FR hour 8, we follow the procedure described in Section 4.5 to generate 1,000 sample paths of German and French prices of length $T - \max(P,Q)$, with T denoting the sample size of the data, and P and Q referring to the order of the model for the conditional mean, cf. Eqs. (1) and (4).[9] In Fig. 6, we show quantile plots of actual prices vs. simulated prices, and quantile plots of actual non-zero spreads vs. simulated non-zero spreads, with the results revealing a satisfactory in-sample fit.

Similar in-sample fits were obtained for the other pairs regarding quantile plots of the individual price series. For the spreads on the other hand, the model does not always perform very well when considering matching both tails. For some hours of the day regarding DE–NL, the model assigns slightly more mass to the right tail of the distribution, while being able to capture the left tail nicely. As can be seen in Fig. 2 and Table 1, cases where the Dutch price is above the German price are either extremely few or none at all, explaining perhaps the inability of the model to provide a sufficient explanation of the right tail.



(a) DE price (b) FR price (c) DE–FR spread

Fig. 6: Quantile plots of price and (non-zero) spread data vs. simulated prices and (non-zero) spreads from the fitted model with skew t margins and a Gumbel copula cf. Table 4.

[9] To start the simulation, we use the first $\max(P,Q)$ prices. Regarding $\hat{\sigma}^2_{0,i}$, for $i = 1,2$, we use the sample variance of the data.

5 Financial transmission rights

The presence of heavy tails in the individual price distributions, asymmetric dependence and tail dependence between prices can have significant implications when considering derivatives whose payoffs depend on the price spread. To investigate such effects, we consider in this section the pricing of financial transmission rights as a first application.

A financial transmission right (FTR) is a spread option written on the hourly day-ahead electricity price difference in two interconnected areas. It gives the option holder the right to "transfer electricity across the border" if the price difference is positive for the corresponding hour. In reality, no physical cross-border transfer of electricity is undertaken as the rights are settled financially, hence the name *financial* transmission right.

A particular feature of FTRs is that delivery is limited to certain periods, such as an entire year or month. When entering e.g. a monthly FTR, the option holder actually enters a string of hourly spread options, and has the right to exercise at each hour in that month. Also interesting is that FTRs are linked to the actual available capacity of the interconnector, and they are offered by the transmission system operators (TSO) through explicit auctions.[10]

Let us denote by $V(t_0, T)$ the value of a financial transmission right at time t_0, with delivery at T. $V(t_0, T)$ can be obtained by the usual practice of computing the discounted value of the expected payoff, where the expectation is taken under a pricing measure $\mathbb{Q}$ (see e.g. Benth et al. (2008)):

$$V(t_0, T) = \mathrm{e}^{r(T-t_0)} \mathbb{E}^{\mathbb{Q}}_{t_0}[\max(P^1_T - P^2_T, 0)]. \tag{15}$$

In Eq. (15), P^i_T denotes the hourly day-ahead electricity price corresponding to the time period T in area i, where $i = 1, 2$. Furthermore, r denotes the risk-free rate, which we for simplicity set to zero in the following analysis.

The value of e.g. a monthly FTR is simply obtained as the sum of all individual hourly spread option values in that month, which in turn are obtained using Eq. (15). Thus, the holder of a monthly FTR will have the right to exchange electricity in area 2 for electricity in area 1, if the price difference is positive, during each hour of that month.

Having established a modeling framework for pairs of day-ahead electricity prices in interconnected areas, FTR prices can be obtained by Monte Carlo simulation, following the simulation procedure described in Section 4.5. We note however that computing expectations based on simulations from a model that has been fitted to historical data would yield an option price under the objective measure $\mathbb{P}$, and not the pricing measure $\mathbb{Q}$. We will make the simplifying assumption of $\mathbb{P} = \mathbb{Q}$, arguing that a change of measure would not alter the overall conclusions regarding the impacts of heavy tails and choice

[10] For more details we refer to Füss et al. (2015).

of copula on option prices. For further discussions regarding measure change in the context of electricity markets we refer to Burger et al. (2004), Benth and Meyer-Brandis (2009), and Benth and Schmeck (2014).

To investigate the effects of heavy tails and choice of copula model, we consider here the pricing of FTRs for the months June and December 2015 as examples. The valuation date is set to $t_0 = 20$ May 2015, which corresponds to the end of our sample. Since we model each hour of the day separately, we will limit ourselves to pricing monthly FTRs with delivery during one single hour of the day, for all days in June and December. This way, we avoid averaging effects to some extent and showcase the different behavior among the hours of the day. Specifically, we consider hour 2, 8, 14 and 20, for all pairs of prices, and flows in both directions. The pricing formula for e.g. the flow direction DE→FR during each hour 2 in June is given by

$$V^{\mathrm{DE}\to\mathrm{FR}}(t_0, T_1, T_2, h = 2) = \sum_{t=T_1}^{T_2} \mathbb{E}_{t_0}^{\mathbb{P}}[\max(P_t^{FR} - P_t^{DE}, 0)],$$

where t_0 =20 May 2015, T_1 =1 June 2015, and T_2 =30 June 2015. P_t^{FR} and P_t^{DE} denote the hour 2 prices in France and Germany, respectively, for a given day $t \in [T_1, T_2]$.

We perform option price calculations using three different models: First, a model with normal margins and a Gaussian copula (Model 1); second, a model with skew t margins and a Gaussian copula (Model 2), and third, a model with skew t margins and the preferred copula cf. Table 4 (Model 3). By construction, Model 1 is a naive model since it imposes no heavy tails, no asymmetry and no tail dependence. Model 2 on the other hand allows for heavy tails but imposes symmetry and no tail dependence, while Model 3 allows for heavy tails, and can allow for tail dependence and asymmetry (note that preferred copulas are in some cases the Gaussian copula according to Table 4). In the interest of clarity, we summarize the characteristics of the above-mentioned models in Table 6.

	Model 1 Normal margins & Gaussian copula	**Model 2** Skew t margins & Gaussian copula	**Model 3** Skew t margins & optimal copula
Heavy tails (margins)	×	✓	✓
Tail dependence	×	×	× or ✓
Asymmetric dependence	×	×	× or ✓

Table 6: Model characteristics for the three different models used to compute FTR prices. The optimal copula entering Model 3 refers to the preferred specification cf. Table 4; this model may or may not allow for tail dependence and asymmetric dependence.

The pricing results are displayed in Tables 7 – 8, and are based on 200,000 Monte Carlo simulations from the fitted regime-switching AR–GARCH cop-

ula models from Section 4. Price 1, 2 and 3 in Tables 7 – 8 refer to prices obtained from Model 1, 2 and 3, respectively, and are given in EUR/MWh. Δ_1 refers to the relative difference between Price 1 and 2, measuring the effect of introducing heavy tails in the model. Δ_2 refers to the relative difference between Price 2 and 3, measuring the effect of using a better fitting copula, which in contrast to the Gaussian copula, allows for tail dependence and in some cases asymmetry. Notice that if the Gaussian copula was chosen as the optimal copula cf. Table 4, no price for the corresponding option was computed with Model 3, since Model 3 then corresponds to Model 2.

Regarding the robustness of option price estimates, we note that the same option prices were estimated based on 100,000 and 200,000 Monte Carlo simulations, yielding estimates that are generally accurate to two or three decimal places. Furthermore, the same random number generator was used to compute option prices across competing models, thus allowing us to solely concentrate on the effects induced by the different marginal distributions and copulas.

According to Tables 7 and 8, introducing heavy tails can have a significant effect in terms of magnitude, with the maximum reduction in the option price being as high as 45%, and the maximum increase reaching 28%. In terms of sign, the effects are mixed, which might be surprising at first. In the case of e.g. an out-of-the-money option written on one underlying asset, heavy tails would imply an increase in the option price, since more mass is assigned to extreme events. In the case of spread options however, where the payoff depends on the price differential, heavy tails might or might not cancel out, leading to both decreases and increases. Another factor influencing both the sign and the magnitude is the dependence implied by the Gaussian copula, which is not exactly the same for the models with normal margins as for the those with skew t margins. To illustrate this variation, we compute the Kendall's τ implied by the models specified in Table 6. The results are reported in Table 9, revealing differences which naturally impact the option prices.

Also noticeable in the context of heavy tails is that when comparing the same option in the two different flow directions, the effect is always magnified for the direction that is least in-the-money. The more difference there is between the moneyness of the options, the more difference there is in the effect of heavy tails. This is of course expected, since out-of-the-money options are more sensitive to changes in the tail behavior of a distribution.

When considering the effects of using a better fitting copula, the results generally imply a reduction in the option prices, with a maximum drop amounting to 13%. As with heavy tails, the effects are magnified for out-of-the-money options. These findings are expected if considering the isolated effect of tail dependence: Allowing prices to move together in extreme events lowers the probability of obtaining extreme price differences in the simulated

	Delivery during June									
	Price 1	Price 2	Price 3	Δ_1	Δ_2	Price 1	Price 2	Price 3	Δ_1	Δ_2
	FR → DE					**DE → FR**				
Hour 2	3.0621	2.8952	2.8714	-5.45%	-0.82%	1.9482	1.7830	1.7608	-8.48%	-1.25%
Hour 8	2.5707	2.5518	2.5298	-0.74%	-0.86%	1.3009	1.2818	1.2657	-1.47%	-1.26%
Hour 14	0.7294	0.6528	0.6365	-10.50%	-2.50%	5.9410	5.8640	5.8461	-1.30%	-0.31%
Hour 20	3.7612	3.7539	3.7451	-0.19%	-0.23%	0.9362	0.9282	0.9191	-0.85%	-0.98%
	NL → DE					**DE → NL**				
Hour 2	0.7644	0.4850	–	-36.55%	–	6.8804	6.6040	–	-4.02%	–
Hour 8	0.1948	0.2106	0.1939	8.11%	-7.93%	5.1074	5.1223	5.1198	0.29%	-0.05%
Hour 14	0.1297	0.0898	–	-30.76%	–	12.4058	12.3658	–	-0.32%	–
Hour 20	0.2574	0.2616	0.2677	1.63%	2.33%	6.0492	6.0517	6.0680	0.04%	0.27%
	BE → NL					**NL → BE**				
Hour 2	2.0140	1.9936	–	-1.01%	–	0.8381	0.8169	–	-2.53%	–
Hour 8	1.1731	1.1488	1.1253	-2.07%	-2.05%	1.5565	1.5307	1.5103	-1.66%	-1.33%
Hour 14	4.3010	4.3039	4.2978	0.07%	-0.14%	0.2761	0.2776	0.2696	0.54%	-2.88%
Hour 20	2.7042	2.6650	2.6495	-1.45%	-0.58%	0.8726	0.8322	0.8163	-4.63%	-1.91%
	DK1 → DE					**DE → DK1**				
Hour 2	1.5204	1.3093	1.2060	-13.88%	-7.89%	3.9435	3.7318	3.6144	-5.37%	-3.15%
Hour 8	4.4038	4.3768	4.3417	-0.61%	-0.80%	2.5883	2.5616	2.5260	-1.03%	-1.39%
Hour 14	1.2316	1.1128	1.0772	-9.65%	-3.20%	5.7916	5.6728	5.6374	-2.05%	-0.62%
Hour 20	6.0021	6.0090	–	0.11%	–	0.9293	0.9357	–	0.69%	–

Table 7: Option prices in EUR/MWh with delivery in June 2015, during the specified hour. Price 1 refers to an option price obtained from models with normal margins and a Gaussian copula. Price 2 refers to models with skew t margins and a Gaussian copula, and Price 3 refers to models with skew t margins and the optimal copula cf. Table 4. Δ_1 refers to the difference between Price 1 and 2, while Δ_2 refers to the difference between Price 2 and 3.

	Delivery during December									
	Price 1	Price 2	Price 3	Δ_1	Δ_2	Price 1	Price 2	Price 3	Δ_1	Δ_2
	FR → DE					**DE → FR**				
Hour 2	0.1771	0.1155	0.1113	-34.78%	-3.64%	11.0058	10.9432	10.9545	-0.57%	0.10%
Hour 8	0.1885	0.1909	0.1933	1.27%	1.26%	6.9024	6.9054	6.8953	0.04%	-0.15%
Hour 14	0.4627	0.4193	0.4139	-9.38%	-1.29%	7.6324	7.5900	7.5841	-0.56%	-0.08%
Hour 20	0.1303	0.1363	0.1332	4.60%	-2.27%	9.1461	9.1526	9.1494	0.07%	-0.03%
	NL → DE					**DE → NL**				
Hour 2	0.3427	0.1868	–	-45.49%	–	9.8521	9.6975	–	-1.57%	–
Hour 8	0.0986	0.1261	0.1092	27.89%	-13.40%	6.4267	6.4546	6.4380	0.43%	-0.26%
Hour 14	0.2723	0.2016	–	-25.96%	–	10.3227	10.2520	–	-0.68%	–
Hour 20	0.1058	0.1145	0.1178	8.22%	2.88%	8.0795	8.0894	8.0925	0.12%	0.04%
	BE → NL					**NL → BE**				
Hour 2	0.7064	0.7091	–	0.38%	–	2.3100	2.3129	–	0.13%	–
Hour 8	0.6977	0.6845	0.6644	-1.89%	-2.94%	2.4609	2.4479	2.4278	-0.53%	-0.82%
Hour 14	1.8068	1.7556	1.7470	-2.83%	-0.49%	1.3769	1.3271	1.3136	-3.62%	-1.02%
Hour 20	1.1500	1.1291	1.1186	-1.82%	-0.93%	2.5007	2.4806	2.4697	-0.80%	-0.44%
	DK1 → DE					**DE → DK1**				
Hour 2	1.3579	1.1409	1.0204	-15.98%	-10.56%	4.6786	4.4630	4.3492	-4.61%	-2.53%
Hour 8	7.0992	7.1558	7.1327	0.80%	-0.32%	1.4306	1.4904	1.4680	4.18%	-1.66%
Hour 14	4.0953	4.0324	3.9957	-1.54%	-0.91%	2.3401	2.2786	2.2422	-2.63%	-1.29%
Hour 20	8.4618	8.5114	–	0.59%	–	0.5290	0.5807	–	9.77%	–

Table 8: Option prices in EUR/MWh with delivery in December 2015, during the specified hour. Price 1 refers to an option price obtained from models with normal margins and a Gaussian copula. Price 2 refers to models with skew t margins and a Gaussian copula, and Price 3 refers to models with skew t margins and the optimal copula cf. Table 4. Δ_1 refers to the difference between Price 1 and 2, while Δ_2 refers to the difference between Price 2 and 3.

	Kendall's τ_1	Kendall's τ_2	Kendall's τ_3
		DE and FR	
Hour 2	0.2454	0.2315	0.2116
Hour 8	0.2446	0.2434	0.2350
Hour 14	0.1961	0.2093	0.2126
Hour 20	0.2113	0.2124	0.2132
		DE and NL	
Hour 2	0.1802	0.1917	–
Hour 8	0.3149	0.2961	0.2742
Hour 14	0.1279	0.1499	–
Hour 20	0.1445	0.1397	0.1242
		NL and BE	
Hour 2	0.2658	0.2706	–
Hour 8	0.3035	0.2951	0.2812
Hour 14	0.2398	0.2277	0.2137
Hour 20	0.1930	0.1929	0.1941
		DE and DK1	
Hour 2	0.3984	0.3545	0.3452
Hour 8	0.3243	0.3070	0.3099
Hour 14	0.2883	0.3071	0.3130
Hour 20	0.2546	0.2482	–

Table 9: Kendall's τ implied by the three models specified in Table 6, for the same pairs of prices and hours of the day as in Tables 7 – 8.

future payoffs, which ultimately lowers the option price. Furthermore, the moneyness of the option plays an important role in this context, since it decides the importance of the joint tail behavior for the option price.

Although tail dependence is included in the effect measured by Δ_2, we stress that tail dependence is not the only influencing factor, and hence its impact cannot be separated and measured explicitly. Equally important is that Δ_2 can include the effects of asymmetry, and simply measures the impact of assuming a dependence structure that is different from the Gaussian one. To emphasize this point, we again refer to Kendall's τ computed in Table 9, which makes the results comparable across copula models. According to this measure, some differences are observed in the dependence implied by Models 2 and 3. With everything else being equal, a reduction in Kendall's τ implies an increase in the option price, which would counteract the effects of tail dependence. Similarly, an increase in Kendall's τ implies a decrease in the option price, possibly adding to the reduction implied by tail dependence.

Overall, heavy tails and a better fitting copula are most important to consider in the context of out-of-the-money options - and as we can see from Tables 7 and 8, there are many such options for the interconnected electricity markets we analyze here.

Lastly, we note that the seasonal behavior of prices and the electricity flow patterns discussed in Section 2 are of course reflected in the option prices. For

the pair DE–FR for example, we observe that the direction FR→DE is more in-the-money than DE→FR for the summer month June, during hours 2, 8 and 20. During hour 14, when solar production in Germany is at its peak, the situation reverses. During the winter month December, FR→DE is deep out-of-the-money compared to DE→FR, which is due to increased demand and a high wind power production in Germany. Also for the pair DE–NL, the fact that Netherlands is a significant importer of German power is reflected in the option prices, with NL→DE being very cheap compared to DE→NL.

6 Forecasting of tail quantiles

To provide more empirical evidence for the proposed modeling framework, we consider in this section the forecasting of tail quantiles as a second application. This can be of interest to parties with physical or financial positions in two interconnected markets, portfolio managers, and also risk managers who wish to asses worst-case scenarios.

Clearly, the proposed regime-switching AR–GARCH copula model is a sophisticated approach to forecasting the tail quantiles of the price spread distributions. Therefore, aside from wishing to illustrate the effects of heavy tails and a copula that allows for asymmetry and tail dependence, we also find it relevant to compare the performance of the proposed copula framework with a much simpler approach. We thus perform the forecasting exercise with the following three models:

- First, we consider the regime-switching AR–GARCH model in Eqs. (1) – (6) with skew t margins and a SJC copula. This model will be referred to as Skew t/SJC.
- Second, we consider the regime-switching AR–GARCH model in Eqs. (1) – (6) with normal margins and a Gaussian copula. This model will be referred to as N/Gaussian.
- Third, we consider a naive approach that is very different from the copula framework. Specifically, we apply a regime-switching AR(1) model to the deseasonalized spread z:

  $$z_t = \phi_{s_t} z_{t-1} + \epsilon_{t,s_t}, \qquad \epsilon_t \sim \mathrm{N}(0, \sigma^2_{s_t}).$$

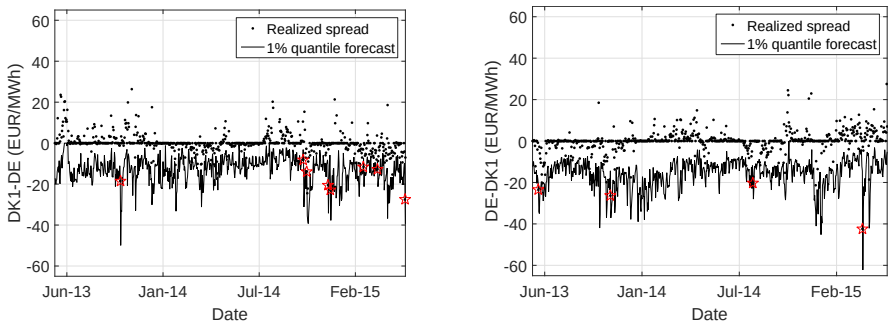
  To correct the spread time series for deterministic seasonality, we apply the seasonal function in Eq. (13). The regimes in this univariate model are the same as before: Whenever $s_t = 0$, we are in the zero spread (equal prices) regime; if $s_t = 1$, we are in the non-zero spread (non-equal prices) regime. Further, we again let the regimes be determined by a Markov chain of order one. As far as the regimes are concerned,

the only difference from the copula approach is that whenever $s_t = 0$ in the univariate spread model, all the model parameters are zero. This model will be referred to as Univariate.

To evaluate the forecasting power of each model, we perform out-of-sample day-ahead forecasts of the tail quantiles for the pair DE–DK1 at hours $0, 1, \ldots, 6$. This choice is motivated by the high degree of asymmetry and tail dependence obtained for this pair during most off-peak hours, cf. Fig. 5. On one hand, forecasting with the Skew t/SJC and the N/Gaussian models enable us to investigate the effects of heavy tails and those of using a copula that allows for asymmetry and tail dependence. On the other hand, employing a very different and much simpler model (the Univariate model) reveals whether or not the copula modeling framework is at all advantageous in this context.

Specifically, we consider the 5%, 1% and 0.5% quantiles of the spread distributions for a period of two years (730 days), corresponding to 21 May 2013 - 20 May 2015. We employ an "expanding window" estimation, where an out-of-sample forecast for period $t + s$ is obtained by fitting the models to data in the interval $[1, t + s - 1]$. The models are re-estimated at each step as we progress through the out-of-sample period, and each forecast is based on 100,000 Monte Carlo simulations.

The coverage results for the forecasts are summarized in Tables 10 and 11, where both tails of the spread distributions are considered, i.e. the 5%, 1% and 0.5% quantile forecasts for the distribution of both DE-DK1 and DK1-DE prices. To provide a visual example, we also plot in Fig. 7 the results obtained for hour 0 with the Skew t/SJC model.



(a) DK1-DE Hour 0

(b) DE-DK1 Hour 0

Fig. 7: Out-of-sample day-ahead forecasts for the 1% quantile of spread distributions with a Skew t/SJC model. The red stars signal exceedances.

While the coverage results in Tables 10 and 11 are informative, they are not sufficient as to draw conclusions on model adequacy. To asses the per-

		DK1-DE							
		$q = 0.05 (36.5)$				$q = 0.01 (7.3)$		$q = 0.005 (3.65)$	
		Coverage	p^{uc}-val.	p^{ind}-val.	p^{cc}-val.	Coverage	p^{uc}-val.	Coverage	p^{uc}-val.
Hour 0	Skew t/SJC	0.0548 (40)	0.5523	0.2416	0.4222	0.0110 (8)	0.7947	0.0082 (6)	0.2582
	N/Gaussian	0.0466 (34)	0.6738	0.7393	0.8659	0.0123 (9)	0.5392	0.0096 (7)	0.1182
	Univariate	0.0274 (20)	**0.0023**	0.5721	**0.0082**	0.0055 (4)	0.1805	0.0055 (4)	0.8544
Hour 1	Skew t/SJC	0.0644 (47)	0.0855	0.5678	0.1937	0.0178 (13)	0.0556	0.0096 (7)	0.1182
	N/Gaussian	0.0644 (47)	0.0855	0.5678	0.1937	0.0192 (14)	**0.0266**	0.0123 (9)	**0.0180**
	Univariate	0.0288 (21)	**0.0045**	0.6310	**0.0157**	0.0082 (6)	0.6204	0.0068 (5)	0.5007
Hour 2	Skew t/SJC	0.0616 (45)	0.1604	0.6027	0.3260	0.0137 (10)	0.3396	0.0096 (7)	0.1182
	N/Gaussian	0.0616 (45)	0.1604	0.6027	0.3260	0.0137 (10)	0.3396	0.0110 (8)	**0.0485**
	Univariate	0.0301 (22)	**0.0082**	0.6912	**0.0281**	0.0123 (9)	0.5392	0.0041 (3)	0.7268
Hour 3	Skew t/SJC	0.0493 (36)	0.9389	0.5067	0.7998	0.0178 (13)	0.0556	0.0110 (8)	**0.0485**
	N/Gaussian	0.0411 (30)	0.2588	0.8203	0.5151	0.0178 (13)	0.0556	0.0123 (9)	**0.0180**
	Univariate	0.0342 (25)	**0.0396**	0.2704	0.0656	0.0096 (7)	0.9135	0.0055 (4)	0.8544
Hour 4	Skew t/SJC	0.0644 (47)	0.0855	0.1015	0.0596	0.0205 (15)	**0.0120**	0.0123 (9)	**0.0180**
	N/Gaussian	0.0562 (41)	0.4480	0.2787	0.4171	0.0205 (15)	**0.0120**	0.0123 (9)	**0.0180**
	Univariate	0.0356 (26)	0.0618	0.0711	**0.0343**	0.0123 (9)	0.5392	0.0110 (8)	**0.0485**
Hour 5	Skew t/SJC	0.0521 (38)	0.7936	0.0526	0.1478	0.0164 (12)	0.1087	0.0096 (7)	0.1182
	N/Gaussian	0.0438 (32)	0.4403	0.0562	0.1200	0.0192 (14)	**0.0266**	0.0110 (8)	**0.0485**
	Univariate	0.0329 (24)	**0.0244**	**0.0448**	**0.0106**	0.0110 (8)	0.7947	0.0110 (8)	**0.0485**
Hour 6	Skew t/SJC	0.0397 (29)	0.1899	0.6204	0.2582	0.0096 (7)	0.9135	0.0082 (5)	0.5007
	N/Gaussian	0.0260 (19)	**0.0011**	NaN	NaN	0.0096 (7)	0.9135	0.0096 (7)	0.1182
	Univariate	0.0521 (38)	0.7936	0.4205	0.6986	0.0123 (9)	0.5392	0.0096 (7)	0.1182

Table 10: Coverage and test results for out-of-sample day-ahead quantile forecasts of DK1-DE. The coverage results are given with number of exeedences in parenthesis, and the reported p-values correspond to the tests in Christoffersen (1998) for unconditional coverage (uc), independence (ind) and conditional coverage (cc). We consider the 5%, 1% and 0.5% quantiles of the spread distribution at seven hours of the day. The results are based on an out-of-sample period of 730 observations, from 21 May 2013 to 20 May 2015. The p-values in bold highlight the cases where independence and correct (un)conditional coverage are rejected at a 5% significance level.

		DE-DK1							
		$q = 0.05 (36.5)$				$q = 0.01 (7.3)$		$q = 0.005 (3.65)$	
		Coverage	p^{uc}-val.	p^{ind}-val.	p^{cc}-val.	Coverage	p^{uc}-val.	Coverage	p^{uc}-val.
Hour 0	Skew t/SJC	0.0274 (20)	**0.0023**	0.1141	**0.0028**	0.0055 (4)	0.1805	0.0055 (4)	0.8544
	N/Gaussian	0.0219 (16)	**0.0001**	**0.0448**	**0.0001**	0.0055 (4)	0.1804	0.0055 (4)	0.8544
	Univariate	0.0233 (17)	**0.0002**	0.4067	**0.0008**	0.0096 (7)	0.9135	0.0082 (6)	0.2582
Hour 1	Skew t/SJC	0.0233 (17)	**0.0002**	0.4067	**0.0008**	0.0055 (4)	0.3447	0.0027 (2)	0.3447
	N/Gaussian	0.0178 (13)	**0.0000**	NaN	NaN	0.0068 (5)	0.3662	0.0041 (3)	0.7268
	Univariate	0.0288 (21)	**0.0045**	NaN	NaN	0.0123 (9)	0.5392	0.0082 (6)	0.2582
Hour 2	Skew t/SJC	0.0356 (26)	0.0618	NaN	NaN	0.0082 (6)	0.6204	0.0041 (3)	0.7268
	N/Gaussian	0.0301 (22)	**0.0082**	NaN	NaN	0.0082 (6)	0.6204	0.0068 (5)	0.5007
	Univariate	0.0342 (25)	**0.0396**	NaN	NaN	0.0192 (14)	**0.0266**	0.0164 (12)	**0.0005**
Hour 3	Skew t/SJC	0.0438 (32)	0.4403	0.7077	0.6921	0.0123 (9)	0.5392	0.0055 (4)	0.8544
	N/Gaussian	0.0370 (27)	0.0929	0.3560	0.1592	0.0123 (9)	0.5392	0.0068 (5)	0.5007
	Univariate	0.0384 (28)	0.1351	NaN	NaN	0.0205 (15)	**0.0120**	0.0123 (9)	**0.0180**
Hour 4	Skew t/SJC	0.0452 (33)	0.5516	0.2504	0.4327	0.0110 (8)	0.7947	0.0055 (4)	0.8544
	N/Gaussian	0.0384 (28)	0.1351	0.4034	0.2309	0.0110 (8)	0.7947	0.0096 (7)	0.1182
	Univariate	0.0534 (39)	0.6681	0.2077	0.4125	0.0192 (14)	**0.0266**	0.0123 (9)	**0.0180**
Hour 5	Skew t/SJC	0.0479 (35)	0.8041	0.3305	0.6040	0.0096 (7)	0.9135	0.0014 (1)	0.0995
	N/Gaussian	0.0425 (31)	0.3423	0.1835	0.2631	0.0110 (8)	0.7947	0.0027 (2)	0.3447
	Univariate	0.0452 (33)	0.5516	0.6787	0.7687	0.0192 (14)	**0.0266**	0.0137 (10)	**0.0061**
Hour 6	Skew t/SJC	0.0260 (19)	**0.0011**	0.5149	**0.0041**	0.0055 (4)	0.1805	0.0014 (1)	0.0995
	N/Gaussian	0.0164 (12)	**0.0000**	0.1857	**0.0000**	0.0055 (4)	0.1805	0.0041 (3)	0.7268
	Univariate	0.0671 (49)	**0.0423**	0.1184	**0.0376**	0.0329 (24)	**0.0000**	0.0288 (21)	**0.0000**

Table 11: Coverage and test results for out-of-sample day-ahead quantile forecasts of DE-DK1. The coverage results are given with number of exeedences in parenthesis, and the reported p-values correspond to the tests in Christoffersen (1998) for unconditional coverage (uc), independence (ind) and conditional coverage (cc). We consider the 5%, 1% and 0.5% quantiles of the spread distribution at seven hours of the day. The results are based on an out-of-sample period of 730 observations, from 21 May 2013 to 20 May 2015. The p-values in bold highlight the cases where independence and correct (un)conditional coverage are rejected at a 5% significance level.

formance of the models, we hence conduct formal statistical testing. We consider the three likelihood ratio tests proposed in Christoffersen (1998), namely the test of correct unconditional coverage (LR_{uc}), independence (LR_{ind}), and correct conditional coverage (LR_{cc}). While LR_{uc} and LR_{ind} are concerned with exceedance proportions and the order of exceedance occurrences, respectively, LR_{cc} is a complete test that enables the joint testing of coverage and independence. When conditioning on the fist observation – as we do here – Christoffersen (1998) shows that $LR_{cc} = LR_{uc} + LR_{ind}$.[11]

For the 5% out-of-sample quantile forecasts, we report in Tables 10 and 11 the p-values obtained by performing all three tests mentioned above. When no concecutive exceedances are present, we note that the independence test and hence the conditional coverage test are not defined. In such situations, the p-value is set to NaN. Since the case of no consecutive exceedances is extremely common for the 1% and 0.5% quantile forecasts, we confine ourselves to reporting the results corresponding to the test of unconditional coverage for these quantiles.

Generally, the 5% out-of-sample quantile forecasts pass the independence test across all models. We only encounter two cases where independence is rejected: Once for the N/Gaussian model, and once for the Univariate model. Although not reported in Tables 10 and 11, the independence test is also passed for the remaining quantile forecasts, whenever this test is defined. Thus, it is to a high extent the rejection of unconditional coverage that leads to the rejection of conditional coverage.

To provide a clearer picture of the model performance results, we report in Table 12 model acceptance percentages that are computed for different groupings: Overall, for each one of the tails, across the three quantiles, and across the seven hours of the day. For each quantile, hour of the day and tail of the spread distribution, a model is labeled as "accepted" if all three tests are passed (given that they are defined).

The results in Table 12 suggest that the Skew t/SJC model performs the best, with the overall acceptance and the acceptance across tails and quantiles always being superior to the other models. As far as acceptance across the hours of the day is concerned, the Skew t/SJC model always outperforms the Univariate model, and performs better or at least as good as the N/Gaussian model. When comparing the N/Gaussian model with the Univariate model, we find stronger empirical evidence for the former, which performs better in all but two instances cf. Table 12. Interestingly, we also point out that while the performance of the Skew t/SJC model is consistent across the two tails of the distribution, the performance of the other models is asymmetric. Specifically, the N/Gaussian performs much better when considering the 5%,

[11] For technical details and a thorough description of all three tests we refer the interested reader to Christoffersen (1998).

	Acceptance percentages					
				Across quantiles		
	Overall	DK1-DE	DE-DK1	0.05	0.01	0.005
Skew t/SJC	**86**	86	86	79	93	86
N/Gaussian	**69**	57	81	64	79	64
Univariate	**48**	62	33	29	64	50
			Across hours			
	H 0	H 1	H 2	H 3	H 4	H 5	H 6
Skew t/SJC	83	83	100	83	67	100	83
N/Gaussian	83	50	67	83	67	67	67
Univariate	67	67	33	50	33	33	50

Table 12: Acceptance percentages based on the forecast results in Table 11. A model is labeled as "accepted" if *all* tests are passed, conditional on the test output being different from NaN.

1% and 0.5% quantiles of the DE-DK1 distribution as opposed to the DK1-DE distribution, while the reverse is true for the Univariate model.

Overall, the two-year out-of-sample evaluations of tail quantile forecasts for the DK1–DE pair strengthen further the previously drawn conclusions regarding the usefulness of the proposed copula framework, and the importance of capturing heavy tails, asymmetric dependence and tail dependence. The out-of-sample performance of the Skew t/SJC model is reasonable when evaluated on its own, and superior when compared with the considered alternative models. While the results are encouraging, they do leave place for improvement. There are many changes and extensions to the proposed regime-switching AR–GARCH copula model that are easily implementable and may improve the tail quantile forecast performance. For example, we mention replacing GARCH with EGARCH (see Knittel and Roberts (2005) and Chan and Gray (2006) for the rationale behind this), considering even more flexible copulas, e.g. copula mixtures, and allowing the copula family to change through the out-of-sample forecasting period. We leave however such comparisons for future study.

7 Conclusion and outlook

In this paper we propose a regime-switching AR–GARCH copula for modeling pairs of day-ahead electricity prices in interconnected markets. The regimes are observable, distinguishing the case of identical prices from the case of non-identical prices. The AR–GARCH filters account for the serial dependence in the conditional mean and the conditional variance, and the copula is introduced to easily relax the assumption of normal margins, and to investigate the dependency between prices beyond linear correlation.

Aside from being able to accommodate the structural change in the joint price dynamics that followed the price coupling of areas, we obtain strong in-sample evidence for the modeling framework. The same type of regime-switching AR–GARCH filter is applied to four pairs of prices and each hour of the day, a total of 96 models, with all results confirming the superiority of the skewed t distribution over the normal distribution, and also whiteness of residuals. The copula analysis performed on the filtered data belonging to the non-identical price regime reveals that the characterization of the dependence structure with one single number, namely the linear correlation, is indeed not sufficient in most of the cases. In the context of copulas, we find significant evidence of tail dependence in all pairs of prices we consider. Moreover, we also find evidence of asymmetry in the dependence structure.

As a first application we consider the pricing of financial transmission rights, where the effects of heavy tails and a better fitting copula (compared to the Gaussian one) are highlighted. We find heavy tails to have a significant effect on option prices in terms of magnitude, but are of mixed sign. The moneyness of the option influences the magnitude of the effect, with out-of-the-money options being most influenced. Like in the case of heavy tails, the choice of copula model affects out-of-the-money options the most. Generally, lower option prices are produced with the preferred copula specification compared to the Gaussian one.

As a second application, the tail quantile forecast performance of three competing models is evaluated. For the forecast exercise, two of the models are copula-based, while the last model represents a much simpler and very different approach. Forecast performance is evaluated based on the likelihood ratio tests proposed in Christoffersen (1998), with the results revealing that the best performing model is a copula-based model where heavy tails, asymmetric dependence and tail dependence are incorporated.

Although major advances were made over the past years toward achieving the goal of a single European electricity market, the completion of such a project lies far out in the future. Developing models for the joint dynamics of prices in interconnected markets that comply with the present institutional framework is hence of great relevance. However, just as the price coupling of areas has led to a structural change, ongoing developments challenge further keeping the models up-to-date. For example, recall that our sample window ends on 20 May 2015, which represents the change in market coupling algorithm for the CWE region (transitioning from the ATC to the more efficient FB methodology). We wish to briefly shed some light on implications caused by this change in the context of the future use of our proposed model. To this end, we enlarge our sample window, allowing it to end on 25 September 2016, and fit all regime-switching AR–GARCH models to this data, confirming of course the goodness-of-fit of the skew t distribution and whiteness of residuals. Then, we perform an analysis similar to that in Sec. 4.3, namely,

we test for a break in the rank correlation at $t^* = 20$ May 2015. Specifically, we wish to test

$$H_0 : \rho^{ATC} = \rho^{FB} \quad \text{vs.} \quad H_1 : \rho^{ATC} \neq \rho^{FB},$$

where ρ^{ATC} denotes the rank correlation between filtered prices belonging to the non-identical price regime in the interval $[1, t^*]$, and ρ^{FB} denotes the rank correlation in the interval $(t^*, T]$.

We apply the test to all four pairs of prices, although the change in methodology does not concern the DE–DK1 pair. The results are presented in Table 13, and show significant evidence against the null of equal rank correlations for the CWE pairs. In the case of DE–FR, the null is rejected 17 out of 24 times in favor of the alternative hypothesis. In the case of DE–NL and NL–BE, the null is rejected 8 and 12 times, respectively. As expected, there is not enough evidence for a break in the case of DE–DK1 at any hour of the day.

In terms of our modeling framework, such a break can be accommodated by e.g. retaining the same copula model, but allowing the copula dependence parameter to change between the two periods. Also interesting to notice is that the transition to the FB methodology generally implies a stronger dependence for the CWE pairs. With everything else being equal, the continuation of such developments in the future would imply that the effects of e.g. tail dependence highlighted in Sections 4.4, 5 and 6 will be amplified, and thus be even more important to account for. At the present time, in-depth studies of the tail dependence after the launch of the FB market coupling are not possible due to data shortage, and we leave such extensions to future research.

Finally, we mention that spot electricity prices have a special feature in the sense that a panel of hourly prices is observed over time. In this paper, we propose a model that only considers each hour of the day separately. There is however a lot of cross-sectional dependence for the hourly prices across each day, and this dependence is not accounted for. As a topic for further research, it would be interesting to consider how this cross-sectional dependence can be modeled in the context of copulas, by e.g. applying vine copulas.

Acknowledgments

The authors would like to thank Jesper Jung and Esben Høg for providing valuable comments and suggestions. Two anonymous referees are also thanked for their constructive criticism and suggestions, which improved the presentation of this paper.

Hour	DE–FR			DE–NL			NL–BE			DE–DK1		
	$\hat{\rho}^{ATC}$	$\hat{\rho}^{FB}$	p-val	$\hat{\rho}^{ATC}$	$\hat{\rho}^{FB}$	p-val	$\hat{\rho}^{ATC}$	$\hat{\rho}^{FB}$	p-val	$\hat{\rho}^{ATC}$	$\hat{\rho}^{FB}$	p-val
0	0.24	0.39	**0.02**	0.21	0.29	0.22	0.36	0.42	0.41	0.38	0.40	0.72
1	0.31	0.38	0.23	0.25	0.36	0.08	0.37	0.42	0.41	0.39	0.41	0.79
2	0.32	0.40	0.25	0.28	0.26	0.77	0.40	0.52	0.07	0.45	0.43	0.74
3	0.34	0.51	**0.01**	0.27	0.38	0.08	0.33	0.50	**0.02**	0.48	0.46	0.71
4	0.39	0.56	**0.01**	0.23	0.44	**0.00**	0.31	0.51	**0.01**	0.43	0.48	0.40
5	0.35	0.48	0.09	0.25	0.43	**0.02**	0.32	0.55	**0.00**	0.39	0.48	0.18
6	0.30	0.49	**0.01**	0.32	0.46	0.07	0.36	0.53	**0.02**	0.38	0.38	0.99
7	0.34	0.62	**0.00**	0.39	0.48	0.22	0.37	0.44	0.32	0.40	0.43	0.53
8	0.35	0.59	**0.00**	0.39	0.52	**0.05**	0.41	0.46	0.51	0.44	0.47	0.44
9	0.27	0.54	**0.00**	0.27	0.41	**0.02**	0.44	0.54	0.14	0.46	0.48	0.68
10	0.30	0.47	**0.00**	0.24	0.35	0.08	0.38	0.50	0.07	0.50	0.41	0.09
11	0.29	0.48	**0.00**	0.20	0.28	0.19	0.34	0.53	**0.00**	0.48	0.40	0.17
12	0.29	0.42	**0.03**	0.21	0.21	0.99	0.33	0.53	**0.00**	0.47	0.40	0.23
13	0.31	0.45	**0.02**	0.24	0.14	0.08	0.32	0.57	**0.00**	0.48	0.44	0.50
14	0.31	0.42	0.06	0.21	0.15	0.34	0.37	0.51	**0.03**	0.44	0.38	0.29
15	0.33	0.49	**0.01**	0.22	0.22	0.99	0.32	0.56	**0.00**	0.43	0.41	0.66
16	0.36	0.48	0.07	0.21	0.24	0.63	0.28	0.57	**0.00**	0.42	0.39	0.62
17	0.36	0.49	0.07	0.20	0.30	0.15	0.25	0.52	**0.00**	0.41	0.41	0.99
18	0.37	0.55	**0.01**	0.25	0.46	**0.01**	0.39	0.44	0.42	0.42	0.43	0.84
19	0.33	0.61	**0.00**	0.27	0.51	**0.00**	0.36	0.46	0.17	0.36	0.34	0.76
20	0.31	0.51	**0.00**	0.17	0.47	**0.00**	0.28	0.35	0.37	0.34	0.33	0.81
21	0.35	0.45	0.07	0.16	0.43	**0.00**	0.33	0.47	**0.04**	0.37	0.30	0.18
22	0.26	0.39	**0.04**	0.15	0.25	0.14	0.31	0.36	0.41	0.35	0.30	0.33
23	0.24	0.37	**0.03**	0.21	0.23	0.65	0.38	0.44	0.32	0.33	0.28	0.41

Table 13: Test results for a one time break in the rank correlations. p-values in bold correspond to cases where the null of no break in the rank correlation is rejected at a 5% significance level.

Funding

Anca Pircalabu acknowledges support from Innovation Fund Denmark, grant number 4135-00082B. Fred Espen Benth acknowledges support from FINEWSTOCH, funded by the Norwegian Research Council.

A Additional tables and figures

A.1 Order of marginal models

Hour	**Order of marginal models**			
	DE–FR	DE–NL	NL–BE	DE–DK1
0	$P = 5, Q = 0$	$P = 4, Q = 0$	$P = 6, Q = 0$	$P = 5, Q = 1$
1	$P = 3, Q = 0$	$P = 4, Q = 0$	$P = 6, Q = 0$	$P = 5, Q = 0$
2	$P = 3, Q = 0$	$P = 6, Q = 0$	$P = 6, Q = 0$	$P = 5, Q = 1$
3	$P = 3, Q = 0$	$P = 6, Q = 0$	$P = 6, Q = 0$	$P = 5, Q = 1$
4	$P = 3, Q = 0$	$P = 6, Q = 0$	$P = 6, Q = 0$	$P = 5, Q = 0$
5	$P = 3, Q = 0$	$P = 6, Q = 0$	$P = 6, Q = 0$	$P = 4, Q = 1$
6	$P = 6, Q = 0$	$P = 6, Q = 0$	$P = 7, Q = 0$	$P = 7, Q = 1$
7	$P = 6, Q = 0$	$P = 6, Q = 0$	$P = 6, Q = 0$	$P = 6, Q = 0$
8	$P = 6, Q = 0$	$P = 7, Q = 0$	$P = 6, Q = 0$	$P = 6, Q = 0$
9	$P = 7, Q = 0$	$P = 6, Q = 0$	$P = 6, Q = 0$	$P = 6, Q = 0$
10	$P = 7, Q = 0$	$P = 7, Q = 0$	$P = 7, Q = 0$	$P = 3, Q = 0$
11	$P = 7, Q = 0$	$P = 7, Q = 0$	$P = 7, Q = 0$	$P = 3, Q = 0$
12	$P = 4, Q = 0$	$P = 7, Q = 0$	$P = 6, Q = 0$	$P = 6, Q = 0$
13	$P = 6, Q = 0$	$P = 7, Q = 0$	$P = 7, Q = 0$	$P = 6, Q = 0$
14	$P = 7, Q = 0$	$P = 7, Q = 0$	$P = 7, Q = 0$	$P = 6, Q = 0$
15	$P = 7, Q = 0$	$P = 7, Q = 0$	$P = 6, Q = 0$	$P = 7, Q = 0$
16	$P = 7, Q = 0$	$P = 6, Q = 0$	$P = 7, Q = 0$	$P = 7, Q = 0$
17	$P = 7, Q = 0$	$P = 7, Q = 0$	$P = 7, Q = 0$	$P = 6, Q = 0$
18	$P = 6, Q = 0$	$P = 6, Q = 0$	$P = 7, Q = 1$	$P = 7, Q = 0$
19	$P = 6, Q = 0$	$P = 6, Q = 0$	$P = 7, Q = 1$	$P = 6, Q = 0$
20	$P = 6, Q = 0$	$P = 6, Q = 0$	$P = 6, Q = 0$	$P = 6, Q = 0$
21	$P = 6, Q = 0$	$P = 6, Q = 0$	$P = 7, Q = 0$	$P = 6, Q = 0$
22	$P = 3, Q = 0$	$P = 4, Q = 0$	$P = 6, Q = 0$	$P = 6, Q = 0$
23	$P = 3, Q = 0$	$P = 3, Q = 0$	$P = 6, Q = 0$	$P = 4, Q = 0$

Table 14: Selected order for all marginal models based on BIC. The variable P refers to the order of autoregression and the variable Q refers to the order of cross-equation effects, cf. Eqs. (1) and (4).

A.2 Parameter estimates for selected models

Model for the pair DE–NL hour 8								
Equal price regime ($s_t = 0$)			**Non-equal price regime** ($s_t = 1$)					
			Area 1: DE			**Area 2: NL**		
Conditional mean (optimal order $P = 7$, $Q = 0$)								
$\hat{\phi}_1^{(0)}$	0.3451	(0.0424)	$\hat{\phi}_1^{(1)}$	0.2923	(0.0573)	$\hat{\phi}_1^{(2)}$	0.2419	(0.0517)
$\hat{\phi}_2^{(0)}$	0.1753	(0.0419)	$\hat{\phi}_2^{(1)}$	0.0905	(0.0573)	$\hat{\phi}_2^{(2)}$	0.1469	(0.0507)
$\hat{\phi}_3^{(0)}$	0.0588	(0.0381)	$\hat{\phi}_3^{(1)}$	-0.0197	(0.0571)	$\hat{\phi}_3^{(2)}$	-0.0581	(0.0501)
$\hat{\phi}_4^{(0)}$	0.0588	(0.0374)	$\hat{\phi}_4^{(1)}$	0.0277	(0.0515)	$\hat{\phi}_4^{(2)}$	0.0950	(0.0493)
$\hat{\phi}_5^{(0)}$	0.0162	(0.0362)	$\hat{\phi}_5^{(1)}$	0.0073	(0.0516)	$\hat{\phi}_5^{(2)}$	0.0093	(0.0482)
$\hat{\phi}_6^{(0)}$	0.1306	(0.0345)	$\hat{\phi}_6^{(1)}$	0.1279	(0.0516)	$\hat{\phi}_6^{(2)}$	0.1192	(0.0468)
$\hat{\phi}_7^{(0)}$	0.0898	(0.0340)	$\hat{\phi}_7^{(1)}$	0.0454	(0.0481)	$\hat{\phi}_7^{(2)}$	0.1195	(0.0447)
Conditional variance								
$\hat{\omega}^{(0)}$	9.2195	(4.6335)	$\hat{\omega}^{(1)}$	36.7656	(13.7786)	$\hat{\omega}^{(2)}$	8.6863	(7.5150)
$\hat{\alpha}^{(0)}$	0.1655	(0.0672)	$\hat{\alpha}^{(1)}$	0.2444	(0.1480)	$\hat{\alpha}^{(2)}$	0.1514	(0.0991)
$\hat{\beta}^{(0)}$	0.5524	(0.1299)	$\hat{\beta}^{(1)}$	0.4139	(0.2149)	$\hat{\beta}^{(2)}$	0.7432	(0.1795)
Marginal distribution (skew t)								
$\hat{\nu}^{(0)}$	5.7397	(1.6234)	$\hat{\nu}^{(1)}$	4.4073	(1.0694)	$\hat{\nu}^{(2)}$	5.4083	(1.5802)
$\hat{\lambda}^{(0)}$	0.0847	(0.0431)	$\hat{\lambda}^{(1)}$	0.0248	(0.0518)	$\hat{\lambda}^{(2)}$	0.1642	(0.0514)

Table 15: Parameter estimates for the DE–NL hour 8 model, with s.e. given in parenthesis.

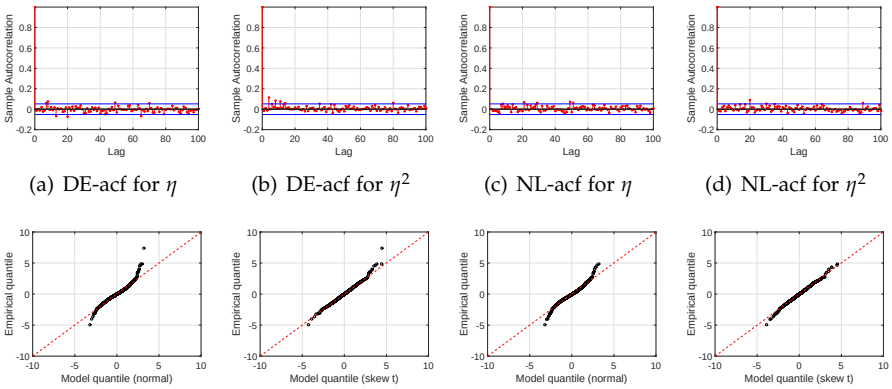
Model for the pair NL–BE hour 8								
Equal price regime ($s_t = 0$)			**Non-equal price regime** ($s_t = 1$)					
			Area 1: NL			**Area 2: BE**		
Conditional mean (optimal order $P = 6$, $Q = 0$)								
$\hat{\phi}_1^{(0)}$	0.3690	(0.0376)	$\hat{\phi}_1^{(1)}$	0.3421	(0.0700)	$\hat{\phi}_1^{(2)}$	0.3533	(0.0698)
$\hat{\phi}_2^{(0)}$	0.1643	(0.0369)	$\hat{\phi}_2^{(1)}$	0.2277	(0.0704)	$\hat{\phi}_2^{(2)}$	0.1006	(0.0633)
$\hat{\phi}_3^{(0)}$	0.1028	(0.0350)	$\hat{\phi}_3^{(1)}$	-0.0410	(0.0731)	$\hat{\phi}_3^{(2)}$	0.1039	(0.0600)
$\hat{\phi}_4^{(0)}$	0.0474	(0.0321)	$\hat{\phi}_4^{(1)}$	-0.0302	(0.0665)	$\hat{\phi}_4^{(2)}$	0.1308	(0.0536)
$\hat{\phi}_5^{(0)}$	-0.0182	(0.0303)	$\hat{\phi}_5^{(1)}$	0.1502	(0.0648)	$\hat{\phi}_5^{(2)}$	0.0374	(0.0526)
$\hat{\phi}_6^{(0)}$	0.1561	(0.0279)	$\hat{\phi}_6^{(1)}$	0.2186	(0.0648)	$\hat{\phi}_6^{(2)}$	0.1237	(0.0504)
Conditional variance								
$\hat{\omega}^{(0)}$	15.7739	(4.6886)	$\hat{\omega}^{(1)}$	82.7166	(15.1057)	$\hat{\omega}^{(2)}$	23.0052	(9.8562)
$\hat{\alpha}^{(0)}$	0.2673	(0.0769)	$\hat{\alpha}^{(1)}$	0.1863	(0.1492)	$\hat{\alpha}^{(2)}$	0.4751	(0.1965)
$\hat{\beta}^{(0)}$	0.3470	(0.1157)	$\hat{\beta}^{(1)}$	0.1467	(0.1932)	$\hat{\beta}^{(2)}$	0.2754	(0.2018)
Marginal distribution (skew t)								
$\hat{\nu}^{(0)}$	6.1638	(1.6688)	$\hat{\nu}^{(1)}$	6.0508	(3.1611)	$\hat{\nu}^{(2)}$	4.1789	(1.3552)
$\hat{\lambda}^{(0)}$	0.0946	(0.0400)	$\hat{\lambda}^{(1)}$	0.0561	(0.0661)	$\hat{\lambda}^{(2)}$	0.0706	(0.0599)

Table 16: Parameter estimates for the NL–BE hour 8 model, with s.e. given in parenthesis.

Model for the pair DE–DK1 hour 8								
Equal price regime ($s_t = 0$)			**Non-equal price regime** ($s_t = 1$)					
			Area 1: DE			**Area 2: DK1**		
Conditional mean (optimal order $P = 6$, $Q = 0$)								
$\hat{\phi}_1^{(0)}$	0.3678	(0.0605)	$\hat{\phi}_1^{(1)}$	0.3419	(0.0381)	$\hat{\phi}_1^{(2)}$	0.3709	(0.0434)
$\hat{\phi}_2^{(0)}$	0.1709	(0.0664)	$\hat{\phi}_2^{(1)}$	0.1432	(0.0366)	$\hat{\phi}_2^{(2)}$	0.0463	(0.0407)
$\hat{\phi}_3^{(0)}$	0.0586	(0.0624)	$\hat{\phi}_3^{(1)}$	0.0721	(0.0392)	$\hat{\phi}_3^{(2)}$	0.0854	(0.0372)
$\hat{\phi}_4^{(0)}$	0.0242	(0.0602)	$\hat{\phi}_4^{(1)}$	0.0226	(0.0368)	$\hat{\phi}_4^{(2)}$	0.0338	(0.0381)
$\hat{\phi}_5^{(0)}$	0.0862	(0.0574)	$\hat{\phi}_5^{(1)}$	0.0054	(0.0375)	$\hat{\phi}_5^{(2)}$	0.0426	(0.0369)
$\hat{\phi}_6^{(0)}$	0.1507	(0.0540)	$\hat{\phi}_6^{(1)}$	0.1094	(0.0351)	$\hat{\phi}_6^{(2)}$	0.0792	(0.0446)
Conditional variance								
$\hat{\omega}^{(0)}$	0.0008	(10.1640)	$\hat{\omega}^{(1)}$	4.2682	(3.8344)	$\hat{\omega}^{(2)}$	70.2888	(20.4931)
$\hat{\alpha}^{(0)}$	0.1323	(0.0718)	$\hat{\alpha}^{(1)}$	0.1811	(0.0556)	$\hat{\alpha}^{(2)}$	0.1715	(0.1159)
$\hat{\beta}^{(0)}$	0.8325	(0.1479)	$\hat{\beta}^{(1)}$	0.7515	(0.0826)	$\hat{\beta}^{(2)}$	0.1856	(0.2307)
Marginal distribution (skew t)								
$\hat{\nu}^{(0)}$	6.5004	(3.5006)	$\hat{\nu}^{(1)}$	5.7861	(1.2491)	$\hat{\nu}^{(2)}$	3.7827	(0.5552)
$\hat{\lambda}^{(0)}$	0.0417	(0.0644)	$\hat{\lambda}^{(1)}$	0.0368	(0.0424)	$\hat{\lambda}^{(2)}$	0.2301	(0.0358)

Table 17: Parameter estimates for the DE–DK1 hour 8 model, with s.e. given in parenthesis.

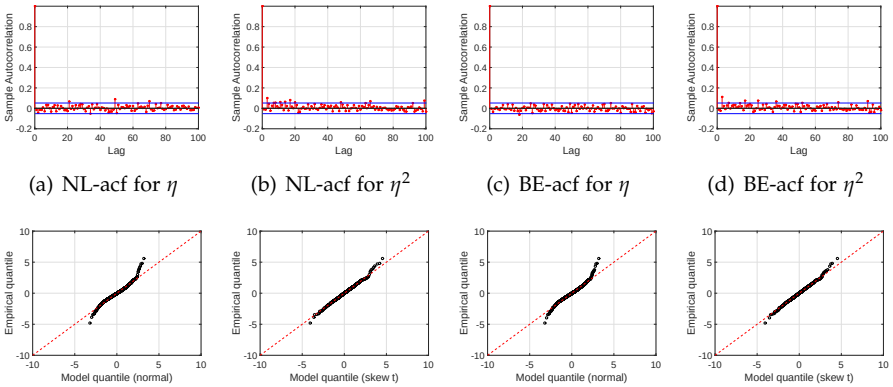
A.3 Model check for selected models



(a) DE-acf for η (b) DE-acf for η^2 (c) NL-acf for η (d) NL-acf for η^2

(e) DE-qq normal (f) DE-qq skew t (g) NL-qq normal (h) NL-qq skew t

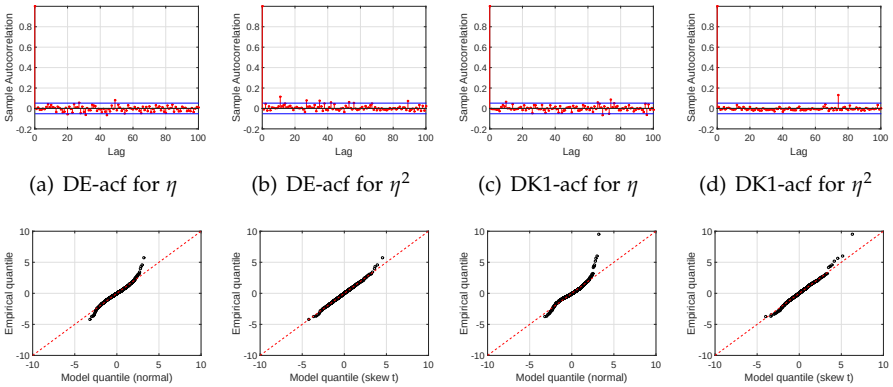
Fig. 8: Sample autocorrelation and quantile plots of standardized residuals resulting from fitting regime-switching AR–GARCH models to the DE–NL pair for hour 8.



(a) NL-acf for η (b) NL-acf for η^2 (c) BE-acf for η (d) BE-acf for η^2

(e) NL-qq normal (f) NL-qq skew t (g) BE-qq normal (h) BE-qq skew t

Fig. 9: Sample autocorrelation and quantile plots of standardized residuals resulting from fitting regime-switching AR–GARCH models to the NL–BE pair for hour 8.



(a) DE-acf for η (b) DE-acf for η^2 (c) DK1-acf for η (d) DK1-acf for η^2

(e) DE-qq normal (f) DE-qq skew t (g) DK1-qq normal (h) DK1-qq skew t

Fig. 10: Sample autocorrelation and quantile plots of standardized residuals resulting from fitting regime-switching AR–GARCH models to the DE–DK1 pair for hour 8.

References

K. Aas, C. Czado, A. Frigessi, and H. Bakken. Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics*, 44(2):182 – 198, 2009.

K. Avdulaj and J. Barunikl. Are benefits from oil–stocks diversification gone? New evidence from a dynamic copula and high frequency data. *Energy Economics*, 51:31 – 44, 2015.

F. E. Benth and P. C. Kettler. Dynamic copula models for the spark spread. *Quantitative Finance*, 11(3):407 – 421, 2011.

F. E. Benth and T. Meyer-Brandis. The information premium for non-storable commodities. *Journal of Energy Markets*, 2(3):111 – 140, 2009.

F. E. Benth and M. D. Schmeck. Pricing futures and options in electricity markets. In S. Ramos and H. Veiga, editors, *Lecture Notes in Energy*, volume 54, pages 233 – 260. Springer Verlag, 2014.

F. E. Benth, J. Šaltytė Benth, and S. Koekebakker. *Stochastic Modelling of Electricity and Related Markets*. World Scientific, 2008.

R. Börger, A. Cartea, R. Kiesel, and G. Schindlmayr. Cross-commodity analysis and applications to risk management. *Journal of Futures Markets*, 29(3): 197 – 217, 2009.

M. Burger, B. Klar, A. Muller, and G. Schindlmayr. A spot market model for pricing derivatives in electricity markets. *Quantitative Finance*, 4(1):109 – 122, 2004.

R. Carmona, M. Coulon, and D. Schwarz. Electricity price modeling and asset valuation: A multi-fuel approach. *Mathematics and Financial Economics*, 7(2): 167 – 202, 2013.

K. F. Chan and P. Gray. Using extreme value theory to measure value-at-risk for daily electricity spot prices. *International Journal of Forecasting*, 22:283 – 300, 2006.

P. F. Christoffersen. Evaluating interval forecasts. *International Economics Review*, 39(4):841 – 862, 1998.

A. Dias and P. Embrechts. Testing for structural changes in exchange rates' dependence beyond linear correlation. *The European Journal of Finance*, 15 (7–8):619 – 637, 2009.

C. Elberg and S. Hagspiel. Spatial dependencies of wind power and interrelations with spot price dynamics. *European Journal of Operational Research*, 241(1):260 – 272, 2015.

EPEX. Flow-Based methodology for CWE market coupling successfully launched. epexspot.com, 2015.

R. Füss, S. Mahringer, and M. Prokopczuk. Electricity market coupling in Europe: Status quo and future challenges. *University of St. Gallen, School of Finance Research Paper No. 1512*, 2015.

R. Füss, S. Mahringer, F. Paraschiv, and M. Prokopczuk. Electricity spot and derivatives pricing under market coupling. *University of St. Gallen, School of Finance Research Paper No. 1323*, 2017.

O. Grothe and J. Schnieders. Spatial dependence in wind and optimal wind power allocation: A copula-based analysis. *Energy Policy*, 39(9):4742 – 4754, 2011.

N. Haldrup and M. Ø. Nielsen. A regime switching long memory model for electricity prices. *Journal of Econometrics*, 135(1–2):349 – 376, 2006.

N. Haldrup, F. S. Nielsen, and M. Ø. Nielsen. A vector autoregressive model for electricity prices subject to long memory and regime switching. *Energy Economics*, 32(5):1044 – 1058, 2010.

J. Janczura, S. Trück, R. Weron, and R. C. Wolff. Identifying spikes and seasonal components in electricity spot price data: A guide to robust modeling. *Energy Economics*, 38:96 – 110, 2013.

H. Joe. *Multivariate Models and Dependence Concepts*. Chapman and Hall, 1 edition, 1997.

D. Keles, M. Genoese, D. Möst, and W. Fichtner. Comparison of extended mean-reversion and time series models for electricity spot price simulation considering negative prices. *Energy Economics*, 34(4):1012 – 1032, 2012.

R. Kiesel and M. Kustermann. Structural models for coupled electricity markets. *Journal of Commodity Markets*, 3(1):16 – 38, 2016.

C. R. Knittel and M. R. Roberts. An empirical examination of restructured electricity prices. *Energy Economics*, 27:791 – 817, 2005.

J. J. Lucia and E. S. Schwartz. Electricity prices and power derivatives: Evidence from the Nordic Power Exchange. *Review of Derivatives Research*, 5 (1):5 – 50, 2002.

R. B. Nelsen. *An Introduction to Copulas*. Springer, 1999.

F. Paraschiv. Price dynamics in electricity markets. In R. M. Kovacevic, G. Ch. Pflug, and M. T. Vespucci, editors, *Handbook of Risk Management in Energy Production and Trading*, pages 47 – 69. Springer, 2013.

A. J. Patton. Modelling time-varying exchange rate dependence using the conditional copula. *UCSD Dept. of Economics, working paper*, 2001.

A. J. Patton. Modelling asymmetric exchange rate dependence. *International Economic Review*, 47(2):527 – 556, 2006.

A. J. Patton. Copula methods for forecasting multivariate time series. In Graham Elliott and Allan Timmermann, editors, *Handbook of Economic Forecasting*, volume 2B, pages 899 – 960. Elsevier B.V., 2013.

A. Sklar. Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de L'Université de Paris*, 8:229 – 231, 1959.

TenneT. Market review 2013. tennet.eu, 2013.

Paper IV

A non-Gaussian Ornstein-Uhlenbeck model for pricing wind power futures

List of authors: Fred Espen Benth* and Anca Pircalabu†,‡
* Department of Mathematics, University of Oslo
† Department of Mathematical Sciences, Aalborg University
‡ Quantitative Analytics, Neas Energy

Submitted to
Applied Mathematical Finance

The layout has been revised.

ABSTRACT

The recent introduction of wind power futures written on the German wind power production index has brought with it new interesting challenges in terms of modeling and pricing. Some particularities of this product are the strong seasonal component embedded in the underlying, the fact that the wind index is bounded from both above and below, and also that the futures are settled against a synthetically generated spot index. Here, we consider the non-Gaussian Ornstein-Uhlenbeck type processes proposed by Barndorff-Nielsen and Shephard (2001) in the context of modeling the wind power production index. We discuss the properties of the model and estimation of the model parameters. Further, the model allows for an analytical formula for pricing wind power futures. We provide an empirical study, where the model is calibrated to 37 years of German wind power production index that is synthetically generated assuming a constant level of installed capacity. Also, based on one year of observed prices for wind power futures with different delivery periods, we study the market price of risk. Generally, we find a negative risk premium whose magnitude decreases as the length of the delivery period increases. To further demonstrate the benefits of our proposed model, we address the pricing of European options written on wind power futures, which can be achieved through Fourier techniques.

1 Introduction

Following the significant expansion in wind turbine installations that some European countries have experienced over the past years, the demand for financial instruments that can be used to address the problem of volumetric risk in wind power generation has grown. This has led to the launch of a standardized product written on the wind power production index, namely the so-called wind power futures (or wind index futures). Currently, wind power futures can be traded on NASDAQ OMX and the European Energy Exchange (EEX) on the German wind power production index. The index is obtained by measuring the German wind power generation relative to the available installed capacity; hence, the index has a lower bound of 0 and an upper bound of 1, corresponding to a 0% and a 100% wind power utilization, respectively.

To clarify the payoff structure of wind power futures, let us denote by $F(t,T)$ the wind power futures price at time t and delivery during day T, with $0 \leq t < T$ and $0 \leq F(t,T) \leq 1$. Further, let $P(T)$ be the wind index measured at day T. Then, a long position in a wind power futures contract entered at time $t \leq T$ for delivery at T yields the payoff

$$24(P(T) - F(t,T)) \cdot x,$$

where 24 denotes the usual number of hours in a day and x denotes a known fixed tick size. For the wind power futures traded at NASDAQ OMX and EEX, $x = 100$ EUR. Moreover, the futures are settled against an externally provided spot index $P(T)$, which is synthetically generated based on weather data and an individual power curve for every grid point in Germany.

Natural sellers of wind power futures are the wind power producers and companies with considerable wind park portfolios, as they are interested in protection against the low wind scenarios, which are likely to lower revenues. Although one could argue that day-ahead electricity prices tend to increase in times of low wind, wind power generators usually receive a fixed price per generated unit of electricity, and do not participate in the wholesale market themselves. Hence, volumetric risk is the only risk source left to be addressed, and wind power futures can be an obvious tool for stabilizing the revenue of the wind power generators. Typical buyers are conventional power plants acting in e.g. the day-ahead market, whose profitability drops in times of high wind due to the negative relation between wind power production and spot electricity prices.

In this paper, we propose a non-Gaussian Ornstein-Uhlenbeck process in the spirit of Barndorff-Nielsen and Shephard (2001) to model the wind power production index. The model is very straightforward, allowing for an easy estimation of the parameters and analytical pricing of wind power futures, with the latter facilitating the study of the market price of risk. Based on one year of observed German wind power futures curves, we perform an empirical analysis of the risk premia in this newly established market.

Wind power futures are characterized as weather derivatives, and fall in this category together with derivatives written on temperature, rainfall, snowfall, humidity, etc. While the existing literature on temperature derivatives is extensive and broad in terms of modeling approach (see e.g. Davis (2001), Brody et al. (2002), Cao and Wei (2004), Campbell and Diebold (2005), Platen and West (2005), Härdle and Cabrera (2012) and Benth and Benth (2011)), literature related specifically to wind derivatives is very scarce. To the best of our knowledge, the first study concerned with the pricing of wind derivatives is that of Benth and Benth (2009), which was motivated by the introduction of futures and options on wind speed indexes at different wind farm locations in the US back in 2007. However, trade in these products never really picked up, explaining perhaps the scarcity of related studies.

Almost ten years after the first attempt to establish a market for wind derivatives, the introduction of the German wind power futures on NASDAQ OMX and EEX awakens interest again. The study of Gersema and Wozabal (2017) is the first to provide a thorough introduction to the German wind power futures market, the market players and their risks. Further, Gersema and Wozabal (2017) propose an equilibrium pricing model, and based on different case studies they conclude that a negative risk pre-

mium is to be expected in wind power futures markets. A second related study concerning wind power futures is that of Pircalabu and Jung (2017), where the authors focus on the hedging benefits of wind power futures in the context of energy trading companies entering into long-term agreements with wind power generators, where the fluctuating wind power production is bought at a pre-determined fixed price. Here, wind power futures are not the main focus of the analysis, and they are thus treated on a conceptual basis, disregarding some practical aspects concerning the data foundation in their pricing application. In this paper, we shall address this aspect in detail, highlighting its importance.

The paper is structured as follows: In Section 2, we present the data and comment on key features as to motivate the model choice. In Section 3, we introduce the model for the wind power production index and provide an empirical study where the model is applied to German data. Analytical futures prices are derived in Section 4, and based on one year of market prices for wind power futures with different delivery periods, the market price of risk is studied. In Section 5, we elaborate on further applications of the proposed model in derivatives valuation. Section 6 concludes.

2 Data presentation

Since wind power futures are only traded on the German wind power index at the moment, the empirical analysis performed in this paper is based on German data. We consider a time series of daily wind power production indexes for the German market, which was synthetically constructed by MeteoGroup for a period of 37 years (1 January 1979 to 31 December 2015). The synthetic index is displayed in Fig. 1, and measures how the utilization of installed wind power capacity would have looked like in the German market zone in the past, conditional on the *present level* of available capacity and geographical location of wind turbines. Specifically, the *present level* we consider here corresponds to September 2016. To construct such an index, a bottom-up approach was implemented based on historical weather data and power curves. Clearly, since the wind index measures the wind power production relative to the installed capacity, it must be bounded between zero and one. For the data in Fig. 1, the lowest and highest values recorded are 0.35% and 83.05%, respectively.

In the context of pricing wind power futures, which is the main focus of the present paper, we argue that fitting a model to the type of data in Fig. 1 seems much more reasonable than considering the historical evolution of the wind index. This is an essential point, since the wind power futures price today is clearly not influenced by how the available installed capacity evolved over time in Germany, but rather on the present and ideally the

future installed capacity level.

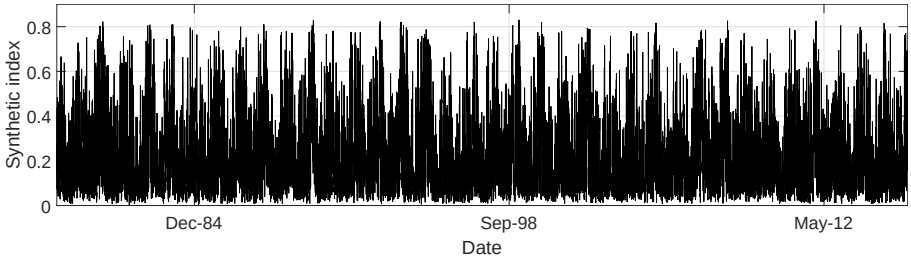


Fig. 1: Index constructed based on the total installed wind power capacity observed in September 2016.

2.1 Seasonality

Aside from the wind index being bounded on [0,1], another key feature is the yearly seasonality we observe in the data illustrated in Fig. 1. To emphasize the annual pattern, we complement the time series plot in Fig. 1 with the empirical autocorrelation function of the index in Fig. 2(a). Following the related literature (see e.g. Benth et al. (2008), Härdle and Cabrera (2012), and Benth and Benth (2011)), the yearly seasonality can be addressed by the following seasonality function:

$$\Lambda(t) = a_1 + a_2 \sin(2\pi t/365) + a_3 \cos(2\pi t/365). \tag{1}$$

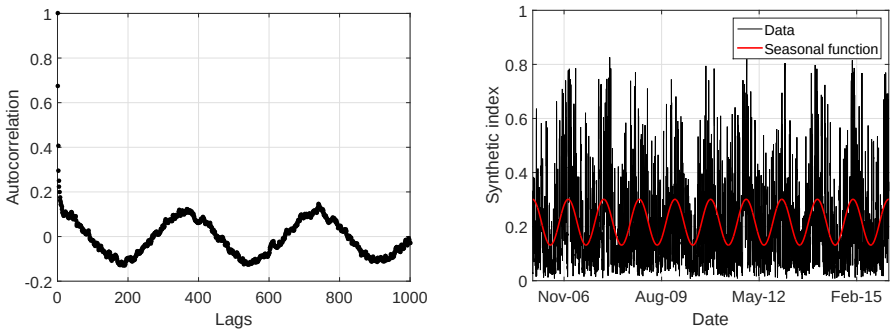
Fitting this function to the wind index by ordinary least squares yields the parameter estimates reported in Table 1. In Fig. 2(b), the wind power production index is plotted together with the fitted seasonal function. For better clarity, we display a snapshot of the last 10 years, i.e., from year 2006 to 2015.

	Estimate	Standard error
$\hat{a}_1$	0.2164	0.0014
$\hat{a}_2$	0.0102	0.0020
$\hat{a}_3$	0.0839	0.0020

Table 1: OLS estimates for the parameters of the seasonal function.

3 A model for the wind power production index

Motivated by the two key features of the wind index enhanced in Section 2, i.e., boundedness on [0,1] and yearly seasonality, we specify a model for



(a) Empirical autocorrelation function (b) Fitted seasonal function

Fig. 2: Empirical autocorrelation function of the synthetic wind power production index and fitted seasonal function.

the wind power production index as follows. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space with a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ satisfying the usual conditions. We denote by P the wind power production index obtained by measuring the wind power production relative to the available installed capacity, implying that $P(t) \in [0, 1]$ for all t.

We define $P(t)$ as

$$P(t) = \Lambda(t) \exp(-X(t)), \tag{2}$$

where $\Lambda(t)$ describes the deterministic seasonal component of the wind power production index and $X(t)$ follows a non-Gaussian Ornstein-Uhlenbeck process as in the stochastic volatility model proposed by Barndorff-Nielsen and Shephard (2001). Specifically,

$$\mathrm{d}X(t) = \alpha(\mu - X(t))\mathrm{d}t + \mathrm{d}L(t), \tag{3}$$

with L being a driftless subordinator, and $\mu > 0$, $\alpha > 0$ denoting two constants. From standard theory, the solution of the Ornstein-Uhlenbeck process is

$$X(t) = X(0)\mathrm{e}^{-\alpha t} + \mu(1 - \mathrm{e}^{-\alpha t}) + \int_0^t \mathrm{e}^{-\alpha(t-s)}\mathrm{d}L(s),$$

where $X(0) = \ln(\Lambda(0)/P(0))$. The constant μ is connected to $\Lambda(t)$, and its purpose is to ensure that $P(t)$ never exceeds 1. In order to elaborate on this, we include the following Proposition regarding the stationarity of $X(t)$.

Proposition 3.1. *Let $\ell(dz)$ denote the Lévy measure corresponding to the Lévy process $L(t)$. If*

$$\int_{|z|>2} \ln|z|\ell(dz) < \infty,$$

then $X(t)$ given by the Ornstein-Uhlenbeck process in Eq. (3) *has a limiting distribution. The stationary solution of $X(t)$ is*

$$X(t) = \mu + \int_{-\infty}^{t} \mathrm{e}^{-\alpha(t-s)} dL(s),$$

where L here is a two-sided Lévy process.

We refer to Sato (1999), Thm. 17.5, for a proof and more details on this result. Regarding the stationary solution X, we refer to an extensive discussion in Basse-O'Connor et al. (2014).

Returning to the connection between μ and $\Lambda(t)$, let $M = \max(\Lambda(t))$. Then, we must have that

$$\begin{aligned} \max(P(t)) &= \max(\Lambda(t) \exp(-X(t)) \\ &\leq M \exp(-\min(X(t))). \end{aligned}$$

Owing to L being a subordinator, it follows from the stationary solution in Proposition 3.1 that $X(t) \geq \mu$. Further, since we also have that $0 \leq P(t) \leq 1$, we choose μ such that $M \exp(-\mu) = 1$. Thus, we let

$$\mu = \ln M, \tag{4}$$

and obtain an exact upper bound of 1 as a possible case. This way of introducing seasonality in the model has its advantages and disadvantages, and we refer to Appendix A for a detailed discussion on the subject.

Next, we state the limiting distribution of $X(t) - \mu$ for a specific case, since this will be used in our empirical study.

Proposition 3.2. *If $L(t)$ is a compound Poisson process with exponentially distributed jumps,*

$$L(t) = \sum_{k=1}^{N(t)} J_k, \tag{5}$$

where $N(t)$ is a Poisson process with frequency λ and J_k are independent identically distributed exponential random variables with density function

$$f_J(x) = \kappa \mathrm{e}^{-\kappa x}, \tag{6}$$

then the limiting distribution of $X(t) - \mu$, where $X(t)$ evolves according to Eq. (3)*, is the Gamma distribution with density function given by*

$$f_\Gamma(x) = \frac{\kappa^{\lambda/\alpha} x^{\lambda/\alpha - 1} \mathrm{e}^{-\kappa x}}{\Gamma(\frac{\lambda}{\alpha})}. \tag{7}$$

Proof. See Appendix B.1.

While the model proposed in this section captures key features of the wind index, there are other alternatives when it comes to modeling data with range [0,1]. In particular, we mention the Jacobi processes. A Jacobi process, which is in fact an extension of the Heston model, will have values in any desirable positive interval [Ackerer et al. (2017)]. In our case, we could consider a process of the type

$$\mathrm{d}P(t) = -a(P(t) - b)\mathrm{d}t + \sqrt{cP(t)(1 - P(t))}\mathrm{d}W(t),$$

where $a > 0$, $c > 0$, $0 < b < 1$ and W denotes a Brownian motion.

On one hand, the Jacobi approach is simpler compared to our proposed model in that the wind index is modeled directly, and the Lévy process is replaced by a Brownian motion. On the other hand, since we let $\ln(\Lambda(t)/P(t))$ be an Ornstein-Uhlenbeck, our approach is advantageous from a calibration perspective. In fact, estimation of model parameters for the Jacobi process is not straightforward cf. Gouriéroux and Valéry (2002). Furthermore, the marginal distribution of the Jacobi process is fixed to the beta distribution, whereas our model allows for great flexibility in choosing marginal distributions. In terms of derivatives pricing, both models have their advantages when it comes to the pricing of wind power futures. For the Jacobi process, making use of its polynomial property could result in simple (possibly explicit) expressions for the futures price. Regarding our proposed process, explicit pricing formulas for wind power futures are attainable, as we shall illustrate later in the paper. Unlike our model however, it is unclear how a measure change is included in the Jacobi model as to preserve the Jacobi-structure. In light of the discussion above, we favor the model in Eqs. (2)-(3), and shall not pursue the Jacobi processes in the present paper. Nevertheless, we stress that the Jacobi approach is an interesting and unexplored alternative for modeling the wind power production index.

3.1 An empirical analysis on German wind index data

In this section, we turn to the empirical study of the German wind index time series in Fig. 1. Recalling that the seasonal function entering Eq. (2) has already been estimated in Section 2.1, an estimate for μ immediately follows from Eq. (4). We obtain $\hat{M} = 0.3009$, implying that

$$\hat{\mu} = -1.2010. \tag{8}$$

Using the expression for $P(t)$ in Eq. (2), the variable $X(t) - \mu$ is then constructed by

$$X(t) - \mu = -\left(\ln\left(\frac{P(t)}{\Lambda(t)}\right) + \mu\right). \tag{9}$$

According to Appendix A, our way of incorporating seasonality in the model introduces the potential of having

$$X(t) - \mu < 0. \tag{10}$$

When considering the time series $X(t) - \mu$, we do indeed observe negative values; however, the percentage of negative data points is very low, corresponding to 0.95%, which we find acceptable.

Next, we consider the parameter α entering the dynamics of the $X(t)$ process cf. Eq. (3), and note that

$$X(t+1) - \mu = \mathrm{e}^{-\alpha}(X(t) - \mu) + \int_t^{t+1} \mathrm{e}^{-\alpha(t+1-s)} \mathrm{d}L(s). \tag{11}$$

Clearly, it follows from Eq. (11) that α can be obtained by fitting an AR(1) model to $X(t) - \mu$, and by using the relation $\phi = \exp(-\alpha)$, where ϕ denotes the slope coefficient in the AR(1). However, this procedure requires residuals to be normally distributed, which is not the case here. Recalling that the lag s correlation between observations s periods apart can be expressed as

$$\mathrm{Corr}(X(t+s), X(t)) = \phi^s = \mathrm{e}^{-\alpha s},$$

we fit instead the function $\exp(-\alpha t)$ to the sample autocorrelation of $X(t) - \mu$ using nonlinear least squares[1], and obtain

$$\hat{\alpha} = 0.5455 \tag{12}$$

with a standard error of 0.0011. For comparison purposes, we provide in Appendix C detailed results from implementing the AR(1)-estimation approach, including a residual analysis.

To get an idea of the goodness-of-fit of the proposed exponential function, we plot in Fig. 3 the empirical autocorrelation function together with the fitted exponential. The fit is satisfactory, capturing rather well the sample autocorrelations at the first lags, which are also the most significant. We do however note that the fitted exponential drops to zero slightly quicker than the sample autocorrelation does.

The remaining part of the fitting procedure relates to estimating the parameters of the stationary distribution of $X(t) - \mu$, i.e. the Gamma distribution cf. Proposition 3.2. The choice of a Gamma distribution is motivated by its correspondence to an $L(t)$ being a compound Poisson process with exponential jumps, as well as its reasonable description of the data which we shall illustrate shortly.

Due to the presence of dependence, fitting the Gamma distribution to the actual data (the positive part) would not necessarily yield accurate estimates. Consequently, we wish to fit the Gamma distribution to an *iid* sample

[1] We applied the **nlinfit** function in Matlab.

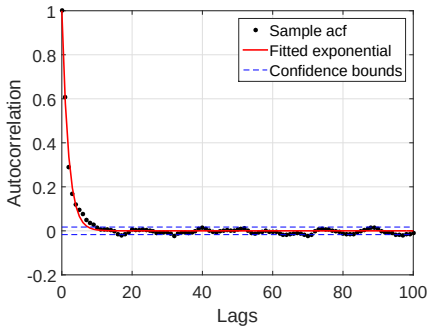


Fig. 3: Fitted exponential to the empirical autocorrelation function of $X(t) - \mu$.

generated from the actual data. To achieve this, we consider the 'opposite' of a block bootstrap, in the sense that we do not wish to generate a bootstrapped sample that preserves the autocorrelation structure the we observe in the data; on the contrary, we wish to ensure independence. Specifically, we follow the procedure described below to obtain a sample with the desired properties:

1. Estimate an optimal block-length l by following the procedure in Politis and White (2004) and Patton et al. (2009).[2]

2. Draw a number $x_1 \geq 0$ from the empirical distribution of $X(t) - \mu$ and let $B_{1,l} = \{\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_l\}$ denote the block consisting of l consecutive indexes, with $\tilde{x}_1$ corresponding to the position of x_1 in $X(t) - \mu$.

3. Let $\tilde{T}$ equal the length of the original time series ($\tilde{T} = 13,514$ cf. Fig. 1) and repeat the following for $j = 2, \ldots, \tilde{T}$.

   (a) Draw (with replacement) a new number $z_j \geq 0$ from the empirical distribution of $X(t) - \mu$, and let $\tilde{z}_j$ be the corresponding index.

   (b) If $\tilde{z}_j \subseteq B_{j-1,l}$, discard the draw and repeat step (a). Otherwise, set $x_j = z_j$, $B_{j,l} = \{\tilde{x}_j, \ldots, \tilde{x}_{j+l-1}\}$ and proceed.

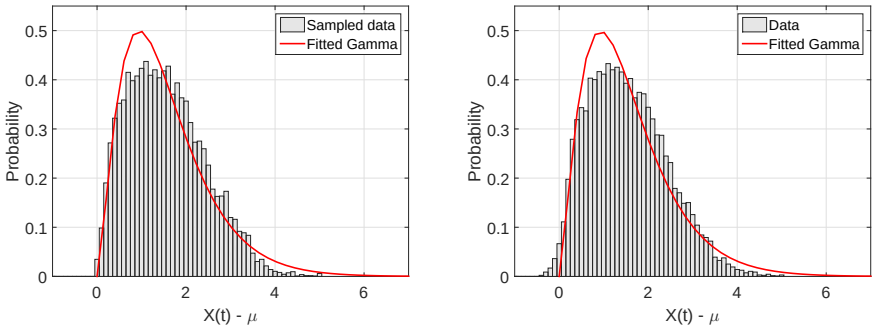
We implement the above procedure with $\hat{l} = 45$, and fit a Gamma distribution to the generated bootstrap sample of the data of length $\tilde{T}$. Stressing that the parameter α in Eq. (3) coincides with α in Eq. (7), we retrieve $\hat{\lambda}$ and $\hat{\kappa}$ conditional on $\hat{\alpha}$ cf. Eq. (12). By repeating this $N = 10,000$ times, a bootstrapped distribution of $\{(\hat{\lambda}_i, \hat{\kappa}_i)\}_{n=1}^{N}$ is produced. Based on these bootstrapped distributions, we then obtain the estimates reported in the first column block of

[2]The procedure is intended for e.g. carrying out the so-called stationary block bootstrap introduced in Politis and Romano (1994), which is generally applicable for stationary weakly dependent time series.

Table 2. For comparison, we also fit a Gamma distribution to the actual data (the positive part). The results are displayed in the second column block of Table 2, and we find that they are very similar to the ones obtained with the bootstrap method. To provide some evidence for the goodness-of-fit of the Gamma distribution, we plot in Fig. 4 an example of a bootstrapped sample of the data and the empirical distribution of $X(t) - \mu$, together with corresponding fitted Gamma distributions. Disregarding the few negative values in the empirical distribution of $X(t) - \mu$, the results show that the Gamma distribution provides an acceptable fit to the data.

	Bootstrap procedure		**Empirical distribution of $X(t) - \mu$**	
	Estimate	Standard error	Estimate	Standard error
$\hat{\lambda}$	1.3649	0.0183	1.3645	0.0157
$\hat{\kappa}$	1.6201	0.0207	1.6187	0.0206

Table 2: Parameter estimates for the Gamma distribution, conditional on $\hat{\alpha}$.



(a) Example of a bootstrapped sample

(b) Empirical distribution

Fig. 4: Bootstrapped and empirical distribution of $X(t) - \mu$ together with the corresponding fitted Gamma distribution.

We remark in passing that other stationary distributions could be chosen as long as they are within the class of self-decomposable distributions. However, a different choice of a stationary distribution does not always result in the Lévy process being easily characterisable, as is the case with the Gamma distribution. For detailed discussions on self-decomposability, we refer the interested reader to Barndorff-Nielsen and Shephard (2001) and Halgreen (1979).

Based on the empirical results obtained in this section, we conclude that the model proposed in Eqs. (2)-(3) provides a good overall fit and is thus a reasonable model for the German wind power production index.

4 Pricing of wind power futures

Motivated by the recent introduction of futures written on the German wind power production index, we derive in this section futures prices based on our proposed model. We denote by $F(t,T)$ the wind power futures price at day $t \geq 0$, with delivery at day $T \geq t$. As it is usual in these types of markets, if we want to consider martingale pricing, we must *define* the futures price as the conditional expectation of the wind index at delivery, since the buy-and-hold argument does not hold. The expectation is not to be taken under the objective measure $\mathbb{P}$, but under a pricing measure $\mathbb{Q}$ that is equivalent to $\mathbb{P}$, and hence, $\mathbb{Q}$-dynamics for the wind power production index must be established.

Since wind is naturally not a tradable asset, there are many potential martingale measures $\mathbb{Q}$. In order to choose such one, we consider here the class of parametrized equivalent measures that can be obtained from the Esscher transform [Esscher (1932)]. Restricting our discussion to a constant market price of risk which we shall denote by θ, and following Benth et al. (2008), we define the probability $\mathbb{Q}$ through

$$\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_t} = \exp\left(\theta L(t) - \psi_{L(1)}(-i\theta)t\right), \tag{13}$$

with $\psi_{L(1)}$ being the cumulant function of $L(1)$ defined as

$$\psi_{L(1)}(x) = \ln \mathbb{E}[e^{ixL(1)}]. \tag{14}$$

Furthermore, to ensure that the Esscher transform is well-defined, we assume that there exists a non-negative constant c such that

$$\mathbb{E}[e^{cL(1)}] < \infty. \tag{15}$$

Hence, the Esscher transform is well-defined for all $\theta \leq c$.

Narrowing the discussion down to our context, where we let $L(t)$ be a compound Poisson process with exponentially distributed jumps, the cumulant function of $L(1)$ becomes

$$\psi_{L(1)}(x) = \lambda \frac{ix}{\kappa - ix}. \tag{16}$$

For a detailed derivation of this result we refer to the proof in Appendix B.1. Also, since we have established that the limiting distribution is the Gamma distribution, we get explicit conditions for the non-existence of the cumulant, as we shall illustrate shortly.

In the following proposition, we derive an explicit expression for the futures price.

Proposition 4.1. *Let $0 \leq t \leq T$ and assume that $P(t)$ and $X(t)$ evolve according to the model in Eq.* (2) *and Eq.* (3)*, respectively. Further, let $L(t)$ be a compound Poisson process as specified in Proposition 3.2. Assuming that the exponential moment condition in Eq.* (15) *holds for a $c \geq 0$, we have that the wind power futures price $F(t,T)$ is given by*

$$F(t,T) = \Lambda(T) H_\theta(t,T) \left(\frac{P(t)}{\Lambda(t)} \right)^{\exp(-\alpha(T-t))}, \tag{17}$$

where

$$H_\theta(t,T) = \exp\left(-\mu(1 - \mathrm{e}^{-\alpha(T-t)}) \right) \left(\frac{\kappa_\theta + \mathrm{e}^{-\alpha(T-t)}}{\kappa_\theta + 1} \right)^{\lambda_\theta/\alpha},$$

and

$$\kappa_\theta = \kappa - \theta,$$
$$\lambda_\theta = \frac{\lambda\kappa}{\kappa - \theta}.$$

Proof. See Appendix B.2.

Since $P(t) \leq 1$, it follows that $\mathbb{E}^{\mathbb{Q}}[P(T)|\mathcal{F}_t] \leq 1$, and so $F(t,T) \leq 1$; moreover, if $P(t) \geq 0$, we also have that $F(t,T) \geq 0$.

Considering the expression for $F(t,T)$ in Proposition 4.1, we note that the condition $\theta < \kappa$ must be imposed to ensure exponential integrability of L and thus the existence of an Esscher transform. Since the estimated κ is positive cf. Table 2, we have no sign restriction on the market price of risk θ. Also note that while the distributional properties of the jump process remain unchanged, the jump intensity and jump size are impacted by the Esscher transform: $L(t)$ is still a compound Poisson process with exponentially distributed jumps, but now with intensity λ_θ and mean jump size $1/\kappa_\theta$. A positive θ will emphasize the jump intensity and the mean jump size, while a negative θ will have the opposite effect.

According to Eq. (17), the shape of the futures curve $T \to F(t,T)$ depends explicitly on the seasonal function Λ, a function H_θ that incorporates the market price of risk θ and a term that includes today's spot wind power index $P(t)$. The seasonal component gives a contribution to the futures curve corresponding to the fitted seasonal function plotted in Fig. 2(b).

To illustrate the contribution from the second term entering the expression for $F(t,T)$, we plot in Fig. 5 the evolution of $H_\theta(t,T)$ as a function of θ for five different maturities, and using the parameter estimates for α, λ and κ obtained in Sec. 3.1 for the German data. When considering the different maturities T, we observe that $H_\theta(t,T)$ converges very fast to a fixed shape as

T increases; with no market price of risk, that is $\theta = 0$, the contribution from the second term is very close to 1 meaning that the futures price is almost unaffected by this term. Generally, we observe that a negative θ implies a value of $H_\theta(t,T) > 1$ and hence an increase in the futures price. Equivalently, a positive θ implies a decrease in the futures price.

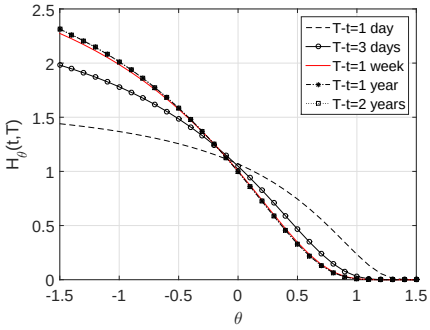


Fig. 5: Values of $H_\theta(t,T)$ for $\theta \in [-1.5, 1.5]$ and different maturity periods. α, λ and κ are fixed to the estimated values obtained in Sec. 3.1.

Unlike the first two terms, the third term in Eq. (17) gives rise to a stochastically varying shape for the futures curve in the short end. As the time to maturity increases, this term will either decrease or increase to 1 depending on whether $P(t) > \Lambda(t)$ or $P(t) < \Lambda(t)$. To depict this behavior, we set $t = 0$ and α to its estimated value from Sec. 3.1, and plot in Fig. 6 two situations: First, we let $P(0) = 0.40$ and $\Lambda(0) = 0.30$ and second, we let $P(0) = 0.20$ and $\Lambda(0) = 0.30$.

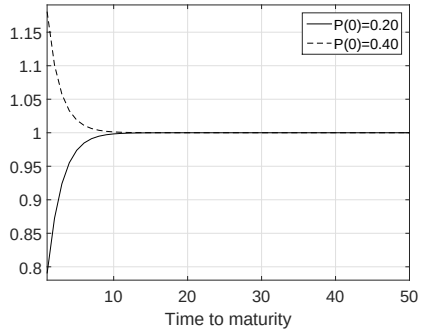
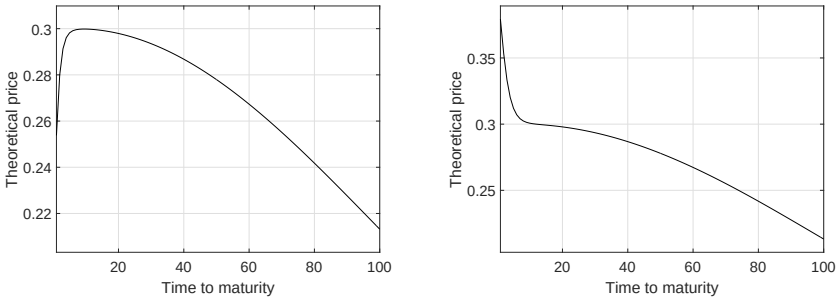


Fig. 6: The shape of the third term entering the expression for the futures price in Eq. (17). We consider the case where $\Lambda(0) = 0.30$, and $P(0) = 0.20$ and $P(0) = 0.40$, respectively.

Combining the three terms discussed above, we obtain two futures curves which we plot in Fig. 7. The same values for $\Lambda(0)$ and $P(0)$ as in Fig. 6 are employed, with $\theta = 0$ and the parameter estimates obtained for the German

data. On the short end, the shape of the futures curve is highly influenced by the behavior of the contributing term from Fig. 6. On the long end, the shape is mostly influenced by the seasonal function, as $F(t,T) \sim v\Lambda(T)$ for a constant v and $T \gg t$. The decreasing pattern of both curves in Fig. 7 in the long end is due to the yearly seasonal cycle and the fact that the initial value of the seasonality curve, i.e. $\Lambda(0)$, corresponds to data as of 1 January. We stress that the annual pattern of the term structure is not clear in Fig. 7 as we restrict our attention to 100 days.



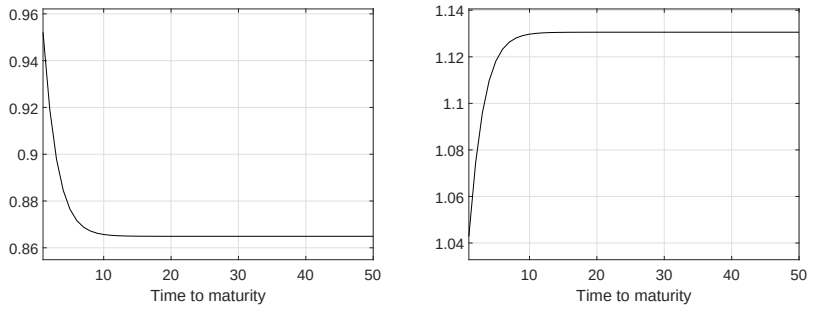
(a) Futures curve with $P(0) < \Lambda(0)$ (b) Futures curve with $P(0) > \Lambda(0)$

Fig. 7: Theoretical futures curves implied by the proposed model fitted to the German data. The market price of risk θ is set to zero, and the start values for $\Lambda(0)$ and $P(0)$ set to illustrate the same instances as in Fig. 6.

Since the market price of risk is rarely zero in reality, we also investigate the contribution to the futures curve implied by $\theta = \pm 0.1$; that is, we again compute $H_\theta(t,T)$, but now for two fixed values of θ and maturities $T \in [1,50]$. The results are displayed in Fig. 8, showing that the contribution of a constant market price of risk $\theta \neq 0$ corresponds to a function that is either decreasing or increasing exponentially.

4.1 An empirical study of the market price of risk

Since wind power futures on the German wind power index have been traded for a while now, historical futures prices quoted in the market are available, allowing us to perform an empirical study of the market price of risk. Like with commodity futures such as power or gas futures, delivery periods for wind power futures are usually an entire week, month, year, etc. This is in contrast with the type of curve implied by our proposed model, which is smooth and made up of daily futures prices (contracts with non-overlapping delivery periods). To convert the single-day delivery prices $F(t,T)$ obtained with our model to prices of contract types quoted in the market, we assume



(a) Market price of risk $\theta = 0.1$ (b) Market price of risk $\theta = -0.1$

Fig. 8: Contribution of a constant market price of risk to the futures curve. Here, we consider the cases $\theta = \pm 0.1$.

that

$$f(t, T_1, T_2) = \frac{1}{T_2 - T_1 + 1} \sum_{\tau=T_1}^{T_2} F(t, \tau),$$

where T_1 and T_2 denote start and end delivery dates, respectively.

As we shall illustrate shortly, wind power futures prices are given in EUR/wph (wind production hour) with a tick size fixed to EUR 100, which is used to convert differences between futures and spot values into a monetary measure (see NASDAQ OMX (2017)). Given a wind power futures contract with delivery during e.g. a week, a difference of 0.01 (1%) between the value of the futures contract at time t and the average realized index for the same delivery will yield a profit or loss of 1 (EUR/wph) $\times$ 24 (hours) $\times$ 7 (days) $=$ 168 (EUR). With our model, we established that $0 \leq F(t, T) \leq 1$, and we will simply multiply this value by 100 such that theoretical and quoted prices are comparable.

To provide an example of a wind power futures curve quoted in the market, we illustrate in Fig. 9 the observed curve on $t = 1$ September 2016. The observations correspond to NASDAQ OMX closing prices and are plotted using horizontal lines from start to end delivery, where time is measured in days. We note that we make up the observed curve using 13 contracts, namely 3 front weeks, 5 front months, 4 front quarters and 1 front year, relative to the valuation date t. For comparison purposes, we also add in Fig. 9 prices implied by our model with $\theta = 0$.

Generally, our model produces prices that are above the quoted prices in the market, translating to the fact that $\theta > 0$ according to our discussion earlier. Further, note the strong seasonality pattern in both the theoretical and the market quoted futures curves, with winter contracts being much more expensive than summer contracts.

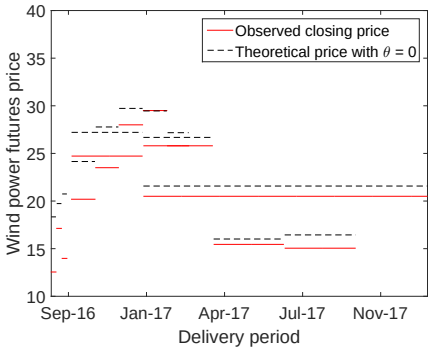


Fig. 9: Observed wind power futures curve on 1 September 2016 together with the corresponding theoretical curve implied by our proposed model with zero market price of risk. All prices are given in EUR/wph, and a tick size equal to EUR 100, i.e. we use the conversion 1% = 1 EUR/wph.

Having an explicit futures price formula facilitates the calibration of θ, which can be achieved through a minimization of the distance between theoretical and observed prices. To distinguish between theoretical prices implied by our model and market prices, let $f_\theta(t, T_1, T_2)$ denote the theoretical price, emphasizing its dependence on θ. Further, let $f^{Obs}(t, T_1, T_2)$ denote the corresponding closing price observed in the market. To extract the implied market price of risk associated with the contracts in Fig. 9, we consider the following:

$$\hat{\theta}(t, T_1, T_2) = \min_{\theta} |f^{Obs}(t, T_1, T_2) - f_\theta(t, T_1, T_2)|.$$

Implementing this procedure[3] yields a market price of risk per contract and valuation date. The obtained values are tabulated in Table 3, confirming that the implied values for $\hat{\theta}$ are generally positive.

Next, we briefly turn our attention to the risk premium, defined as

$$RP(t, T_1, T_2) = f^{Obs}(t, T_1, T_2) - f_{\theta=0}(t, T_1, T_2).$$

Owing to our model construction, notice that we will generally have an alternating sign between the implied market price of risk and the risk premium, i.e. $\theta > 0$ implies $RP < 0$ and vice versa.

So far in our analysis, we have restricted our attention to a single observed futures curve. Based on this, it is of course difficult to comment on general tendencies regarding the market price of risk in the wind power futures market. In a stylized situation, to have a time series for a given contract could be very interesting, since this would render the time series properties of the market price of risk visible. However, it may be problematic that the

[3]We applied the **fmincon** function in Matlab.

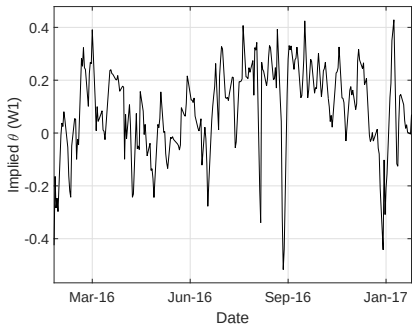
Contract type	Delivery period	Implied $\hat{\theta}$ 1 September 2016
1	Week 1	0.2411
2	Week 2	0.0980
3	Week 3	0.2385
4	Month 1	0.1213
5	Month 2	0.1140
6	Month 3	0.0437
7	Month 4	-0.0005
8	Month 5	0.0382
9	Quarter 1	0.0684
10	Quarter 2	0.0252
11	Quarter 3	0.0268
12	Quarter 4	0.0634
13	Year 1	0.0376

Table 3: The implied market price of risk on 1 September 2016.

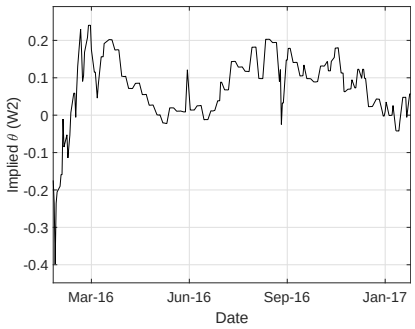
contracts move in time to maturity (time to start of delivery), suggesting that the various market prices of risk are not directly comparable. An alternative approach would be to find one market price of risk every day, given by a θ that minimizes the distance of the theoretical curve to all the available futures contracts that day. Then, we would get a series of market prices of risk for the whole market. This hides potential dependencies on time to delivery and length of delivery, but will nonetheless reveal the risk premium sign, and potentially if there are any interesting time series properties for the market price of risk. To gain more insight, we address next both types of investigations mentioned here.

The data we consider are observed wind power futures curves for the period from 1 February 2016 to 31 January 2017, amounting to a total of 257 curves. Each curve consists of 13 observed prices corresponding to the contract types specified in Table 3. Further, we consider static parameter estimates, that is, the ones obtained in Sec. 3.1. Ideally, the model should be recalibrated each day in the interval from 1 February 2016 to 31 January 2017, but lack of a synthetic index time series constructed for each of the valuation dates impedes such analysis. Nevertheless, we do not believe that the market has undergone significant changes in the period 1 February 2016 to 31 January 2017 relative to September 2016, thus justifying our study.

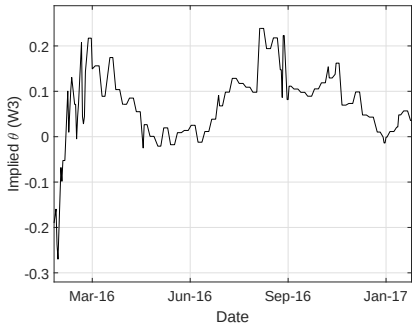
Performing the same analysis as the one given in Table 3 on all wind power futures curves yields a time series of implied $\hat{\theta}$s for each one of the 13 contract types. In Fig. 10, we plot some examples. Despite the lack of direct comparability caused by the presence of a strong seasonal effect, Fig. 10 highlights some interesting features. First, notice that the implied market price of



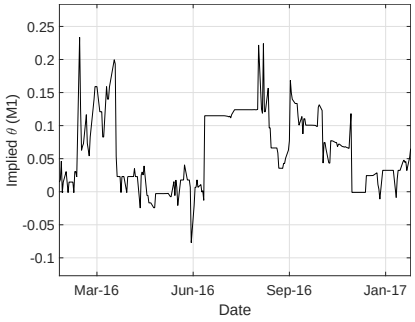
(a) Front week (W1)



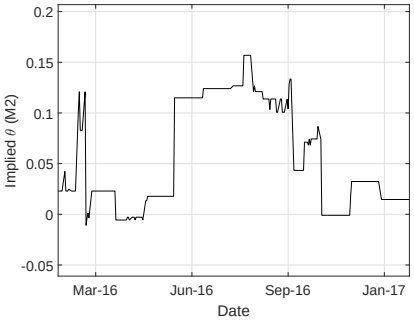
(b) Second front week (W2)



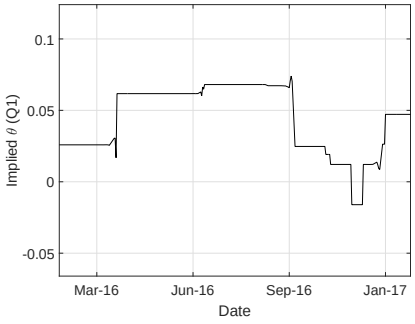
(c) Third front week (W3)



(d) Front month (M1)



(e) Second front month (M2)



(f) Front quarter (Q1)

Fig. 10: Implied $\hat{\theta}$s obtained by minimizing the distance between the observed futures prices and the theoretical prices for a given day and contract type. The time series stretches from 1 February 2016 to 31 January 2017.

risk corresponding to shorter deliveries is more volatile. Especially the front week contract series exhibits a singular behavior, with one possible explanation being the valuable information encompassed in short-term weather forecasts. Since this forward-looking information is not included in our model, the computed values for $\hat{\theta}$ for the front week contracts contain both a market price of risk as well as a sort of information premium. For longer delivery periods or start deliveries that lie further away from the valuation date, the information from weather forecasts becomes less reliable and hence its effect diminishes. Second, we mention that the market for wind power futures is still very illiquid, and especially Figs. 10(e) and 10(f) illustrate this through the long periods with an unchanged implied $\hat{\theta}$. For contracts with shorter delivery lengths, we stress that these contracts roll more often, thus 'forcing' the price to change regardless of the trading activity.

Averaging across the implied $\hat{\theta}$s for each of the contract types produces the values displayed in Fig. 11. We observe that all mean values are positive, consolidating our earlier findings relating to a positive θ (and hence a negative risk premium). Also notable is the decay in mean values with the length of delivery period. Possible explanations for this behavior can be different actors operating in different segments of the market, viable weather forecasts for very near and short delivery periods, illiquidity and seasonality.

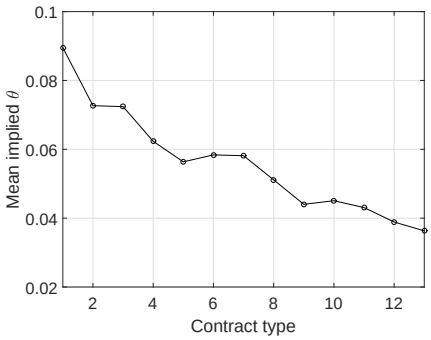


Fig. 11: Mean implied $\hat{\theta}$ per contract type, obtained by averaging across the daily implied θs corresponding to each of the contract types. The numbering of contract types coincides with that of Table 3.

Next, we compute a single implied θ at each t, based on all contracts making up the futures curve. That is, we consider the following minimization problem:

$$\hat{\theta}(t) = \min_{\theta} \sum_{k=1}^{K} \left| f^{Obs}\left(t, T_1^{(k)}, T_2^{(k)}\right) - f_\theta\left(t, T_1^{(k)}, T_2^{(k)}\right) \right|,$$

where $K = 13$ in our case, since each curve consists of 13 wind power futures contracts. As mentioned previously, an investigation of this type would

produce more comparable values for the market price of risk. The results are presented in Fig. 12, yielding that the implied $\hat{\theta}$ for the whole market is positive, which is not surprising considering our previous empirical findings.

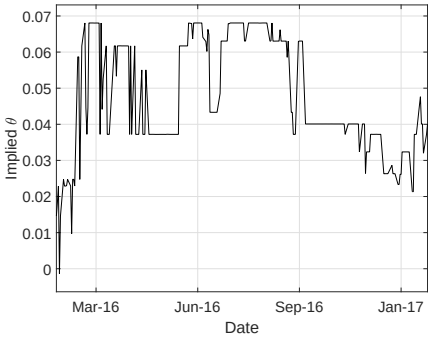


Fig. 12: Daily implied $\hat{\theta}$s obtained by minimizing the distance between all market prices and theoretical prices for a given valuation date. We consider historical wind power futures curves for the period 1 February 2016 to 31 January 2017.

Correlations and sign of risk premia

All the empirical analyses performed in this section point convincingly towards a negative risk premium, implying a wind power futures market that is in backwardation. Generally in typical commodity markets, the normal backwardation case is an expected situation, since the hedgers are usually the producers who are willing to accept a lower price (e.g. the futures price) than what is predicted in the spot. This seems to be the case for the German wind power futures market as well, and a possible explanation, as also stated in Gersema and Wozabal (2017), goes as follows: The production of single (or a collection of) wind parks is generally much more correlated to the average German wind power production than e.g. the production of a gas-fired power plant. Thus, wind power futures are a more powerful hedging tool for wind power generators than for conventional generators. As a result, the former group exhibits a higher demand and is willing to accept a lower price when selling wind power futures – and hence the negative risk premium in the German wind power futures market.

To substantiate the claims stated above, we perform a concise empirical investigation: On one hand, we compute the correlation between the German synthetic wind index (cf. Fig. 1) and the historical wind power production index of 26 different German wind parks. This data consists of daily measurements from 1 January 2012 to 31 December 2015, and is provided to us by Neas Energy. Further, the 26 wind parks we consider differ in e.g. total installed capacity, number of wind turbines in the park and geographical

location.

On the other hand, we compute the correlation between the German synthetic wind index and the historical day-ahead spark spread. Note that the decision to run/not run of gas-fired power plants depends on whether or not the spark spread is positive, and hence the spark spread is a measure for the profitability of such plants. We compute the spark spread as the difference between the day-ahead electricity price in Germany and the day-ahead gas price in the NetConnect Germany hub scaled by a heat rate h, with $h \in [1.9, 2.4]$. This interval corresponds to an efficiency between approximately 42% and 53%, which reflects a realistic level according to e.g. figure 14 in the report by Ecofys (2014).

Since the interval from 1 January 2012 to 31 December 2015 is the 'common denominator' for the many different time series we consider here, all correlations are computed based on this time interval. The correlations between the generation of the 26 wind parks and the German synthetic wind index are illustrated in Fig. 13, hereby also showing the approximate geographical location of each individual wind park we consider.

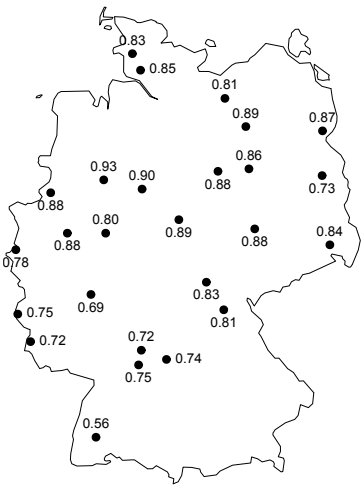


Fig. 13: Linear correlations between the German synthetic index and the actual wind power production index of 26 distinct wind parks.

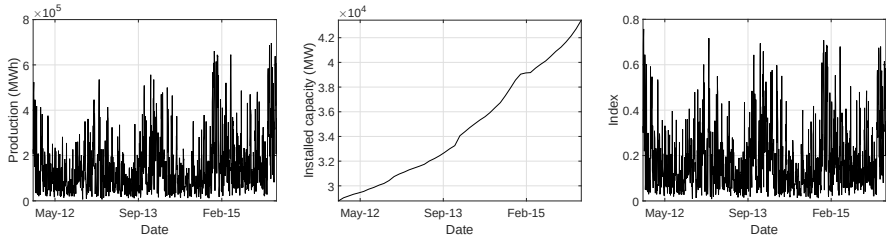
Fig. 13 shows a very strong positive relation (generally) between the production index of one specific wind park and the German index. When computing the linear correlation between the day-ahead spark spread and the German index, we get a value of -0.47 for the lowest h, i.e. $h = 1.9$. Increas-

ing h, that is assuming a less efficient gas-fired power plant, weakens the negative relation between the spark spread and the German index, rendering wind power futures less attractive as hedging instruments (for $h = 2.4$, we get a value of -0.44).

4.2 Synthetic vs. historical wind index

An essential point in the context of pricing wind power future is the distinction between the synthetic index (illustrated in Fig. 1) and the historical evolution of the actual German wind power production index. To highlight the importance of this distinction, we perform a comparison study in what follows.

Let us start with introducing the historical data we shall use for comparison: Because data for the actual historical index is not directly available to us, we construct the index using its two underlying data components. First, we consider the total wind power production on a daily basis corresponding to the period from 1 January 2012 to 31 December 2015 (a total of 1461 observations). Second, we consider monthly observations for the total installed wind power capacity for the same period, with monthly measurements having the 1st of each month as time stamps (a total of 49 observations, counting the measurement corresponding to 1 January 2016). The two data components are illustrated in Fig. 14(a) and Fig. 14(b), respectively, revealing the impressive growth that Germany has experienced over the considered period.



(a) Daily wind power production (b) Monthly installed capacity (c) Daily wind power production index

Fig. 14: Historical data for Germany.

A proxy for the actual daily index $\tilde{P}$ is then obtained as

$$\tilde{P}(t) = \frac{W(t)}{24C(t)}, \tag{18}$$

where $W(t)$ denotes the total wind power production in Germany at day t, and $C(t)$ denotes the total installed capacity in Germany at day t. Since the installed capacity data is measured at monthly intervals, intermediate daily

values are obtained by linear interpolation. The evolution of the index $\tilde{P}$ is displayed in Fig. 14(c), confirming that all measurements lie above 0 and below 1, as expected. The minimum measurement corresponds to a value of approx. 0.01 (a 1% utilization of the installed capacity), and the maximum measurement reaches approx. 0.76 (a 76% utilization of the installed capacity).

Next, we perform a linear regression of the actual index against the synthetic index, based on data in the interval 1 January 2012 to 31 December 2015 (corresponding to the four years that the two indexes have in common). We obtain an estimate for the intercept of 0.0055 (with standard error 0.0011) and an estimate for the slope of 0.8303 (with standard error 0.0039); a scatter plot of the actual index against the synthetic index is displayed in Fig. 15.

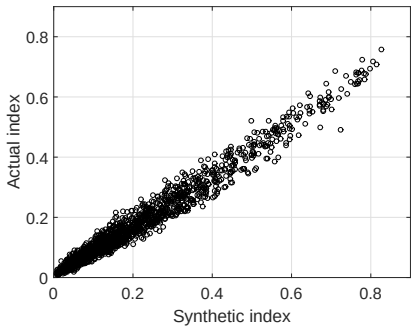


Fig. 15: Scatter plot of the actual index against the synthetic index.

The regression results translate to the fact that using the actual instead of the synthetic index data for model calibration would lead to an underestimation of the wind power futures price, which is indeed not surprising. One factor that helps explain this finding is that the synthetic index does not include information concerning intentional temporary switch off of turbines to reduce output, whereas the actual data does. With everything else being equal, this entails that the actual historical index must generally yield lower values than the synthetic index. Another contributing factor - and likely the most important - is the fixed installed capacity used to compute the synthetic index as opposed to the varying installed capacity used to compute the actual historical index, cf. Fig. 14(b). In continuation hereof, the expansion in wind turbine installations is centered in the wind-rich northern part of Germany, which is expected to have pulled the German index upwards. Hence, the synthetic index which is based on the newer installed capacity numbers for September 2016 is expected to have a higher mean than the actual index which is based on a varying installed capacity. In the context of pricing wind power futures, we are only interested in the available installed capacity on the valuation date, and not on its historical evolution. Lastly, we mention

that advancements in wind turbine technology can also be of relevance in this context.

To investigate how the use of the actual historical index influences the conclusions drawn in Section 4.1 regarding the risk premium, we have re-run all the computations performed in Sections 2.1, 3.1 and 4.1, based on this data. Surprisingly, not only does the model produce lower futures prices (with $\theta = 0$) as argued above, but the conclusions on the risk premium change drastically. Based on the newly calibrated model, we obtain a risk premium that is generally positive, implying a wind power futures market that is in contango, which is in contrast with our earlier findings. In a nutshell, using the 'wrong' data, i.e. the actual historical index, for the calibration of the model parameters has a significant impact, leading to very misleading conclusions.

5 Pricing options on wind power futures contracts

As a further demonstration of the advantages of our proposed model and the measure change using the Esscher transform, we consider here the pricing of European options written on wind power futures. While we acknowledge that these options are not traded on an exchange at the current time, they are potentially interesting, and hence this section is intended to provide an outlook.

Let us consider a call option on a wind power futures contract, where the exercise time of the option is T, the strike price is K, and r denotes a constant risk-free rate. To simplify calculations in what follows, we further assume that the maturity of the futures contract coincides with the exercise of the option, i.e., the call option is written on the actual wind power production. The call option price $C(t;T,K,T)$ can be expressed as the discounted conditional expectation of the future payoff under $\mathbb{Q}$, which is the pricing measure under the Esscher transform cf. Section 4. Hence,

$$\begin{aligned} C(t;T,K,T) &= \mathrm{e}^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[\max(F(T,T)-K,0)|\mathcal{F}_t] \\ &= \mathrm{e}^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[\max(P(T)-K,0)|\mathcal{F}_t] \\ &= \mathrm{e}^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[\max(A(T)\mathrm{e}^{Z(T)}-K,0)|\mathcal{F}_t], \end{aligned}$$

where

$$\begin{aligned} A(T) &= \Lambda(T)\exp(-X(t)\mathrm{e}^{-\alpha(T-t)} - \mu(1-\mathrm{e}^{-\alpha(T-t)})), \\ Z(T) &= -\int_t^T \mathrm{e}^{-\alpha(T-s)}\mathrm{d}L(s). \end{aligned}$$

Note that $A(T)$ can easily be computed given $P(t)$, the estimated seasonality function $\hat{\Lambda}$, and the speed of mean reversion $\hat{\alpha}$. To compute the call option

price based on our model framework, it is convenient to employ Fourier techniques, as suggested in Benth et al. (2008). Following Benth et al. (2008), we define the Fourier transform of a function $g \in L^1(\mathbb{R})$ as

$$\widehat{g}(y) = \int_{\mathbb{R}} g(x) \mathrm{e}^{-iyx} \mathrm{d}x. \tag{19}$$

If $\widehat{g} \in L^1(\mathbb{R})$, the inverse Fourier transform can be expressed as

$$g(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{g}(y) \mathrm{e}^{iyx} \mathrm{d}y.$$

Before proceeding to computing the price $C(t;T,K,T)$, we state the payoff function in terms of the Fourier transform.

Lemma 5.1. *For $a > 1$, we define*

$$g_T(x) = e^{-ax} \max\left(A(T)e^x - K, 0\right).$$

Then, we have that

$$\widehat{g}_T(y) = \frac{K}{(a-1+iy)(a+iy)} \left(\frac{K}{A(T)}\right)^{-(a+iy)},$$

where $\widehat{g}_T$ is the Fourier transform of g_T.

The result in Lemma 5.1 follows from employing the definition in Eq. (19). We note that the factor $\exp(-ax)$ in the definition of g_T is introduced due to the call option payoff not being a square-integrable function. For more details, we refer to Benth et al. (2008), Lemma 9.1, and Carr and Madan (1999). In the next Proposition, we derive the price $C(t;T,K,T)$.

Proposition 5.1. *Let $C(t;T,K,T)$ denote the price of a call option written on a wind power futures contract with strike K, exercise T, and delivery period of the futures contract T. The price $C(t;T,K,T)$ at time $t \leq T$ is given as*

$$C(t;T,K,T) = e^{-r(T-t)} \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{g}_T(y) \Xi(t,T) dy, \tag{20}$$

where

$$\Xi(t,T) = \left(\frac{\kappa_\theta + (a+iy)e^{-\alpha(T-t)}}{\kappa_\theta + a + iy}\right)^{\lambda_\theta/\alpha}.$$

Proof. See Appendix B.3.

We note that by having an analytical expression for the cumulant $\psi^{\mathbb{Q}}_{L(1)}$, the call option price $C(t;T,K,T)$ can easily be determined by solving the

integral in Eq. (20) numerically. Concerning the estimation of put option prices, these follow from the put-call parity.

We conclude this section by illustrating in Fig. 16 call option prices obtained by applying the formula in Proposition 5.1 for a series of strike indexes K. The valuation date t equals 31 December 2015, and we consider two different maturities, 1 July 2016 and 1 December 2016, as to emphasize the seasonal effects. Not surprisingly, the yearly seasonality in the wind index translates to the call options being cheaper for delivery during summer than during winter. Lastly, we note that the option prices could be multiplied by a tick size of EUR 100 in order to achieve comparability with the forward prices quoted in the market, see e.g. Fig 9.

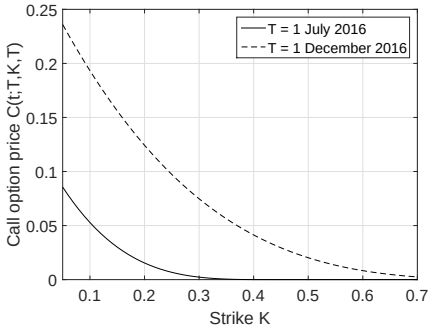


Fig. 16: Estimated call option prices as functions of the strike K at $t = 31$ December 2015. The computations are performed with $r = 0$, $a = 1.1$ and the parameter estimates obtained in Sections 2.1 and 3.1. We assume $\theta = 0$, and thus $\kappa_\theta = \kappa$, $\lambda_\theta = \lambda$.

6 Conclusion

In this paper we propose a non-Gaussian Ornstein-Uhlenbeck model for the wind power production index. The model has appealing characteristics, among others straightforward estimation of model parameters and analytical tractability. Motivated by the recent introduction of the German wind power futures on NASDAQ OMX and EEX, we employ the proposed model to conduct an empirical study on German data. First, the model is fitted to a synthetically generated time series of German wind power production indexes, revealing a good overall fit. Then, explicit prices for wind power futures are derived in the framework of no-arbitrage pricing. This facilitates the study of the market price of risk, which can be obtained by the usual practice of minimizing the distance between theoretical prices produced with our model and actual prices observed in the market. Based on historical wind power futures curves made up of closing prices from NASDAQ OMX, we perform different studies of the market price of risk.

Generally, we find evidence of a negative risk premium, whose magnitude decreases as the length of the delivery period increases. The negative risk premium suggests that wind power producers are willing to accept a lower price when selling wind power futures. As also argued in Gersema and Wozabal (2017), this behavior is due to wind power futures being a more powerful hedging tool for wind power generators than for conventional generators. This argument is enhanced by a brief empirical study, which demonstrates that the production of individual wind parks at different locations in Germany is more correlated to the German index than the production of conventional generators (here gas-fired power plants). Also, we find that the market price of risk is more volatile for shorter delivery periods, and argue that this behavior might be related to liquidity aspects and the information contained in short-term weather forecasts, which our model does not incorporate.

In this paper, we have restricted our attention to a constant market price of risk θ; admittedly, it is possible to allow for e.g. a seasonally varying θ in the Esscher transform. While it remains unclear whether this is backed by the data, one could potentially imagine a seasonality in the market. For a more general measure change, stochastic θ's (even being state dependent) could be considered as well, however this aspect is left for future research.

To highlight the importance of fitting our proposed model to a wind index that is generated assuming a constant as opposed to a varying level of installed capacity, we show through an empirical example that building on the 'wrong' data foundation can lead to the opposite conclusion regarding the sign of the risk premium. Finally, we address the pricing of European options written on wind power futures contracts, as to elaborate further on the benefits of the proposed modeling approach. Since an analytical expression for the cumulant is readily available, we show that the pricing of calls and puts can be achieved without difficulty.

Acknowledgments

Two anonymous referees are thanked for their constructive criticism and suggestions.

Funding

Fred Espen Benth acknowledges support from FINEWSTOCH, funded by the Norwegian Research Council. Anca Pircalabu acknowledges support from Innovation Fund Denmark, grant number 4135-00082B.

A Seasonality in the non-Gaussian Ornstein-Uhlenbeck model

The purpose of this appendix is to elaborate on issues related to seasonality in the non-Gaussian Ornstein-Uhlenbeck model for the wind power production index. Let us start by considering a simplified version of our model proposed in Eqs. (2)–(3), where $\Lambda(t) = 1$ and $\mu = 0$:

$$P(t) = \exp(-X(t)),$$
$$dX(t) = -\alpha X(t)dt + dL(t).$$

Since L is a subordinator, it follows that $X(t)$ is non-negative, and thus $P(t) \in [0, 1]$ is not violated with this model specification. While it is highly important to comply with the bound restrictions for $P(t)$, we cannot ignore the shortcoming of the above model regarding seasonality: The wind power production index has a strong seasonal component embedded in its dynamics, causing $\Lambda(t) \neq 1$ and $\mu \neq 0$ in reality.

To include seasonality, one possibility is to relax the assumptions imposed on $\Lambda(t)$ and μ above. This has led to our model proposed in Eqs. (2)–(3), and based on the variable $P(t)$, we have argued that $\mu = \ln M$, where $M = \max(\Lambda(t))$. However, if we let $m = \min(\Lambda(t))$ and instead regard the variable $P(t)/\Lambda(t)$, a renewed analysis yields that $P(t)/\Lambda(t) \in [0, 1/m]$, while $\exp(-X(t)) \in [0, \exp(-\mu)] = [0, 1/M]$. Since $m < M$, we have that $1/m > 1/M$, implying that the span of the data $P(t)/\Lambda(t)$ will be bigger than what our model, that is $\exp(-X(t))$, can capture. Hence, there is a potential of having

$$X(t) - \mu < 0,$$

which is somehow in contradiction with Proposition 3.2, where we obtained a Gamma distribution as the limiting distribution of $X(t) - \mu$.

Despite the slight inconsistency produced with our model specification in Eqs. (2)–(3), there are strong arguments in favor of this model. First, the above mentioned issue proves to have very limited impact in practice and second, the task of parameter estimation becomes more straightforward, since $\Lambda(t)$ can be fitted by ordinary least squares, and an estimate for μ immediately follows from the established relation $\mu = \ln M$.

Lastly, we mention that there are of course other alternatives concerning the inclusion of a seasonal component in the model. In our case for example, where we assume $L(t)$ to be a compound Poisson process with exponentially distributed jumps, such an alternative could be to consider seasonal instead of constant intensity. Such an approach will, in mean, provide us with a seasonally varying $P(t)$, while satisfying $P(t) \in [0, 1]$. However, this will come at the cost of having a complex empirical analysis, for which reason we have not pursued this approach in the present paper.

B Proofs

B.1 Proof of Proposition 3.2

Because

$$X(t) = X(0)\mathrm{e}^{-\alpha t} + \mu(1-\mathrm{e}^{-\alpha t}) + \int_0^t \mathrm{e}^{-\alpha(t-s)} dL(s),$$

the characteristic function of $X(t)$ becomes

$$\begin{aligned}\mathbb{E}[\mathrm{e}^{ixX(t)}] &= \exp\left((X(0)\mathrm{e}^{-\alpha t} + \mu(1-\mathrm{e}^{-\alpha t}))ix\right) \mathbb{E}[\mathrm{e}^{ix\int_0^t \mathrm{e}^{-\alpha(t-s)} dL(s)}] \\ &= \exp\left(ixX(0)\mathrm{e}^{-\alpha t} + ix\mu(1-\mathrm{e}^{-\alpha t}) + \int_0^t \psi_{L(1)}(x\mathrm{e}^{-\alpha(t-s)})ds\right).\end{aligned}$$

Further, the cumulant function of $L(1)$ denoted by $\psi_{L(1)}$ above is defined as

$$\begin{aligned}\psi_{L(1)}(x) &= \ln \mathbb{E}[\mathrm{e}^{ixL(1)}] \\ &= \ln \mathbb{E}\left[\mathbb{E}\left[\exp\left(ix\sum_{k=1}^{N(1)} J_k\right)\middle| N(1)\right]\right] \\ &= \ln \sum_{n=0}^{\infty} \mathrm{e}^{-\lambda}\frac{\lambda^n}{n!}\left(\mathbb{E}[\mathrm{e}^{ixJ}]\right)^n \\ &= \ln\left(\mathrm{e}^{-\lambda}\mathrm{e}^{\lambda\mathbb{E}[\mathrm{e}^{ixJ}]}\right) \\ &= \lambda(\mathbb{E}[\mathrm{e}^{ixJ}] - 1),\end{aligned}$$

with the second equality following from the definition of $L(t)$. Since J is an exponentially distributed random variable with density function given in Eq. (6), its characteristic function entering the expression of $\psi_{L(1)}$ is simply

$$\mathbb{E}[\mathrm{e}^{ixJ}] = \int_0^\infty \mathrm{e}^{ixy}\kappa\mathrm{e}^{-\kappa y}dy = \frac{\kappa}{\kappa - ix}.$$

Hence,

$$\psi_{L(1)}(x) = \lambda\frac{ix}{\kappa - ix}.$$

In the limit as $t \to \infty$, we obtain

$$\begin{aligned}\lim_{t\to\infty}\mathbb{E}[\mathrm{e}^{ixX(t)}] &= \exp\left(ix\mu + \int_0^\infty \psi_{L(1)}(x\mathrm{e}^{-\alpha s})ds\right) \\ &= \exp\left(ix\mu + \lambda\int_0^\infty \frac{ix\mathrm{e}^{-\alpha s}}{\kappa - ix\mathrm{e}^{-\alpha s}}ds\right) \\ &= \exp\left(ix\mu + \frac{\lambda}{\alpha}\ln\frac{\kappa}{\kappa - ix}\right) \\ &= \mathrm{e}^{ix\mu}\left(1 - i\frac{x}{\kappa}\right)^{-\frac{\lambda}{\alpha}},\end{aligned}$$

where we recognize the second factor of the above product as the characteristic function of the Gamma distribution. The result follows.

B.2 Proof of Proposition 4.1

Appealing to the adaptedness of $X(t)$ and recalling Eq. (2) and Eq. (3), we have that

$$\begin{aligned} F(t,T) &= \mathbb{E}^{\mathbb{Q}}[P(T)|\mathcal{F}_t] \\ &= \Lambda(T)\exp\left(-X(t)\mathrm{e}^{-\alpha(T-t)} - \mu(1-\mathrm{e}^{-\alpha(T-t)})\right) \\ &\quad \times \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^T \mathrm{e}^{-\alpha(T-s)}dL(s)\right)\,\middle|\,\mathcal{F}_t\right], \end{aligned}$$

where $\mathbb{Q}$ is the pricing measure obtained from the Esscher transform. Further, since L is a $\mathbb{Q}$-Lévy process and thus characterized by independent increments, we find

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^T \mathrm{e}^{-\alpha(T-s)}dL(s)\right)\,\middle|\,\mathcal{F}_t\right] &= \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\theta\int_t^T \mathrm{e}^{-\alpha(T-s)}dL(s)\right)\right] \\ &= \exp\left(\int_t^T \psi^{\mathbb{Q}}_{L(1)}(i\mathrm{e}^{-\alpha(T-s)})ds\right) \\ &= \exp\left(\int_0^{T-t} \psi^{\mathbb{Q}}_{L(1)}(i\mathrm{e}^{-\alpha s})ds\right), \end{aligned}$$

where $\psi^{\mathbb{Q}}_{L(1)}$ denotes the cumulant function of $L(1)$ under the Esscher transformed measure $\mathbb{Q}$. Using the Radon-Nikodym derivative in Eq. (13) and the definition in Eq. (14), it follows that the characteristic function of $L(1)$ under $\mathbb{Q}$ can be expressed as

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}}\left[\mathrm{e}^{ixL(1)}\right] &= \mathbb{E}\left[\mathrm{e}^{ixL(1)+\theta L(1)}\right]\mathrm{e}^{-\psi_{L(1)}(-i\theta)} \\ &= \exp(\psi_{L(1)}(x-i\theta) - \psi_{L(1)}(-i\theta)). \end{aligned}$$

Thus,

$$\psi^{\mathbb{Q}}_{L(1)}(x) = \psi_{L(1)}(x-i\theta) - \psi_{L(1)}(-i\theta).$$

Recalling the expression for $\psi_{L(1)}$ stated in Eq. (16), we obtain

$$\begin{aligned} \psi^{\mathbb{Q}}_{L(1)}(x) &= \lambda\left(\frac{i(x-i\theta)}{\kappa - i(x-i\theta)} - \frac{i(-i\theta)}{\kappa - i(-i\theta)}\right) \\ &= \lambda_\theta\left(\frac{ix}{\kappa_\theta - ix}\right), \end{aligned}$$

where

$$\kappa_\theta = \kappa - \theta,$$
$$\lambda_\theta = \frac{\lambda\kappa}{\kappa - \theta}.$$

Inserting all the information obtained above in the expression for $F(t,T)$ yields

$$\begin{aligned} F(t,T) &= \Lambda(T)\exp\left(-X(t)\mathrm{e}^{-\alpha(T-t)} - \mu(1-\mathrm{e}^{-\alpha(T-t)}) + \int_0^{T-t}\lambda_\theta\frac{-\mathrm{e}^{-\alpha s}}{\kappa_\theta+\mathrm{e}^{-\alpha s}}ds\right) \\ &= \Lambda(T)\mathrm{e}^{-\mu(1-\exp(-\alpha(T-t)))}\left(\frac{\kappa_\theta+\mathrm{e}^{-\alpha(T-t)}}{\kappa_\theta+1}\right)^{\lambda_\theta/\alpha}\left(\frac{P(t)}{\Lambda(t)}\right)^{\exp(-\alpha(T-t))}. \end{aligned}$$

The proposition follows.

B.3 Proof of Proposition 5.1

From Lemma 5.1, it follows that the call option price can be expressed as

$$C(t;T,K,T) = \mathrm{e}^{-r(T-t)}\frac{1}{2\pi}\int_{\mathbb{R}}\widehat{g}(y)\mathbb{E}^{\mathbb{Q}}[\mathrm{e}^{(a+iy)Z(T)}|\mathcal{F}_t]\mathrm{d}y.$$

Recall from the derivation of the wind power futures price in Appendix B.2 that the conditional expectation above is given by

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}}[\mathrm{e}^{(a+iy)Z(T)}|\mathcal{F}_t] &= \mathbb{E}^{\mathbb{Q}}[\mathrm{e}^{-(a+iy)\int_0^{T-t}\mathrm{e}^{-\alpha s}\mathrm{d}L(s)}|\mathcal{F}_t] \\ &= \exp\left(\int_0^{T-t}\psi^{\mathbb{Q}}_{L(1)}((ai-y)\mathrm{e}^{-\alpha s})\mathrm{d}s\right). \end{aligned}$$

The result follows from a straightforward calculation, where the expression for $\psi^{\mathbb{Q}}_{L(1)}(x)$ derived in Appendix B.2 must be employed.

C Further details on the estimation of α

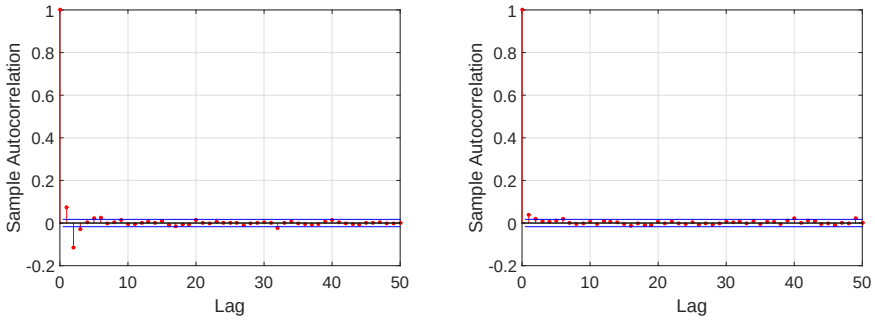
In this appendix, we consider the estimation of α cf. Eq. (3) by performing an AR(1)-estimation. Recalling the discussion and notation in Section 3.1, fitting an AR(1) to $X(t)-\mu$ yields $\hat{\phi} = 0.6067$, with a standard error of 0.0070. This implies that

$$\hat{\alpha} = 0.4997,$$

which is different (but not too far) from the estimate obtained in Section 3.1.

Considering the sample autocorrelation of the resulting residuals (Fig. 17(a)) and residuals squared (Fig. 17(b)), we note that an autoregressive

model of higher order could be beneficial. By computing the partial autocorrelation function of $X(t) - \mu$, this is indeed confirmed. Specifically, the pacf cuts off at lag 3, indicating that an AR(3) is preferred. With its continuous-time analogous, the CAR(3) model, we are however not ensured positivity, which is clearly an essential point in our modeling of the wind power production index. The proposed Ornstein-Uhlenbeck process in Eq. (3) is positive by design, while a similar CAR(3) process is not necessarily so. One could check case by case, but this is somehow cumbersome, and therefore not pursued further in the present study.



(a) Acf of residuals (b) Acf of squared residuals

Fig. 17: The acf of residuals and residuals squared after having fitted an AR(1) to $X(t) - \mu$.

References

D. Ackerer, D. Filipović, and S. Pulido. The Jacobi stochastic volatility model. *Swiss Finance Institute Research Paper*, No. 16-35, 2017.

O. E. Barndorff-Nielsen and N. Shephard. Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 63(2): 167 – 241, 2001.

A. Basse-O'Connor, S.-E. Graversen, and J. Pedersen. Stochastic integration on the real line. *Theory of Probability and Its Applications*, 58(2):193 – 215, 2014.

F. E. Benth and J. Šaltytė Benth. Dynamic pricing of wind futures. *Energy Economics*, 31(1):16 – 24, 2009.

F. E. Benth and J. Šaltytė Benth. Weather derivatives and stochastic modelling of temperature. *International Journal of Stochastic Analysis*, Vol. 2011:Article ID 576791, 21 pages, 2011.

F. E. Benth, J. Šaltytė Benth, and S. Koekebakker. *Stochastic Modelling of Electricity and Related Markets*. World Scientific, 2008.

D. Brody, J. Syroka, and M. Zervos. Dynamical pricing of weather derivatives. *Quantitative Finance*, 2(3):189 – 198, 2002.

S. D. Campbell and F. X. Diebold. Weather forecasting for weather derivatives. *Journal of the American Statistical Association*, 100(469):6 – 16, 2005.

M. Cao and J. Wei. Weather derivatives valuation and market price of weather risk. *Journal of Futures Markets*, 24(11):1065 – 1089, 2004.

P. Carr and D. B. Madan. Option valuation using the fast fourier transform. *Journal of Computational Finance*, 2:61 – 73, 1999.

M. Davis. Pricing weather derivatives by marginal value. *Quantitative Finance*, 1(3):305 – 308, 2001.

Ecofys. International comparison of fossil power efficiency and CO_2 intensity - Update 2014. ecofys.com, 2014.

F. Esscher. On the probability function in the collective theory of risk. *Skandinavisk Aktuarietidskrift*, 15:175 – 195, 1932.

G. Gersema and D. Wozabal. An equilibrium pricing model for wind power futures. *Energy Economics*, 65:64 – 74, 2017.

C. Gouriéroux and P. Valéry. Estimation of a Jacobi process. *Working paper*, version of November, 2002.

C. Halgreen. Self-decomposability of the generalized inverse gaussian and hyperbolic distributions. *Probability Theory and Related Fields*, 47(1):13 – 17, 1979.

W. K. Härdle and B. López Cabrera. Implied market price of weather risk. *Applied Mathematical Finance*, 19(1):59 – 95, 2012.

NASDAQ OMX. Nasdaq Renewables Wind Index Germany (DE). nasdaqomx.com, 2017.

A. Patton, D. N. Politis, and H. White. Correction to "automatic block-length selection for the dependent bootstrap". *Econometric Reviews*, 28(4):372 – 375, 2009.

A. Pircalabu and J. Jung. A mixed C-vine copula model for hedging price and volumetric risk in wind power trading. *Quantitative Finance*, 17(10): 1583 – 1600, 2017.

E. Platen and J. M. West. A fair pricing approach to weather derivatives. *Asia-Pacific Financial Markets*, 11(1):23 – 53, 2005.

D. N. Politis and J. P. Romano. The stationary bootstrap. *Journal of the American Statistical Association*, 89(428):1303 – 1313, 1994.

D. N. Politis and H. White. Automatic block-length selection for the dependent bootstrap. *Econometric Reviews*, 23(1):53 – 70, 2004.

K.-I. Sato. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, 1999.

Paper V

A seasonal copula mixture for hedging the clean spark spread with wind power futures

List of authors: Troels Sønderby Christensen[*,†], Anca Pircalabu[*,†] and Esben Høg[*]
[*] Department of Mathematical Sciences, Aalborg University
[†] Quantitative Analytics, Neas Energy

Submitted to
Energy Economics

The layout has been revised.

ABSTRACT

The recently introduced German wind power futures have brought the opportunity to address the problem of volume risk in wind power generation directly. In this paper, we study the hedging benefits of these instruments in the context of peak gas-fired power plants, by employing a strategy that allows trading in the day-ahead clean spark spread and wind power futures. To facilitate hedging decisions, we propose a seasonal copula mixture for the joint behavior of the day-ahead clean spark spread and the daily wind index. The model describes the data surprisingly well, both in terms of the marginals and the dependence structure, while being straightforward and easy to implement. Based on Monte Carlo simulations from the proposed model, the results indicate that significant benefits can be achieved by using wind power futures. Moreover, a comparison study shows that accounting for asymmetry, tail dependence, and seasonality in the dependence structure is especially important in the context of risk management.

1 Introduction

The sudden change in German energy policy that followed the Fukushima nuclear accident marked a new era for the German power market. Since the nuclear shutdown and the shift to renewables, Germany has experienced an impressive growth in both wind and solar power, and has reached a level that far exceeds the Kyōto climate obligations. This change has undoubtedly brought benefits on several fronts, however, the non-programmable nature of wind and solar electricity production has resulted in a large share of weather-dependent supply of electricity. From a financial point of view, the cash-flows from such non-programmable power plants can be incredibly volatile, not only due to price uncertainty, but also due to the uncertainty associated with the volume produced. While renewable generators are clearly affected by the uncertain volume, they are not the only ones; by market design and economics principle, the presence of renewables in the bid stack will always force conventional generators to produce less. In Germany, where the share of renewable energy is especially high, the conventional producers' competitiveness on e.g. the spot and forward markets has deteriorated, which has in turn invoked the need for far more intricate operation patterns and strategies.

In light of the advancements concerning renewables in Germany and the challenges imposed by volume risks for many different market players, the European Energy Exchange (EEX) recently introduced a financial instrument to mitigate the volume uncertainty associated with wind power generation. This instrument is referred to as a *wind power futures*, and its underlying is the German wind index. Representative agents for the sell and buy sides of wind power futures are the wind electricity producers and the conven-

tional electricity producers, respectively. On one hand, low wind scenarios are unfavorable for wind electricity producers, since they have a lowering effect on cash-flows; on the other hand, conventional generators are exposed to high wind scenarios, since a large share of wind power in the electrical grid displaces the costlier sources.

In this paper, we study the risk-reducing benefits of wind power futures in the context of conventional generators that operate in the day-ahead market whenever profitable. As a representative agent for the conventional generator, we consider the case of a peak gas-fired power plant whose profit per unit of electricity produced is measured in terms of the day-ahead clean spark spread. Since the dependence between the day-ahead clean spark spread and the wind index is essential for assessing the benefits of wind power futures, the contribution of this paper is twofold.

First, we propose a seasonal copula mixture to model the joint behavior of the day-ahead clean spark spread and the daily wind index. The model is fitted to four years of German data, and captures the marginal behavior of the individual variables and also the seasonality in the dependence between the variables very well. Second, we employ the proposed seasonal copula mixture to facilitate hedging decisions and showcase the effectiveness of wind power futures. To highlight the benefits of the seasonal copula mixture, we perform a study where the proposed model is compared against alternative models.

Owing to the recent introduction of the German wind power futures, the related literature is very scarce. The first related study is that of Gersema and Wozabal (2017), where the authors focus mainly on the pricing of wind power futures and explaining risk premia, for which an equilibrium pricing model is proposed. Also concentrating on the pricing aspect is the work of Benth and Pircalabu (2017), who apply a non-arbitrage approach to the pricing of wind power futures, and obtain results concerning the sign of risk premia that support the conclusions drawn in Gersema and Wozabal (2017). In contrast to the two existing studies, which focus mainly on pricing and less on hedging and risk management, we take a simplistic approach to pricing but study in detail aspects related to the risk-reducing ability of wind power futures. Nevertheless, we acknowledge that some of the results in Gersema and Wozabal (2017) and Benth and Pircalabu (2017) are very relevant in the context of the present study, and they shall thus be included in our discussion.

Turning to applications of copulas in energy markets, we mention that these models have gained substantial interest over the past years and have become a popular tool to model the non-linear dependence between different commodities. Some examples concerning applications of bivariate copulas are Börger et al. (2009), Benth and Kettler (2011), Grothe and Schnieders (2011), Avdulaj and Barunikl (2015), and Elberg and Hagspiel (2015). For ap-

plications beyond bivariate copulas, we mention the study of Pircalabu and Jung (2017), and that of Aepli et al. (2017). The present paper contributes to this stream of literature in terms of the application, which to the best of our knowledge has not yet been considered, and also in terms of modeling approach, by proposing an extension that deals with seasonality in the dependence structure.

The remaining of this paper is structured as follows: In Section 2, we introduce the data and elaborate on the construction of the variables. In Section 3, we describe the modeling framework and report estimation results. Section 4 introduces the seasonal copula mixture model and provides evidence for its quality of fit. In Section 5, we employ the proposed model to study the benefits of wind power futures, and perform various comparison studies. Section 6 concludes.

2 Background and data

To investigate the benefits of wind power futures for a gas-fired power plant (GFPP), two data components are of interest in the analysis performed in this paper, namely the day-ahead clean spark spread and the daily wind index. In this section, we address each of these in turn, commenting on their construction.

2.1 Clean spark spread

As an indicator for the profit per unit of electricity generated by a GFPP, we consider the day-ahead clean spark spread (CSS). This measure depends on electricity, gas, and emission prices, and also on the heat rate and the emission factor. The heat rate represents the required number of natural gas MWhs to produce one MWh electricity, i.e., the efficiency at which the GFPP transforms gas to electricity. Further, the emission factor represents the number of tons of CO_2 emitted by producing one MWh electricity.

With GFPPs being mainly peak-operated power plants—that is, power plants dispatching during the peak hours between 8 AM to 8 PM on weekdays and non-holidays—we consider the peak electricity price. Specifically, let S_t^E denote the day-ahead peak load electricity price, S_t^G the day-ahead gas price, and S_t^C the day-ahead emission price, with the subscript t indicating time measured in days. Further, let h be the heat rate and e the emission factor. We define the day-ahead CSS on day t as

$$CSS_t = S_t^E - hS_t^G - eS_t^C, \tag{1}$$

where S_t^E and S_t^G are measured in EUR/MWh, and S_t^C is measured in EUR/tCO_2.

Data preparation for the clean spark spread

To construct a time series for the day-ahead CSS, we consider the following time series:

- S_t^E: The German electricity price, which is computed as the average of all hourly electricity prices between 8 AM to 8 PM on weekdays and non-holidays. The source of this data is EEX.
- S_t^G: The day-ahead gas price for NetConnect Germany (NCG), which corresponds to the closing price. The source of this data is EEX.
- S_t^C: The EU Allowance unit of one tonne of CO_2 (EUA) phase 3 daily futures price. This data is collected from the Intercontinental Exchange, and represents the closing price. For more information regarding the EU emissions trading system, we refer the interested reader to European Commission (2017).

All time series above span 1030 observations in the period from 3 January 2013 to 30 December 2016, and cover weekdays that are non-holidays. To provide a sense of the data, we plot in Fig. 1 the time series corresponding to each of the three data sources described above. Moreover, we plot the CSS obtained by applying Eq. (1), and using the values for h and e reported in Table 1. These numbers are based on ICIS (2016), and shall be used in the remaining of this paper unless explicitly stated otherwise.

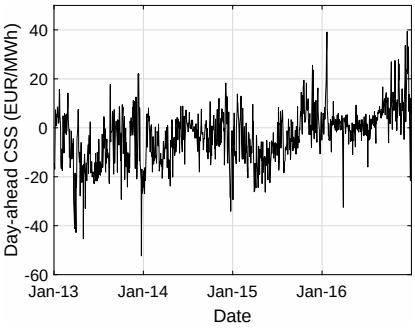
Heat rate h	Emission factor e
2.035	0.375

Table 1: Heat rate (MWhs natural gas per MWh electricity) and emission factor (tCO_2 per MWh electricity) based on ICIS (2016). The chosen heat rate corresponds to an efficiency of 49.13%.

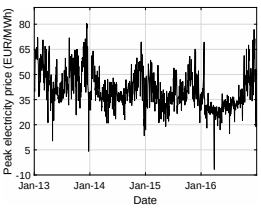
2.2 Wind index

Since the German wind power futures (WPF) were introduced only recently, we find it relevant to provide a brief description of these products and to clarify their payoff structure. WPF contracts are written on the average wind index in Germany, and can be traded at the European Energy Exchange (EEX) and Nasdaq OMX. In this paper, we shall restrict our attention to the WPF traded at EEX.

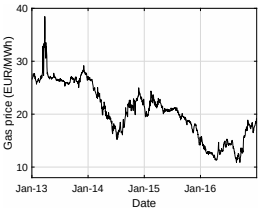
The German wind index is obtained as the ratio between the total wind power generation and the total available installed wind power capacity. Hence, the index is bounded between zero and one, and provides a measure of the German wind utilization. Currently, the delivery periods correspond to



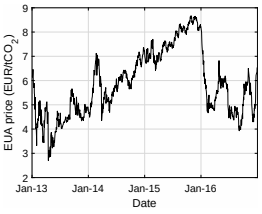
(a) CSS_t: Daily day-ahead CSS (peak load)



(b) S_t^E: German day-ahead electricity price (peak load)



(c) S_t^G: NCG gas day-ahead closing price



(d) S_t^C: EUA day-ahead closing price

Fig. 1: Historical evolution of the daily day-ahead CSS (peak load), the German day-ahead electricity price (peak load), the NCG day-ahead gas closing price, and the EUA day-ahead closing price, from 3 January 2013 to 30 December 2016. The applied heat rate and emission factor to construct the day-ahead CSS are given in Table 1.

weeks, months, quarters and years, and only trading the base load profile is possible. Compared to the definition of the day-ahead CSS data in Eq. (1), there is clearly a mismatch between delivery periods, with wind power futures hedging all hours of every day, and gas turbines generating output during peak hours. However, this reflects the present market conditions, where the volume risk of a GFPP can only be imperfectly hedged.

Assuming a delivery period $[T_1, T_2]$ consisting of H hours, the payoff corresponding to a long position in one WPF contract is given by

$$R^{WPF} = H \left(\underbrace{\frac{1}{T_2 - T_1 + 1} \sum_{t=T_1}^{T_2} W_t}_{=\bar{W}} - W_{t_0} \right) X, \tag{2}$$

where $W_t \in [0,1]$ is the daily wind index, $\bar{W} \in [0,1]$ is the realized average wind index over the delivery period, and $W_{t_0} \in [0,1]$ can be thought of as the "futures price", i.e., the index set at t_0 when entering the contract. Further,

X is a pre-specified constant tick size which is used to convert the index differences into monetary measures. According to EEX, $X = 100$ EUR. We see from Eq. (2) that a short position in WPF will generate a profit in low-wind scenarios, making it a useful hedging instrument for the wind power producer. Conversely, a long position will generate a profit in the high-wind scenarios, to which the GFPPs are exposed.

Data preparation for the wind index

The index that a WPF contract is settled against is externally provided by EuroWind. Since trading in WPF began only recently, the amount of data available on the spot wind index provided by EuroWind is limited. To obtain a longer time series, we consider instead a proxy wind index constructed using the wind power production in Germany on a daily basis, and monthly recordings of the German installed wind power capacity, which are updated at the start of each month. The wind power production data is collected from the four different transmission system operators in Germany, and the source of the installed capacity data is PointConnect.

Specifically, the daily German wind index is constructed as

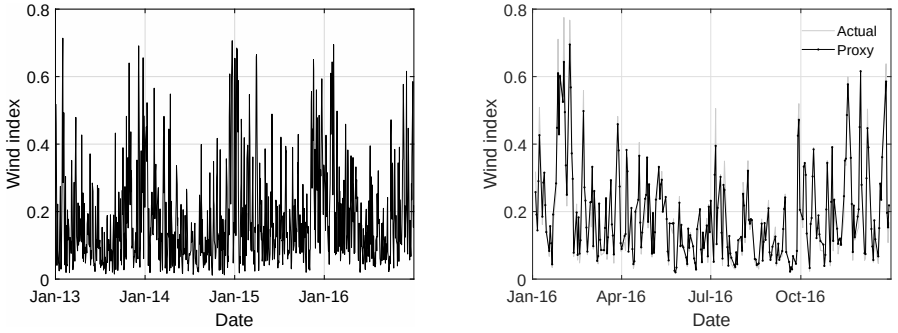
$$W_t = \text{Daily wind index} = \frac{\text{Daily wind power generation (MWh)}}{h \cdot \text{Installed capacity (MW)}},$$

where h denotes the number of hours in a given day, and the installed capacity on a daily basis is obtained by linear interpolation. In order to unify the length of the day-ahead CSS and the wind index, we omit weekends and holidays for the wind index data. Hence, the constructed index spans the period from 3 January 2013 to 30 December 2016, a total of 1030 observations, and is plotted in Fig. 2(a).

To provide some evidence for how the constructed wind index matches the true settlement data, we plot in Fig. 2(b) our proxy together with the one year of actual data from EuroWind that we have available. The time series plot reveals an acceptable resemblance, and to provide a quantitative indication, we compute the mean absolute error to 0.020.

3 Model construction and fit

To model the joint behavior of the day-ahead CSS and the daily wind index (henceforth referred to as simply CSS and wind index, respectively), we consider copula models. Restricting our presentation to the two-dimensional case, a copula is the joint distribution of the random variables U_1 and U_2, where each variable is marginally uniformly distributed as $U(0,1)$. Since our data exhibits seasonality and autocorrelation, we wish to filter out these



(a) Wind index (b) Comparison with EuroWind data

Fig. 2: Historical evolution of the German wind index (W_t) on a daily basis from 3 January 2013 to 30 December 2016, and a comparison of W_t with the actual wind index provided by EuroWind for the year 2016.

effects before applying the copula. Therefore, we are here considering the conditional copula.

Let $F(\cdot|\mathcal{F}_{t-1})$ denote the conditional joint distribution function of the random vector $\mathbf{Y}_t = (Y_{1t}, Y_{2t})$, and let $F_1(\cdot|\mathcal{F}_{t-1})$ and $F_2(\cdot|\mathcal{F}_{t-1})$ denote the conditional continuous marginal distribution functions of Y_{1t} and Y_{2t}, respectively. Then, according to Sklar's theorem [Sklar (1959)] for conditional distributions, there exists a unique copula C such that F can be decomposed as

$$F(y_{1t}, y_{2t}|\mathcal{F}_{t-1}) = C(F_1(y_{1t}|\mathcal{F}_{t-1}), F_2(y_{2t}|\mathcal{F}_{t-1})|\mathcal{F}_{t-1}). \tag{3}$$

The converse also holds, meaning that given two univariate distributions F_1, F_2 and a copula C, F as defined in Eq. (4) is the joint distribution with margins F_1, F_2. Thus, Sklar's theorem not only provides a way of *decomposing* a joint distribution function, but also a way of *composing* it given marginal distributions and a copula, both of which are very useful in practical applications. For the proof of Sklar's theorem for conditional distributions, we refer to Patton (2006).

Recalling the probability integral transform, we note that $U_{it} := F_i(Y_{it}|\mathcal{F}_{t-1}) \sim U(0,1)$, for $i = 1, 2$. Differentiating both sides of Eq. (4) with respect to (y_{1t}, y_{2t}) thus yields

$$f(y_{1t}, y_{2t}|\mathcal{F}_{t-1}) = c\left(u_{1t}, u_{1t}|\mathcal{F}_{t-1}\right) \cdot f_1(y_{1t}|\mathcal{F}_{t-1}) \cdot f_2(y_{2t}|\mathcal{F}_{t-1}), \tag{4}$$

where f denotes the joint density function, c is the copula density, and f_1, f_2 denote marginal density functions.

In our context, copula models are advantageous for various reasons: First, being able to capture dependence beyond the linear correlation can be of utmost importance when illustrating the hedging benefits of WPF, and this

can be achieved with copulas. Second, we can separate the treatment of the dependence structure from that of the marginal behavior of the individual variables (cf. Eq. (4)), since the dependence structure is fully contained in the copula. Third, selecting one type of marginal distribution for the first variable does not restrict our choice of marginal distribution for the second variable.

Turning to the estimation of the model parameters, we let T denote the sample size, $\boldsymbol{\theta}_c$ the copula parameters, and $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ the parameters of the marginal models. From Eq. (4) it follows that the log-likelihood function is

$$\log \mathcal{L} = \sum_{t=1}^{T} \log c\left(u_{1t}, u_{1t} | \mathcal{F}_{t-1}, \boldsymbol{\theta_c}\right) + \sum_{t=1}^{T} \log f_1(y_{1t} | \mathcal{F}_{t-1}, \boldsymbol{\theta_1}) + \sum_{t=1}^{T} \log f_2(y_{2t} | \mathcal{F}_{t-1}, \boldsymbol{\theta_2}).$$

Here, we apply a two-stage maximum likelihood estimation procedure. In a first step, the parameters of the marginal models are obtained; in a second step, the copula parameters are estimated while keeping fixed the estimates from the first step. This provides a far less complicated estimation procedure relative to the one-stage procedure. For more details on the relative efficiency of the two-stage compared to the one-stage maximum likelihood estimation, we refer to Joe (2005).

In the following two sections, we present in detail the marginal models and the constant copulas considered in this paper, and provide empirical evidence for the fit of these models to our data.

3.1 Marginal models

Since both the CSS and the wind index exhibit seasonality, we start the marginal treatment of the individual variables by applying suitable seasonal functions to remove the deterministic seasonal component.

For the CSS, we consider the seasonal function

$$f_t = a_1 + b_1 t + c_1 \sin\left(2\pi t / K\right) + c_2 \cos\left(2\pi t / K\right),$$

where a_1 is a constant, b_1 is the trend coefficient, and c_1 and c_2 are coefficients for the annual cycle. We have on average approximately $K = 258$ observations per year.

Not surprisingly, the seasonality function for the CSS resembles a seasonality function that would typically be considered for the day-ahead electricity price (see e.g. Haldrup and Nielsen (2006), Benth and Benth (2011), and Härdle and Cabrera (2012)). This resemblance is caused by the magnitude of the electricity price compared to the gas and emission price, cf. Fig. 1, causing

the former to have the dominant effect. Aside from electricity prices usually exhibiting a yearly seasonality, a strong within-week seasonality is also often observed. However, with the exclusion of weekends from our data, adding a term that addresses the weekly seasonality (e.g. day-of-week dummies) is unnecessary. Furthermore, adding more trigonometric terms (based on the periodogram) does not improve the fit of the seasonal function substantially.

Turning to the wind index, recall that this series is bounded between 0 and 1, cf. Fig. 2. Following Pircalabu and Jung (2017), we apply the logit-transform to the wind index[1], and consider the following seasonal function for the logit wind index (LWI):

$$f_t = a_1 + c_1 \sin(2\pi t/K) + c_2 \cos(2\pi t/K),$$

which is motivated by the prominent annual cycles we observe in the sample autocorrelation of the LWI. Also here, different meaningful extensions of the seasonal function were experimented with, without yielding a significant improvement.

The seasonality functions are fitted to the data by ordinary least squares, and Table 2 summarizes the results obtained for the CSS and the LWI.

	$\hat{a_1}$	$\hat{b_1}$	$\hat{c_1}$	$\hat{c_2}$
CSS	-10.131	0.014	-3.611	2.990
	(0.579)	(0.001)	(0.411)	(0.404)
LWI	-1.828	-	0.066	0.472
	(0.030)	-	(0.042)	(0.043)

Table 2: OLS estimates for parameters of the seasonal functions for the CSS and the LWI. Standard errors are reported in parenthesis, and are based on a naive OLS calculation.

Next, we apply ARMA-GARCH filters to the deseasonalized data. Given a time series of data y_t, an ARMA(p,q)–GARCH(h,k) model is defined by

$$\begin{aligned} y_t &= \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t, \\ \varepsilon_t &= \sigma_t \eta_t, \\ \sigma_t^2 &= \omega + \sum_{i=1}^{h} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{k} \beta_j \sigma_{t-j}^2, \end{aligned}$$

where $\eta_t \sim iid\ \mathrm{N}(0,1)$. It was Engle (1982) who introduced the ARCH model, and later Bollerslev (1986) who extended the variance equation to include lagged values of σ_t^2. For a review of ARMA and GARCH models, see for

[1]The logit function is given by $\mathrm{logit}(x) = \log(x) - \log(1-x)$.

example Shumway and Stoffer (2011). In the following, we denote by η_t^{CSS} and η_t^{LWI} the standardized residuals resulting from applying the ARMA–GARCH models to the CSS and the LWI, respectively.

Model selection is based on the Bayesian Information Criterion (BIC), and we consider ARMA(p,q)–GARCH(h,k) models for all possible combinations of p,q,h and k, for $p = 0,\ldots,7$, $q = 0,\ldots,7$, $h = 0,1,2$, and $k = 0,1,2$. The optimal order of the models and the corresponding estimated parameters are reported in Table 1.

	CSS	**LWI**
Model	ARMA(2,1) – GARCH(1,1)	ARMA(1,1)
	Conditional mean	
AR1 $\hat{\phi}_1$	1.313 (0.077)	0.360 (0.055)
AR2 $\hat{\phi}_2$	-0.360 (0.058)	-
MA1 $\hat{\theta}_1$	-0.818 (0.064)	0.208 (0.055)
Variance $\hat{\sigma^2}$	-	0.677 (0.030)
	Conditional variance	
Constant $\hat{\omega}$	3.611 (0.636)	-
ARCH $\hat{\alpha_1}$	0.110 (0.022)	-
GARCH $\hat{\beta_1}$	0.827 (0.027)	-

Table 3: Type and order of marginal models, parameter estimates and corresponding standard errors in parenthesis.

Considering the goodness-of-fit of the normal distribution, we find a satisfactory fit in the case of $\hat{\eta}_t^{LWI}$, cf. Figs. 3(e) and 3(f). This is however not the case for $\hat{\eta}_t^{CSS}$. Consequently, we relax the normality assumption for the CSS, and consider instead the normal-inverse Gaussian (NIG) distribution. The probability density function of the NIG distribution is given by

$$g(x|\alpha,\beta,\mu,\delta) = \frac{\alpha\delta G_1\left(\alpha\sqrt{\delta^2+(x-\mu)^2}\right)}{\pi\sqrt{\delta^2+(x-\mu)^2}} e^{\delta\sqrt{\alpha^2-\beta^2}+\beta(x-\mu)},$$

where

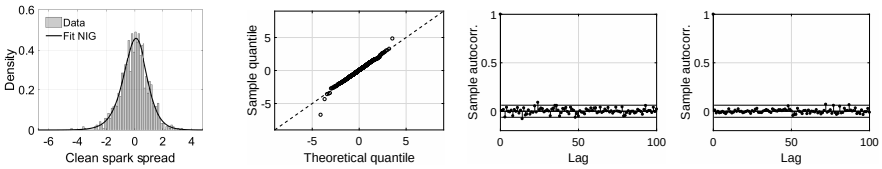
$$G_1(x) = \frac{1}{2}\int_0^\infty e^{-\frac{1}{2}x(t+t^{-1})}dt$$

is the modified Bessel function of third kind and index 1. The NIG distribution is a popular choice in the financial literature (for some examples, see Barndorff-Nielsen (1997a), Rydberg (1997), Barndorff-Nielsen (1997b), and Jensen and Lunde (2001)), and is also often able to provide a good description of commodity data, see e.g. Benth and Benth (2004) and Benth and Kettler

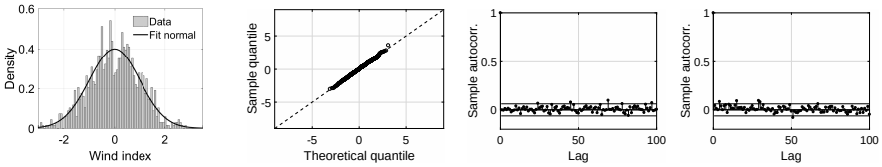
(2011). The NIG distribution is fitted to the residuals from the ARMA(2,1)–GARCH(1,1) model cf. Table 1 by maximum likelihood, and the parameter estimates are reported in Table 4. As it appears from the histogram and quantile plots displayed in Figs. 3(a) and 3(b), the NIG distribution provides a satisfactory fit to the CSS data.

$\hat{\alpha}$	$\hat{\beta}$	$\hat{\mu}$	$\hat{\delta}$
1.584 (0.309)	-0.189 (0.144)	0.189 (0.127)	1.534 (0.282)

Table 4: Maximum likelihood estimates obtained by fitting the NIG distribution to $\hat{\eta}_t^{CSS}$. Corresponding standard errors are given in parenthesis.



(a) Histogram for $\hat{\eta}_t^{CSS}$ (b) Quantile plot for $\hat{\eta}_t^{CSS}$ (c) Acf for $\hat{\eta}_t^{CSS}$ (d) Acf for squared $\hat{\eta}_t^{CSS}$



(e) Histogram for $\hat{\eta}_t^{LWI}$ (f) Quantile plot for $\hat{\eta}_t^{LWI}$ (g) Acf for $\hat{\eta}_t^{LWI}$ (h) Acf for squared $\hat{\eta}_t^{LWI}$

Fig. 3: Diagnostics for the standardized residuals $\hat{\eta}_t^{CSS}$ (first row) and $\hat{\eta}_t^{LWI}$ (second row).

To provide further evidence for the appropriateness of the chosen marginal distributions, we perform the Kolmogorov-Smirnov (K-S) and the Cramer-von Mises (CvM) goodness-of-fit tests. To obtain critical values for the tests, we employ the simulation-based method described in detail in Patton (2013). In the CSS case we obtain p-values of 0.627 and 0.785 for the K-S and CvM test, respectively, and can thus not reject the null that the NIG distribution is well-specified. This is also the conclusion in the LWI case, where we test the goodness-of-fit of the normal distribution. Here, the resulting p-values are 0.746 and 0.915 for the K-S and CvM test, respectively.

Aside from providing evidence for the goodness-of-fit of the marginal distributions, the sample autocorrelations provided in Fig. 3 suggest that no considerable serial dependence is left in the conditional mean and variance,

for either variable. Having verified that the models proposed here are suitable for describing the marginal behavior of the CSS and the LWI, we proceed in the next section to the modeling of the dependence structure.

3.2 Constant copula models

Let F^{NIG} and F^{N} denote the cumulative distribution functions for the NIG and standard normal distribution, respectively. To obtain the approximately uniforms that are the input variables to the copula function, we apply the probability integral transform, i.e.,

$$\hat{u}_t^{CSS} = F^{NIG}(\hat{\eta}_t^{CSS}|\mathcal{F}_{t-1}, \hat{\alpha}, \hat{\beta}, \hat{\mu}, \hat{\delta}),$$
$$\hat{u}_t^{LWI} = F^{N}(\hat{\eta}_t^{LWI}|\mathcal{F}_{t-1}),$$

for $t = 1, \ldots, T$. In Fig. 4 we plot the results, revealing a strong negative relation between the variables.

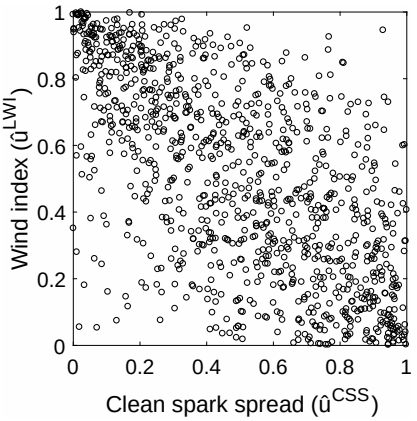


Fig. 4: Empirical copula density.

From an economic point of view, this negative dependence is caused by the strong negative relation between the electricity price and the wind index. A high wind penetration in the electricity grid puts a downward pressure on day-ahead electricity prices owing to the process of day-ahead price formation, which prioritizes cheap electricity producers. With everything else being equal, this lowering effect on the electricity price is then reflected in the CSS, which is also lowered, cf. Eq. (1). Similar arguments apply to the case of a low wind scenario, where electricity prices are typically pushed upwards.

The dependence structure seems to be slightly asymmetric, with the north west corner of Fig. 4 exhibiting more concentration and being sharper in shape compared to the south east corner. That is, there seems to be more

probability of observing the combination of extremely high wind / extremely low CSS than the reverse. Since non-zero dependence in extreme events could have notable implications for the benefits of WPF, being able to capture such behavior in a model for the dependence structure must be considered. Luckily in the context of copulas, such extreme events can be easily captured by considering certain copula families that allow for non-zero *tail dependence*. More specifically, the lower and upper tail dependence can be defined as

$$\Lambda^l = \lim_{q \to 0^+} P(u_t^{CSS} \leq q | u_t^{LWI} \leq q)$$
$$\Lambda^u = \lim_{q \to 1^-} P(u_t^{CSS} > q | u_t^{LWI} > q),$$

where q denotes the quantile. Clearly, since our data is characterized by negative dependence, computing Λ^l and Λ^u as defined above is not meaningful. This can however be resolved by performing suitable rotations of the data, which shall be discussed in more detail shortly.

To investigate which copula best describes the dependence structure illustrated in Fig. 4, we consider first the following standard copulas, which are often employed in the related literature: Gaussian, Gumbel, rotated Gumbel (RGumbel), Clayton, rotated Clayton (RClayton), Frank, symmetrized Joe-Clayton, and Student t.[2] These copula models cover a wide range of dependency structures, with some models being able to capture asymmetric dependence, and also upper and lower tail dependence, i.e., a non-zero probability of extreme events happening simultaneously. In the interest of brevity, we shall not go into detail with the properties of each copula model here, and refer instead to McNeil et al. (2005), Nelsen (2006), and Patton (2006) for a comprehensive description.

To allow for further flexibility compared to the standard copulas enumerated above, we also consider copula mixtures. As in e.g. Rodriguez (2007) and Dias and Embrechts (2009), for a given t we mix copula a, having copula density $c^a(\cdot|\mathcal{F}_{t-1}, \boldsymbol{\theta}^a)$, with copula b, having copula density $c^b(\cdot|\mathcal{F}_{t-1}, \boldsymbol{\theta}^b)$, by using a mixing parameter $0 < \lambda < 1$ and the following form:

$$c^m(\cdot|\mathcal{F}_{t-1}, \boldsymbol{\theta}^a, \boldsymbol{\theta}^b, \lambda) = \lambda c^a(\cdot|\mathcal{F}_{t-1}, \boldsymbol{\theta}^a) + (1-\lambda) c^b(\cdot|\mathcal{F}_{t-1}, \boldsymbol{\theta}^b). \tag{5}$$

As expected, a mixture copula inherits characteristics from its mixing components. In the following proposition, we present an especially useful result relating to the tail dependence of a mixture copula, which we shall use shortly. Notice that we omit the conditioning to ease the notation.

Proposition 3.1. *Let $U_i \sim Unif(0,1)$ for $i = 1, 2$, and let C^m denote the bivariate copula of (U_1, U_2). Further assume C^m is given as the mixture*

$$C^m(u_1, u_2) = \lambda C^a(u_1, u_2) + (1-\lambda) C^b(u_1, u_2),$$

[2] By rotated, we mean a 180 degree rotation of the data.

where C^a and C^b are two bivariate copulas, and $0 < \lambda < 1$. Then, the lower tail dependence Λ^l and the upper tail dependence Λ^u for the mixture C^m are given as

$$\Lambda^l = \lambda\Lambda^{l,a} + (1-\lambda)\Lambda^{l,b},$$

and

$$\Lambda^u = \lambda\Lambda^{u,a} + (1-\lambda)\Lambda^{u,b},$$

where $\Lambda^{l,a}$, $\Lambda^{l,b}$, $\Lambda^{u,a}$, and $\Lambda^{u,b}$ are the respective tail dependence measures for C^a and C^b.

Proof. See Appendix A.1.

Moving on to the estimation aspect, we let $c(\cdot|\mathcal{F}_{t-1}, \boldsymbol{\theta})$ denote the conditional copula density with parameter vector $\boldsymbol{\theta} \in \mathbb{R}^l$, where $l \in \mathbb{N}$ is the number of parameters in the copula. For each copula model, we obtain an estimate for $\boldsymbol{\theta}$ by maximizing the copula log-likelihood, i.e.,

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \sum_{t=1}^{T} \log c(\hat{u}_t^{CSS}, \hat{u}_t^{LWI}|\mathcal{F}_{t-1}, \boldsymbol{\theta}). \tag{6}$$

We note that it is only the Gaussian and Student t copulas that allow for negative dependence. To fit the remaining copulas to our data, we perform suitable rotations of the data. Specifically, we rotate around the LWI variable for the case of the Gumbel, Clayton, Frank, and symmetrized Joe-Clayton copulas, and consider thus the pair $(\hat{u}_t^{CSS}, 1-\hat{u}_t^{LWI})$ as input to Eq. (6) for these models. Regarding tail dependence, the rotation of data implies that lower tail dependence for the estimated copulas corresponds to high wind index / low CSS scenarios (north west corner of Fig. 4), whereas upper tail dependence for the estimated copulas corresponds to low wind index / high CSS scenarios (south east corner of Fig. 4). To fit the RGumbel and the RClayton, we note that a further 180 degree rotation of the pair $(\hat{u}_t^{CSS}, 1-\hat{u}_t^{LWI})$ is performed.

In Table 5, we report the estimation results for all standard copula models and three selected mixtures. Other copula mixtures aside from those reported in Table 5 were considered, but we found no increase in performance. As a model selection criterion, we employ the Akaike Information Criterion (AIC). According to the AIC, the preferred model is the mix of Frank and RGumbel (hereafter denoted FRG copula), confirming the presence of slight asymmetry in the dependence structure illustrated in Fig. 4.

Considering the FRG copula in more detail, its first mixing component, the Frank copula, imposes symmetric dependence and a zero tail dependence. Its second mixing component, the RGumbel, imposes an asymmetric

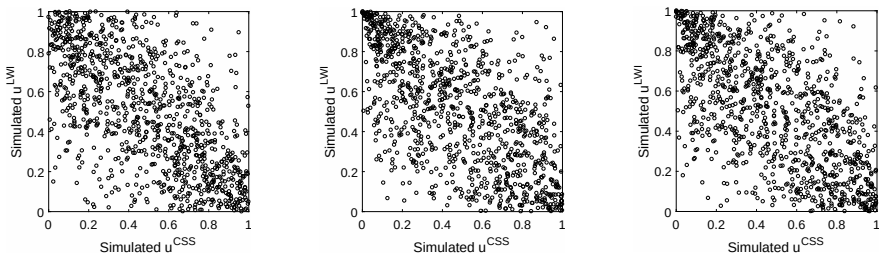
dependence structure, with zero upper tail dependence and lower tail dependence given by

$$\Lambda^{l,RG} = 2 - 2^{1/\theta^{RG}},$$

where θ^{RG} is the parameter for the RGumbel copula. Recalling Prop. 3.1, we thus have that the upper and lower tail dependence for the FRG copula are

$$\begin{aligned} \Lambda^{u,FRG} &= 0, \\ \Lambda^{l,FRG} &= (1-\lambda)\Lambda^{l,RG}. \end{aligned} \tag{7}$$

The fit produced by the FRG translates into a tail dependence coefficient of approximately 0.359 when considering the north west corner of Fig. 4, and hence a rather high probability of extremely high wind index / low CSS happening simultaneously. To illustrate the shape of the FRG copula and how it deviates from the shapes of the individual copulas in the mixture, we plot in Fig. 5 simulations from the fitted Frank, RGumbel and FRG copulas. The simulations reveal that while the fitted Frank copula is too symmetric and the fitted RGumbel is too asymmetric compared to the observed dependence in Fig. 4, the fitted FRG mixture is able to dampen the individual effects, hence providing a better resemblance to the observed dependence structure.



(a) Frank (b) RGumbel (c) FRG

Fig. 5: T simulations from the fitted Frank, RGumbel and FRG copulas, cf. Table 5.

Copula model		Param. (s.e.)	log $\mathcal{L}_c$	AIC
Gaussian	$\hat{\rho}$	-0.636 (0.019)	266.503	-531.006
Gumbel*	$\hat{\theta}$	1.675 (0.049)	226.083	-450.165
RGumbel*	$\hat{\theta}$	1.773 (0.053)	277.829	-553.657
Clayton*	$\hat{\theta}$	1.233 (0.083)	245.947	-489.894
RClayton*	$\hat{\theta}$	0.911 (0.071)	167.862	-333.724
Frank*	$\hat{\theta}$	5.029 (0.247)	267.805	-533.611
Sym. Joe-Clayton*	$\hat{\Lambda}^u$	0.274 (0.048)	265.777	-527.553
	$\hat{\Lambda}^l$	0.539 (0.026)		
Student t	$\hat{\rho}$	-0.646 (0.020)	274.596	-545.192
	$\hat{\nu}$	9.873 (15.008)		
Mix of Gumbel and RGumbel*	$\hat{\theta_1}$	1.964 (0.500)	285.060	-564.120
	$\hat{\theta_2}$	1.797 (0.101)		
	$\hat{\lambda}$	0.219 (0.081)		
Mix of Frank and RGumbel*	$\hat{\theta_1}$	4.552 (1.309)	286.419	-566.837
	$\hat{\theta_2}$	1.920 (0.169)		
	$\hat{\lambda}$	0.365 (0.104)		
Mix of Gaussian and RGumbel*	$\hat{\rho}$	-0.494 (0.092)	285.152	-564.304
	$\hat{\theta}$	2.095 (0.234)		
	$\hat{\lambda}$	0.357 (0.130)		

Table 5: Estimation results for 11 selected copula models. The maximized value of the copula log-likelihood is denoted log $\mathcal{L}_c$. For the functional forms of the considered copulas and other characteristics, we refer to McNeil et al. (2005), Nelsen (2006) and Patton (2006). A copula marked by an asterisk has been estimated using a suitable rotation of the data. Standard errors are based on 999 simulations.

4 Time-Varying Dependence

Up until this point, we have assumed a static model for the dependence structure, which is seldom a realistic representation. Natural follow-up questions are therefore related to the presence and type of time variation in the dependence. In this section, we consider these questions in more detail.

To investigate the time-varying aspect we consider Kendall's τ, which is a measure of concordance. In terms of a bivariate copula C, Kendall's τ can be expressed as (see e.g. McNeil et al. (2005))

$$\tau = 4 \int_0^1 \int_0^1 C(u_1, u_2) \mathrm{d}C(u_1, u_2) - 1 = 4\mathbb{E}[C(U_1, U_2)] - 1. \tag{8}$$

We compute Kendall's τ between $\hat{u}_t^{CSS}$ and $\hat{u}_t^{LWI}$ based on a rolling window of 60 days. Fig. 6 displays the results, and reveals a strong seasonal pattern in the dependence structure. According to Fig. 6, the dependence is strongest around winter and weakest around summer. A possible explanation for this behaviour relates to the power generation mix in Germany and the import/export conditions. During winter, the increased wind power production has a direct lowering effect on the daily electricity price due to the mechanism of day-ahead electricity price formation. As argued in Section 3.2, this lowers the CSS. During summer, the lower wind power production does not have the same direct effect on the daily electricity price. If that were the case, prices should increase. The high photovoltaic production during peak periods combined with the high likelihood of being able to import cheap nuclear power from France prevents however prices from increasing. Consequently, this weakens the dependence between the wind power production and the CSS during the summer months. This is in line with the observations of asymmetry made in Section 3.2: The lack of wind power production does not have the exact opposite effect on the CSS as the excess of it.

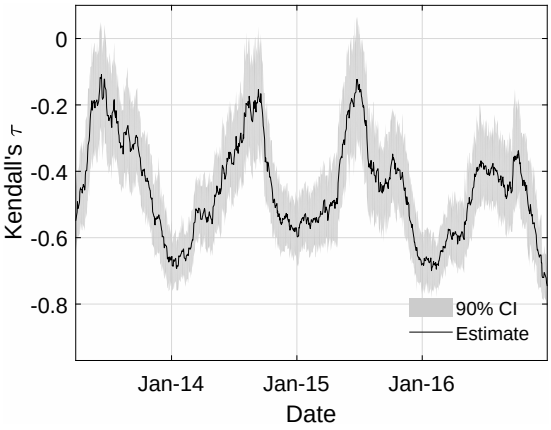


Fig. 6: Kendall's τ between $\hat{u}_t^{CSS}$ and $\hat{u}_t^{LWI}$ based on a 60-days rolling window. The confidence interval is based on 999 bootstraps. Note that the date corresponding to each estimate refers to the last day in the 60-days period.

In light of these findings we consider next extending the static copula mixture, such that the yearly seasonality in the dependence measured by Kendall's τ in Fig. 6 can be accounted for.

4.1 A seasonal copula mixture

Since the FRG copula is the best performing static copula cf. Table 5, we shall restrict our attention to this particular model in order to fix ideas. First, let us

state a general but simple result concerning Kendall's τ for copula mixtures, which is particularly useful in our modeling context. Again here, we omit the conditioning for notational convenience.

Proposition 4.1. *Let $U_i \sim$ Unif(0,1) for $i = 1, 2$, and let C^m denote the bivariate copula of (U_1, U_2). Further, suppose C^m is given as*

$$C^m(u_1, u_2) = \lambda C^a(u_1, u_2) + (1 - \lambda) C^b(u_1, u_2)$$

for two copulas C^a and C^b, and mixing parameter $0 < \lambda < 1$. Then Kendall's τ implied by C^m can be expressed as

$$\tau^m = \lambda \tau^a + (1 - \lambda) \tau^b, \tag{9}$$

where τ^a is Kendall's τ corresponding to copula C^a, and τ^b is Kendall's τ corresponding to copula C^b.

Proof. See Appendix A.2.

It follows from Prop. 4.1 that Kendall's τ for the copula mixture is simply a linear combination of the individual Kendall's τ's corresponding to the copulas comprised in the mixture. Thus, introducing time variation in τ^a and τ^b translates into time variation in τ^m. Further, explicit relations between Kendall's τ and the copula parameter are available for many of the copulas usually employed in the literature, see e.g. McNeil et al. (2005). Considering the FRG copula, we have that Kendall's τ for the Frank copula is given by

$$\tau^F = 1 - \frac{4}{\theta^F} \left(1 - \frac{1}{\theta^F} \int_0^{\theta^F} \frac{t}{e^t - 1} dt \right), \tag{10}$$

and Kendall's τ for the RGumbel copula is given by

$$\tau^{RG} = 1 - \frac{1}{\theta^{RG}}. \tag{11}$$

The superscripts F and RG are added in the above formulas to indicate their link to a particular copula. Notice that in both expressions, Kendall's τ is monotonically increasing as a function of the corresponding copula parameter. Therefore, specifying time variation for τ^F and/or τ^{RG} will also uniquely determine values of θ^F and θ^{RG}. If we instead were to introduce time variation directly in Kendall's τ for the FRG copula, we would not be able to identify θ^F and θ^{RG}.

Based on the discussion above and motivated by the pronounced yearly cycle in Fig. 6, we propose a simple extension to the static FRG model. Specifically, we introduce a yearly cycle in Kendall's τ corresponding to the RGumbel copula, i.e.

$$\tau_t^{RG} = a^{RG} + b^{RG} \sin(2\pi t / K) + c^{RG} \cos(2\pi t / K), \tag{12}$$

where a^{RG}, b^{RG}, and c^{RG} are constant coefficients, and $K = 258$ as was the case with the seasonal functions in Section 3.1. Regarding the Frank contribution in the FRG copula, we keep the corresponding Kendall's τ static. Consequently, the evolution equation for the overall Kendall's τ implied by the *seasonal* FRG copula is

$$\tau_t^{SFRG} = \lambda^{SFRG}\tau^F + (1 - \lambda^{SFRG})\tau_t^{RG}. \tag{13}$$

Given the relations in Eqs. (10) and (11) and the seasonal specification in Eq. (13), the model is estimated by maximizing the FRG copula loglikelihood. The estimation results are given in Table 6, revealing a clear improvement in AIC compared to the static FRG copula.

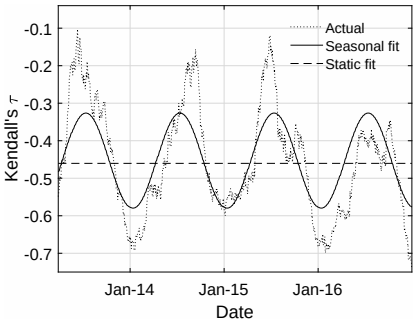
$\hat{\tau}^F$	$\hat{a}^{RG}$	$\hat{b}^{RG}$	$\hat{c}^{RG}$	λ^{SFRG}	$\log \mathcal{L}_c$	AIC
0.542	0.412	-0.111	0.170	0.314	309.711	-609.422
(0.068)	(0.033)	(0.033)	(0.035)	(0.092)		

Table 6: Maximum likelihood estimation results for the seasonal FRG copula described in Eqs. (12)–(13). The maximized value of the copula log-likelihood is denoted $\log \mathcal{L}_c$. Standard errors are reported in parenthesis and are computed following the simulation-based procedure described in detail in Patton (2013), where we note that the estimation error from the marginal models is taken into account.

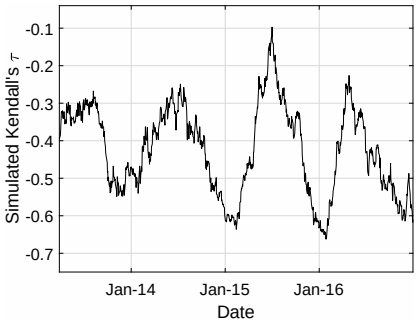
Since the FRG copula has three parameters, there are of course different ways of incorporating yearly seasonality in the model. Other specifications were therefore considered, such as introducing a yearly cycle in τ^F instead of τ^{RG}, and letting both τ^{RG} and τ^F vary over a yearly horizon. In terms of AIC, none of these alternatives outperformed the model in Eqs. (12)–(13). On a different note, by selecting the static FRG copula as the optimal model amongst static alternatives does not guarantee that the seasonal FRG copula will be preferred to time-varying extensions of other copula models. As a result, similar extensions as those proposed in Eqs. (12)–(13) were implemented for most of the copulas in Table 5 to ensure that the seasonal FRG is superior in terms of AIC.

To illustrate the fit of the proposed seasonal FRG model, we plot in Fig. 7(a) the empirical Kendall's τ together with $\hat{\tau}_t^{SFRG}$ implied by the seasonal FRG, using a 60-days moving window, as in Fig. 6. As a standard of comparison, we include the Kendall's τ implied by the static FRG. The results indicate that the dependence implied by the seasonal FRG follows the yearly cycle observed in the actual Kendall's τ rather well. Moreover, it appears from Fig. 7(a) that we would underestimate the strength of the dependence between the CSS and the LWI during autumn and winter with the static FRG. The reverse is observed during spring and summer, with the strength of the dependence being overestimated by the static FRG. To provide further sup-

port for the proposed seasonal FRG, we display in Fig. 7(b) a simulated path over a four-year horizon, which resembles the actual data nicely.

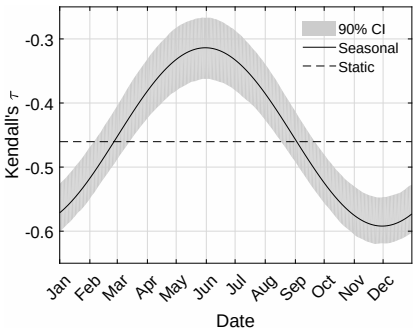


(a) Actual and implied Kendall's τ

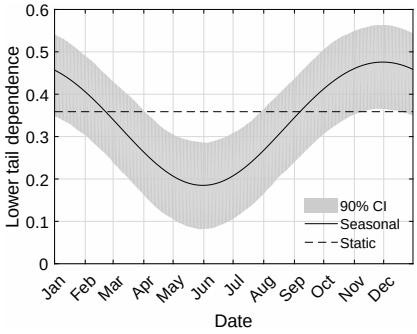


(b) One simulated path from the seasonal FRG

Fig. 7: (left) Actual Kendall's τ and Kendall's τ implied by the static and seasonal FRG copula, based on a 60-days rolling window. The date corresponding to each estimate refers to the last day in the 60-days period. (right) A simulated path of Kendall's τ from the seasonal FRG copula, aggregated using a 60-days rolling window.



(a) Implied Kendall's τ between CSS and LWI



(b) Implied lower tail dependence between CSS and LWI

Fig. 8: Kendall's τ and lower tail dependence implied by the static and seasonal FRG copulas throughout the year. A 90% confidence interval is provided for the seasonal parts.

To complement Fig. 7(a), a clearer picture of the yearly shape of Kendall's τ implied by the seasonal FRG is given in Fig. 8(a), where we illustrate the fit at each time point during a year (i.e., no averaging of Kendall's τ is performed). Equally interesting to consider is the lower tail dependence implied by the fitted seasonal FRG copula, which follows directly from applying Eq. (7). The results are plotted in Fig. 8(b), revealing that the lower tail dependence coefficient reaches its lowest value of approximately 0.2 dur-

ing summer and its maximum value of approximately 0.50 during winter. This entails that there is a rather large difference between the probability of observing the event of extreme high wind index / low CSS during winter compared to summer. In Fig. 8(b), we again provide as benchmark the corresponding static estimate.

5 Application results

Having established a model for the joint behavior of the CSS and the wind index, we consider next the quantification of the benefits that WPF can offer GFPPs. A GFPP acting in the day-ahead market can decide from day to day whether to run or not, and thereby take advantage of the daily variation in the CSS. By the construction of the day-ahead wholesale electricity market, a GFPP will not run in times of a negative CSS; its profit R^{CSS} for a period $t \in [T_1, T_2]$ can thus be represented as

$$R^{CSS} = \sum_{t=T_1}^{T_2} 12 \max(CSS_t, 0)s, \tag{14}$$

where s is the size of the GFPP measured in MW, and 12 is the number of peak load hours during a day. Recalling the payoff in Eq. (2), taking a position $\gamma \in \mathbb{Z}$ in WPF contracts yields the hedged profit of the GFPP, which we denote by R:

$$R = R^{CSS} + \gamma R^{WPF}. \tag{15}$$

We note that by excluding weekends and holidays from our analysis, these are not captured in R^{WPF}. We argue however that this does not alter the overall conclusions drawn below.

To facilitate hedging decisions, we perform Monte Carlo simulations from the proposed model. Specifically, the marginal models fitted in Section 3.1 and the seasonal FRG copula fitted in Section 4.1 are employed to produce simulations of the joint behavior of the CSS and the wind index, i.e., the pair (CSS_t, W_t). The "price" W_{t_0} affecting R^{WPF} in Eq. (15) is computed by averaging across all Monte Carlo simulations of W_t for the delivery $[T_1, T_2]$.[3] While

[3]When constructing the wind index data used in this paper, we considered the historical evolution of its two underlying data components, namely the wind power production and the installed capacity, as discussed in Section 2. This implicitly means that we have captured 1) the variations due to changes in wind speeds and 2) the variations caused by the increase in installed capacity and changes in the geographical distribution of wind turbines. While the latter aspect is important to capture in the modeling part of this paper, we argue that a different wind index series should be used in a pricing context. This is because today's WPF price is not affected by the historical evolution of the installed capacity and the changes in the geographical distribution of turbines, but by the *present* conditions. We argue that this issue does not affect the conclusions drawn in this paper, but can have serious implications in other contexts. For more details, we refer the interested reader to Benth and Pircalabu (2017).

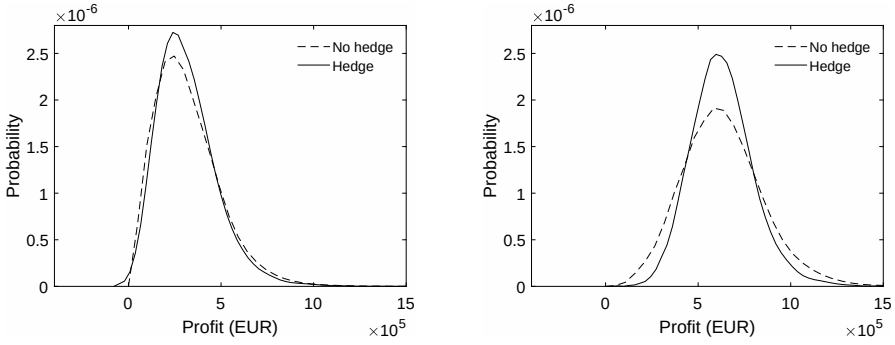
we recognize that this pricing approach is simplistic in that it assumes a zero market price of risk, it simplifies our hedging exercise somewhat, since the mean of the hedged profit R will not be affected by varying the quantity γ. Consequently, instead of the classical mean-variance objective, we can restrict ourselves to the variance minimization criterion in order to determine optimal positions in WPF contracts. Hence, we consider the following objective:

$$\min_{\gamma \in \mathbb{Z}} \mathbb{V}\text{ar}[R]. \tag{16}$$

5.1 Effectiveness of wind power futures

To illustrate the results obtained by applying the hedging approach described above, we fix $s = 200$ MW, $t_0 = 30$ December 2016 (the last date in our sample), and perform 20,000 Monte Carlo simulations of the pair (CSS_t, W_t) one year ahead. The resulting simulated paths are split into monthly periods, and WPF prices corresponding to monthly deliveries are computed as explained earlier. Then, monthly quantities for R^{CSS} and R^{WPF} are constructed for each simulated path, and the minimization in Eq. (8) is applied to each month in turn. The subdivision to monthly profits is motivated by the seasonal pattern observed in the dependence structure cf. Fig. 7(a), and allows us to investigate the effect of the yearly seasonality on hedging-related aspects.

In Fig. 9, we illustrate the simulated unhedged profit distribution R^{CSS} and the hedged profit distribution R obtained by solving Eq. (8) for the months July and October. We observe a compression of the profit distribution in both cases when applying the hedge, which entails that WPF have variance reducing effects. In fact, this finding applies to all 12 months, as will be illustrated shortly.



(a) July (b) October

Fig. 9: Examples of profit distributions before and after hedging with WPF, based on Monte Carlo simulations from the proposed seasonal FRG copula model.

Perhaps unsettling is the fact that losses can occur when considering the hedged profit distribution in Fig. 9(a), whereas the unhedged profits cannot attain negative values by construction (see Eq. (14)). Nevertheless, we find that the probability of a loss when hedging with WPF is approximately 0.4% on average. The downside of performing the hedge is therefore quite small. In the pursuit to impair this concern even further, recall that the price of WPF is computed under the assumption of a zero market price of risk. In reality, the studies of Gersema and Wozabal (2017) and Benth and Pircalabu (2017) find evidence of a negative market price of risk in the German market for WPF, implying that a GFPP buys WPF at a discounted price compared to the one computed here. Accounting for this would shift the hedged profit distributions to the right, potentially excluding losses altogether.

Next, we consider in more detail the reduction in the variance of profit distributions attained by performing the hedge. The results are stated in Fig. 10 for all months of the year, and reveal considerable reductions; even for May and June, where we observe the lowest values, the variance reductions are above 10 %. Further, notice the connection between the yearly pattern of the reductions in Fig. 10 and the implied Kendall's τ in Fig. 8(a): Not surprisingly, the stronger the dependence, the higher the variance reduction.

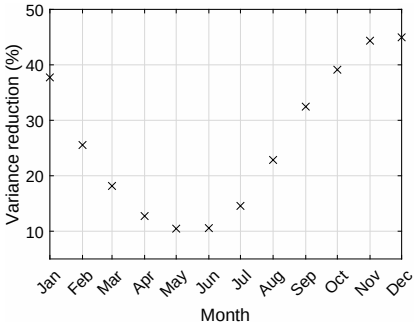


Fig. 10: Variance reduction achieved by hedging with WPF, for each month of 2017. The results are based on Monte Carlo simulations from the proposed seasonal FRG copula model.

Also relevant to consider in this context is the impact on hedging effectiveness from changing the GFPP efficiency. To assess this, we allow the efficiency to vary from 43.13% to 55.13% with a step size of 2%, and let the emission factor vary according to

$$e = 0.184h,$$

which is based on ICIS (2016). Fig. 11 illustrates the variance reductions obtained with the different efficiencies, across all months of the year. It appears that increasing the efficiency (i.e., lowering the heat rate) leads to an increase in the variance reductions for all months. The effect seems to be more pro-

nounced during autumn and winter compared to spring and summer. From Eq. (1), it was already apparent that increasing the efficiency of a GFPP produces a higher CSS and hence increases profitability. The findings presented in Fig. 11 incentivize such action even further: Aside from the higher CSS, an increased hedging effectiveness of WPF can be achieved.

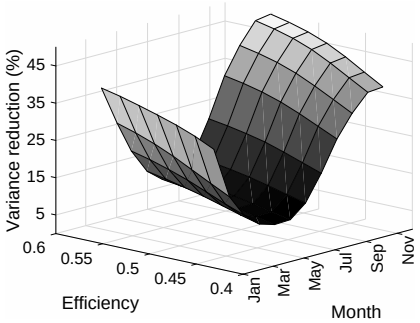


Fig. 11: Variance reduction implied by hedging with WPF, for GFPP efficiencies spanning from 43.13% to 55.13% with a step size of 2%.

5.2 Comparison with alternative models

So far in the hedging application, we have focused on the results obtained with the preferred copula model, that is, the seasonal FRG. In this section, we wish to highlight the benefits of this copula compared to other less optimal alternatives. We consider the following natural progression in comparisons:

1. Frank copula versus FRG copula: In this comparison, we focus on the effect of asymmetry and tail dependence on the benefits of WPF. These features are captured by the FRG copula, as discussed in Section 3.2, but not by the Frank copula, which imposes symmetry and no tail dependence.

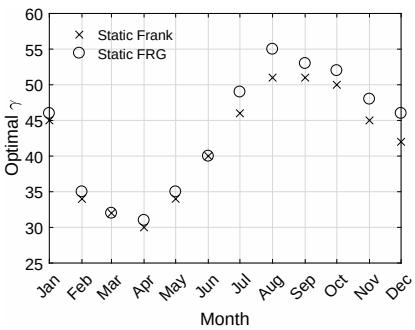
2. FRG copula versus seasonal FRG copula: Here, we concentrate on the effect of seasonal dependence on the hedging benefits.

To perform comparisons, we keep the marginal models proposed in Section 3.1 fixed, and repeat the simulations performed in Section 5.1 with the Frank, FRG, and seasonal FRG copulas, instead of only the seasonal FRG. We note that the same random seed was used to produce Monte Carlo simulations from the three models. Then, we compute optimal hedge quantities γ and associated variance reductions with each model, on a monthly basis.

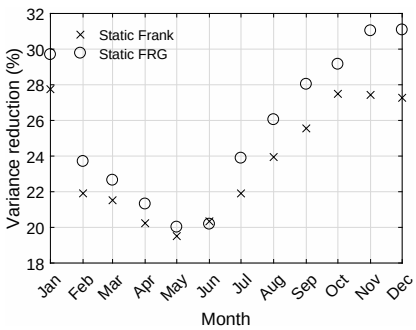
The effects of asymmetry and tail dependence in the copula

Recalling Figs. 5(a) and 5(c), it is apparent that by employing the FRG copula instead of the Frank copula, we introduce a slight asymmetry and assign more probability to the extreme events where high wind and low CSS happen simultaneously. The resulting effects on hedging are depicted in Fig. 12, where we present the optimal hedge quantities and variance reductions produced by the two copulas. Regarding the former, we notice that the Frank copula generally suggests less WPF in the hedging portfolio. Further, the optimal hedge quantities vary across the year, which is a consequence of the seasonality captured in the marginal models.

Turning to the variance reductions, which are depicted in Fig. 12(b), we observe that the values implied by the Frank copula are generally lower compared to those implied by the FRG copula. This finding is expected, since GFPPs seek to cover their exposure to high wind / low CSS scenarios. By assigning more probability mass to precisely these events happening simultaneously, which is done by shifting from the Frank to the FRG copula, we increase the benefits of WPF. At the same time, due to the asymmetric behavior of the FRG copula, we are not increasing the probability of observing the reverse combination of low wind / high CSS, and thus not counteracting the increased benefits of WPF. Briefly put, by believing in a dependence structure described by the Frank copula compared to the FRG copula, we would underestimate the risk-reducing power of WPF.



(a) Optimal hedge quantity for each month of 2017 implied by the Frank and FRG copulas



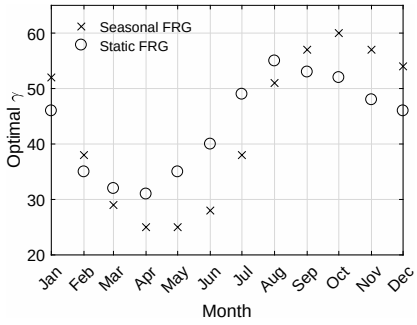
(b) Reduction in the variance of the profit distribution for each month of 2017, when applying the hedges implied by the Frank and FRG copulas

Fig. 12: Comparison of hedging results implied by the Frank and FRG copulas.

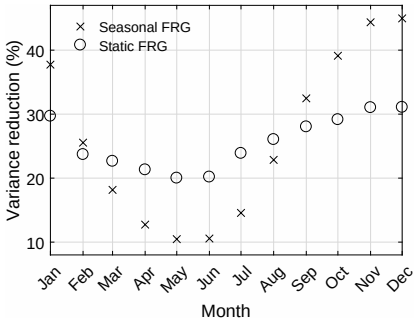
The effects of seasonal time variation in the copula

Proceeding to the comparison of the FRG copula with its seasonal version, we present in Fig. 13 results that are similar to those in Fig. 12. Regarding the optimal hedge quantities in Fig. 13(a), the FRG copula yields higher values than the seasonal FRG copula during spring and summer, while the situation reverses during autumn and winter. This alternating behavior is connected to that of the differences in Kendall's τ implied by the FRG and the seasonal FRG, cf. Fig. 8(a). That is, the hedge quantities decrease (increase) with a decrease (increase) in absolute values of Kendall's τ.

Considering Fig. 13(b), the results reveal a fairly constant level in the variance reduction produced by the FRG copula, compared to the variance reduction levels implied by the seasonal model. Hence, believing in static dependence can lead to very misleading conclusions when managing risks. Again in this context, we mention the link between the difference in percentage reductions and the difference between the Kendall's τ implied by the two models (cf. Fig. 8(a)): The difference in reductions is largest in May/June and November/December, reflecting the fact that the dependence implied by the seasonal FRG model is weakest during May/June and strongest during November/December.



(a) Optimal hedge quantity for each month of 2017 implied by the FRG and seasonal FRG copulas



(b) Reduction in the variance of the profit distribution for each month of 2017, when applying the hedges implied by the FRG and seasonal FRG copulas

Fig. 13: Comparison of hedging results implied by the FRG and seasonal FRG copulas.

Having found clear evidence of seasonal dependence between the CSS and the wind index, we conclude this section by briefly addressing the error we would get by applying a hedge based on the static FRG model in a seasonal time-varying reality. To perform this analysis, we asses the optimal hedge quantities implied by the static FRG copula in a seasonal setting by using the simulated CSS and wind index from the seasonal model. The

variance reductions obtained with this approach are then compared with the reductions implied by the seasonal model shown in Fig. 13(b). The results, presented in Fig. 14, reveal very small errors. The smallest and largest errors occur in February and June, respectively, which is connected to the findings presented in Fig. 13(a); the absolute difference in the optimal number of WPF in the static and seasonal case generates the pattern seen in Fig. 14.

With the small errors in mind, the real error one commits by believing in static dependence, is the belief in a wrong resulting variance reduction. Thus, while the static model creates a misleading picture in a risk management context, our results suggest that it could be employed to determine optimal hedging quantities.

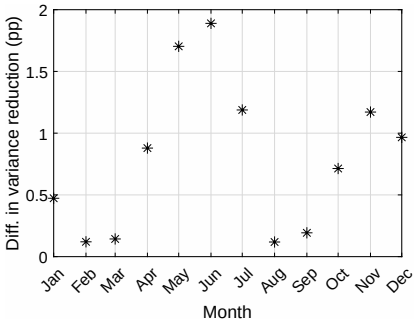


Fig. 14: The effects of using the "wrong" hedge quantity: Difference in variance reduction of the profit distribution when using hedge quantities obtained from 1) the FRG, and 2) the seasonal FRG, both evaluated using simulations from the seasonal FRG copula model.

5.3 Discussion of the proposed hedging strategy

To conclude this section, we turn briefly to the standard hedging principle often employed to hedge the day-ahead CSS. Usually, conventional generators remove their exposure to day-ahead price risk either completely or partially by entering a short position in standard power forwards, and a long position in fuel forwards and carbon credit forwards. In this paper, we have considered a different hedging approach with the purpose of determining the potential of the newly introduced wind power futures, but we stress that our strategy is not incompatible with the industry standard. In fact, our hedging portfolio consisting of wind power futures could be extended to include the additional forwards mentioned above. However, this would require us to switch from our bivariate modeling problem to a multivariate one, since the joint behavior of the wind index, the day-ahead CSS and the different forward clean spark spreads should be considered. While this is outside the scope of the present paper, it is nevertheless an interesting perspective that has not been studied yet, and could possibly be approached with vine copulas. In

the context of the effectiveness of the standard hedging principle that conventional generators usually employ, we mention the study of Charalampous and Madlener (2016).

6 Conclusion

In this paper, we propose a joint model for the day-ahead clean spark spread and the daily wind index that can facilitate hedging decisions for a gas-fired power plant. The modeling procedure is based on two steps: First, the marginal behavior of the variables is considered, where we apply seasonal functions and ARMA–GARCH filters to remove the seasonality and the serial dependence in the conditional mean and variance. While the usual Gaussian assumption for the innovation process works in the case of the daily wind index, the normal-inverse Gaussian distribution provides a better fit for the day-ahead clean spark spread. Second, the standardized residuals from the ARMA–GARCH models are connected through copulas. The data reveals a dependence structure that is slightly asymmetric, and also varying according to an annual cycle. To capture these empirical findings, we propose a seasonal copula mixture, where the mixing components are the rotated Gumbel and the Frank copulas.

Based on Monte Carlo simulations from the proposed model, we show that wind power futures have considerable risk-reducing benefits in the context of a gas-fired power plant operating in the day-ahead market. Further, their hedging effectiveness increases as a function of the efficiency of the gas-fired power plant. To highlight the importance of capturing asymmetry, tail dependence, and seasonality in the dependence structure, we perform comparison studies where the optimal model is compared to less optimal alternatives. Accounting for asymmetry and tail dependence (as opposed to imposing symmetry and zero tail dependence) leads to an increase in the effectiveness of wind power futures. Moreover, we find that the conclusions drawn with a static dependence model deviate to a large extent from those obtained with a seasonal dependence model. With static dependence, the variance reductions of the profit distributions attained by the hedge vary between 20% and 31%; in the seasonal case the corresponding reductions vary between 10% and 45%.

Although we have concentrated on the German market and the case of gas-fired power plants, the results are relevant for other markets, and are also transferable to other conventional electricity producers. Since the amount of electricity generated by wind turbines is expected to grow globally, the dependence between the day-ahead clean spark spread and the daily wind index in other market places will most likely be strengthened in the future. Hence, it is reasonable to assume that more weather-based instruments sim-

ilar to the German wind power futures will be introduced, enabling similar analyses to be performed on other than the German market.

Acknowledgments

The authors would like to thank the Quantitative Analytics team at Neas Energy and the participants attending the Second Conference on the Mathematics of Energy Markets 2017 at the Wolfgang Pauli Institute for providing valuable comments and suggestions. Three anonymous referees are also thanked for their constructive criticism and suggestions, which improved the presentation of this paper.

Funding

Troels Sønderby Christensen is supported by the Innovation Fund Denmark under Grant 5189-00117B. Anca Pircalabu is supported by the Innovation Fund Denmark under Grant 4135-00082B.

A Results for copula mixtures

A.1 Proof of Proposition 3.1

For positive dependent variables, the upper tail dependence for copula C can be written as (see e.g. (McNeil et al., 2005, p. 209))

$$\Lambda^u = \lim_{u\uparrow 1} \frac{\hat{C}(1-u,1-u)}{1-u} = \lim_{u\uparrow 1} \frac{1-2u+C(u,u)}{1-u}, \tag{17}$$

where $\hat{C}(u,v) := P(U > u, V > v) = C(1-u,1-v) + u + v - 1$ is the survival copula. Applying Eq. (17) to the copula mixture, we get that the upper tail dependence is

$$\begin{aligned}
\Lambda^u &= \lim_{u\uparrow 1} \left[\frac{1-2u+\lambda C^a(u,u) + (1-\lambda)C^b(u,u)}{1-u} \right] \\
&= \lim_{u\uparrow 1} \left[\frac{1-2u}{1-u} + \lambda \frac{1-2u+C^a(u,u)}{1-u} + (1-\lambda)\frac{1-2u+C^b(u,u)}{1-u} - \lambda\frac{1-2u}{1-u} - (1-\lambda)\frac{1-2u}{1-u} \right] \\
&= \lambda\Lambda^{u,a} + (1-\lambda)\Lambda^{u,b}.
\end{aligned}$$

Similarly, the lower tail dependence can be written in terms of copula C as (again for positive dependent variables)

$$\Lambda^l = \lim_{u\downarrow 0} \frac{C(u,u)}{u},$$

resulting in the following lower tail dependence for the copula mixture:

$$\begin{aligned}\Lambda^l &= \lim_{u \downarrow 0} \left[\frac{\lambda C^a(u,u) + (1-\lambda) C^b(u,u)}{u} \right] \\ &= \lambda \Lambda^{l,a} + (1-\lambda)\Lambda^{l,b}.\end{aligned}$$

A.2 Proof of Proposition 4.1

From Eq. (8) we have that Kendall's τ implied by copula C^m is

$$\begin{aligned}\tau^m &= 4\mathbb{E}[C^m(U_1,U_2)] - 1 \\ &= 4\left(\lambda \mathbb{E}[C^a(U_1,U_2)] + (1-\lambda)\mathbb{E}[C^b(U_1,U_2)]\right) - (1+\lambda-\lambda) \\ &= \lambda\left(4\mathbb{E}[C^a(U_1,U_2)] - 1\right) + (1-\lambda)\left(4\mathbb{E}[C^b(U_1,U_2)] - 1\right) \\ &= \lambda\tau^a + (1-\lambda)\tau^b.\end{aligned}$$

References

M. D. Aepli, R. Füss, T. E. S. Henriksen, and F. Paraschiv. Modeling the multivariate dynamic dependence structure of commodity futures portfolios. *Journal of Commodity Markets*, 6:66 – 87, 2017.

K. Avdulaj and J. Barunikl. Are benefits from oil–stocks diversification gone? New evidence from a dynamic copula and high frequency data. *Energy Economics*, 51:31 – 44, 2015.

O. E. Barndorff-Nielsen. Normal inverse Gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of Statistics*, 24(1):1–13, 1997a.

O. E. Barndorff-Nielsen. Processes of normal inverse Gaussian type. *Finance and Stochastics*, 2(1):41–68, 1997b.

F. E. Benth and J. Šaltytė Benth. The normal inverse Gaussian distribution and spot price modelling in energy markets. *International Journal of Theoretical and Applied Finance*, 07(02):177, 2004.

F. E. Benth and J. Šaltytė Benth. Weather derivatives and stochastic modelling of temperature. *International Journal of Stochastic Analysis*, Vol. 2011:Article ID 576791, 21 pages, 2011.

F. E. Benth and P. C. Kettler. Dynamic copula models for the spark spread. *Quantitative Finance*, 11(3):407 – 421, 2011.

F. E. Benth and A. Pircalabu. A non-Gaussian Ornstein-Uhlenbeck model for pricing wind power futures. *Working paper*, 2017.

T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327, 1986.

R. Börger, A. Cartea, R. Kiesel, and G. Schindlmayr. Cross-commodity analysis and applications to risk management. *Journal of Futures Markets*, 29(3): 197 – 217, 2009.

G. Charalampous and R. Madlener. Risk management and portfolio optimization for gas- and goal-fired power plants in Germany: A multivariate GARCH approach. *Journal of Energy Markets*, 9(2):69 – 94, 2016.

D. Creal, S. J. Koopman, and A. Lucas. Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, 28(5):777 – 795, 2013.

A. Dias and P. Embrechts. Testing for structural changes in exchange rates' dependence beyond linear correlation. *The European Journal of Finance*, 15 (7–8):619 – 637, 2009.

C. Elberg and S. Hagspiel. Spatial dependencies of wind power and interrelations with spot price dynamics. *European Journal of Operational Research*, 241(1):260 – 272, 2015.

R. F. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4):987 – 1007, 1982.

European Commission. The EU emissions trading system. 2017.

G. Gersema and D. Wozabal. An equilibrium pricing model for wind power futures. *Energy Economics*, 65:64 – 74, 2017.

O. Grothe and J. Schnieders. Spatial dependence in wind and optimal wind power allocation: A copula-based analysis. *Energy Policy*, 39(9):4742 – 4754, 2011.

N. Haldrup and M. Ø. Nielsen. A regime switching long memory model for electricity prices. *Journal of Econometrics*, 135(1–2):349 – 376, 2006.

W. K. Härdle and B. López Cabrera. Implied market price of weather risk. *Applied Mathematical Finance*, 19(1):59 – 95, 2012.

ICIS. European daily electricity markets methodology. 2016.

M. B. Jensen and A. Lunde. The NIG-S&ARCH model: A fat-tailed, stochastic, and autoregressive conditional heteroskedastic volatility model. *The Econometrics Journal*, 4(2):319 – 342, 2001.

H. Joe. Asymptotic efficiency of the two-stage estimation method for copula-based models. *Journal of Multivariate Analysis*, 94(2):401 – 419, 2005.

A. J. McNeil, R. Frey, and P. Embrechts. *Quantitative risk management*. Princeton Series in Finance, 2005.

R. B. Nelsen. *An Introduction to Copulas*. Springer, 2006.

A. J. Patton. Modelling asymmetric exchange rate dependence. *International Economic Review*, 47(2):527 – 556, 2006.

A. J. Patton. Copula methods for forecasting multivariate time series. In Graham Elliott and Allan Timmermann, editors, *Handbook of Economic Forecasting*, volume 2B, pages 899 – 960. Elsevier B.V., 2013.

A. Pircalabu and J. Jung. A mixed C-vine copula model for hedging price and volumetric risk in wind power trading. *Quantitative Finance*, 17(10): 1583 – 1600, 2017.

J. C. Rodriguez. Measuring financial contagion: A copula approach. *Journal of Empirical Finance*, 14(3):401 – 423, 2007.

T. H. Rydberg. The normal inverse Gaussian Lévy process: Simulation and approximation. *Communications in Statistics. Stochastic Models*, 13(4):887 – 910, 1997.

R. H. Shumway and D. S. Stoffer. *Time Series Analysis and Its Applications*. Springer Texts in Statistics, 2011.

A. Sklar. Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de L'Université de Paris*, 8:229 – 231, 1959.

Paper VI

On the spatial hedging effectiveness of German wind power futures for wind power generators

List of authors: Troels Sønderby Christensen[*,†] and Anca Pircalabu[*,†]
[*] Department of Mathematical Sciences, Aalborg University
[†] Quantitative Analytics, Neas Energy

Submitted to
Journal of Energy Markets

ABSTRACT

The wind power futures recently introduced on the German market fill the gap of a standardized product that addresses directly the volume risk in wind power trading. While the German wind power futures entail risk-reducing benefits for wind power generators generally speaking, it remains unclear the extent of these benefits across wind farms with different geographical locations. In this paper, we consider the wind utilization at 31 different locations in Germany, and for each site, we propose a copula model for the joint behavior of the site-specific wind index and the overall German wind index. Our results indicate that static mixture copulas are preferred to the stand-alone copula models usually employed in the economic literature. Further, we find evidence of asymmetric dependence and upper tail dependence. To quantify the benefits of wind power futures at each wind site, we perform a minimum variance hedge, and find that variance reductions can differ greatly depending on the geographical location. Further, different comparison studies reveal that the presence of 1) a negative risk premium in the wind power futures market and 2) upper tail dependence weaken the benefits of wind power futures for wind power generators.

1 Introduction

Wind power generators worldwide have historically been given subsidies in order to incentivize the development of renewable energy sources. The non-programmable nature makes investment in wind power generation unpredictable, and although subsidies simplify investment decisions, the stochastic behavior of wind demands further risk reducing opportunities. As a result, the so-called German wind power futures have recently been introduced on the European Energy Exchange and Nasdaq. These instruments are written on a wind power production index that reflects the average German utilization, and a long (short) position gives a profit in high-wind (low-wind) scenarios. It follows naturally that wind power generators constitute the seller group, as they seek to cover their exposure to the low-wind scenarios that affect their cash-flows negatively.

In this paper, we study the hedging benefits of wind power futures for wind turbines or wind farms with different geographical locations in Germany. To facilitate optimal hedging decisions, we employ copula models for the joint behavior of the site-specific and the German wind power production indexes. Specifically, we base our empirical analysis on data from 31 different sites in Germany, and quantify the hedging benefits of wind power futures for each site, showcasing how these vary across locations. Further, we comment on how the negative risk premium in the wind power futures market affects the wind power generator's hedged profit distribution. For

the data we consider, we find significant evidence of asymmetric dependence and upper tail dependence between the site-specific and the German wind utilization. To highlight the importance of capturing such effects, we include a comparison study where the hedging exercise is performed with the optimal copula model and a less optimal alternative that imposes symmetry and no tail dependence.

Owing to the recent introduction of wind power futures to the market, the related literature is yet scarce. The first studies to consider the German wind power futures are Gersema and Wozabal (2017) and Benth and Pircalabu (2017). Both papers concentrate on the pricing rather than the hedging aspect, and agree on the presence of a negative risk premium which is explained by the fact that the wind power producer's profit is more correlated to the German wind utilization than the conventional generator's profit. Considering the hedging aspect is Christensen et al. (2017), where the benefits of wind power futures are studied in the context of conventional generators. Due to the prioritization of the cheapest energy sources in the day-ahead market, conventional generators are affected by the share of wind power in the system. In high-wind scenarios, conventional generators will produce less, and they can minimize this exposure by taking a long position in wind power futures. While the present paper is similar to the study in Christensen et al. (2017) in that it investigates the hedging power of wind power futures, it differentiates itself by considering the seller rather than the buyer side. Furthermore, we consider multiple potential sellers in order to emphasize the spatial aspect, which in contrast to conventional generators is very important to consider here.

Turning to copulas, their application in the energy markets literature has grown tremendously over the past years, see e.g. Benth and Kettler (2011), Grothe and Schnieders (2011), Avdulaj and Barunikl (2015), Elberg and Hagspiel (2015), Pircalabu and Jung (2017), Aepli et al. (2017) and Liu et al. (2017) to name a few. Closest to the present paper are the studies of Grothe and Schnieders (2011) and Elberg and Hagspiel (2015), who also employ copulas to model the German wind power production. The study of Grothe and Schnieders (2011) concentrates on the optimal geographical allocation of wind turbines, and Elberg and Hagspiel (2015) quantify the value of wind turbines at different locations in Germany, which requires modeling the dependence between the aggregated German wind power generation and the wind power generation from each wind turbine. While our modeling approach is very similar to that in Elberg and Hagspiel (2015), we introduce further flexibility by considering copula mixtures.

The remaining of this paper is structured as follows: Section 2 presents the data and introduces the wind power futures. In Section 3 we briefly describe the modeling approach and present the estimation results. In Section 4 we examine the spatial hedging effectiveness of wind power futures and provide

a comparison study. Section 5 concludes.

2 Data presentation and the wind power futures

The empirical study performed in this paper relies on wind index data from 31 different sites in Germany, and on the overall German wind index, which acts as the underlying of wind power futures. In the following, we elaborate on the data and introduce the German wind power futures.

2.1 German and site-specific wind index data

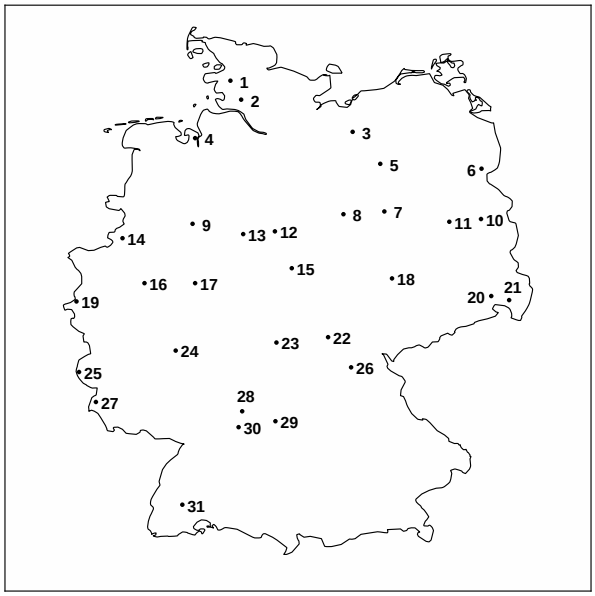
The first data component in our analysis consists of German wind index data provided by Narex. The index represents the aggregated utilization of wind power plants in Germany on a daily basis, and covers the period from 1 January 2012 to 31 December 2015, that is we have a total of 1461 observations. We remark that the index is computed based on wind speed data and a *constant* level of installed capacity in Germany, as to exclude effects that originate from the historical variations in the installed capacity. Specifically, the level is set to that of December 2015, and the German wind index data reflects therefore the distribution of wind generation capacity at that time.

The second data component consists of wind index data at 31 specific wind sites, again for the period 1 January 2012 to 31 December 2015. For each site, we have access to the actual daily wind power generation data and the installed capacity data. To construct the site-specific wind index, we consider the following: Let $P_{t,i}$ denote the day t wind power generation measured in MWh for wind site i, and let $C_{t,i}$ denote the installed capacity for wind site i. The wind index $W_{t,i} \in [0,1]$ is then obtained as

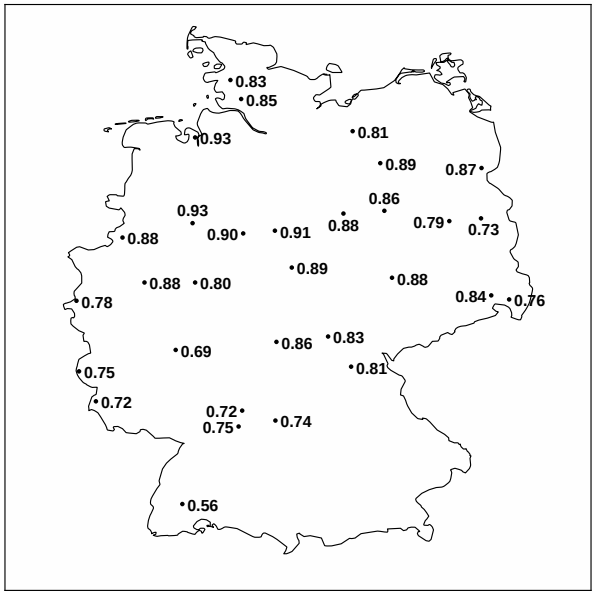
$$W_{t,i} = \frac{P_{t,i}}{h \cdot C_{t,i}}, \tag{1}$$

where h is the number of hours in a given day. We note that the number of turbines included in each wind site, their type, condition, height, and the surrounding terrain vary across the sites. While we acknowledge that circumstances other than the geographical location can have an impact on the site-specific wind index, it is not possible to separate such effects based on the data we have available. Therefore, whenever we refer to differences between the wind sites which are caused by their geographical locations, we implicitly mean other circumstances as well.

To illustrate how the 31 wind sites are spread across Germany, we plot in Fig. 1(a) their approximate geographical location, with corresponding site ID number. Moreover, we report in Fig. 1(b) the linear correlation between the



(a) Location and site ID

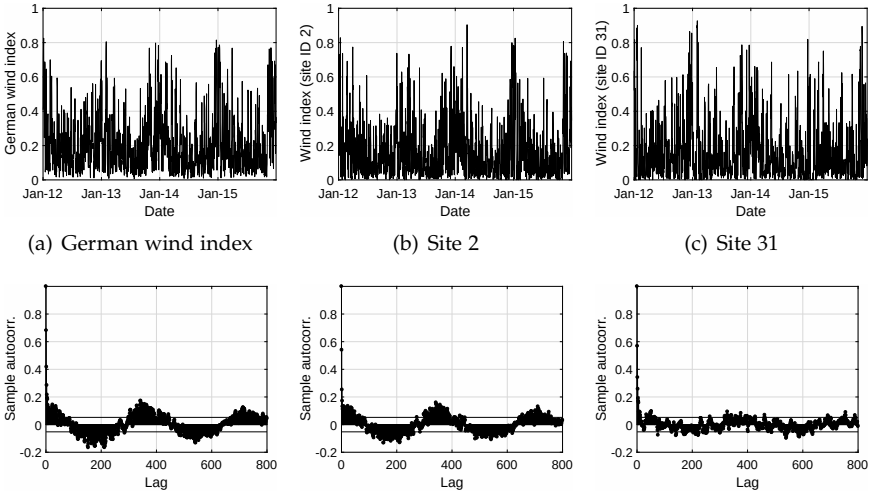


(b) Linear correlation

Fig. 1: Location of wind sites with site ID, and linear correlations between German wind index and each site wind index.

wind index of each site and the German wind index – not surprisingly, we generally find a very strong positive relation.

To provide the reader with more sense of the data, we present in the upper row block of Fig. 2 time series plots of the German wind index and the wind indexes at sites 2 and 31. In the lower row block of Fig. 2, we plot the corresponding sample autocorrelation functions. The plots reveal clear yearly cycles for the German wind index and the wind index at site 2, while the yearly seasonality at site 31 is much less pronounced.



(a) German wind index (b) Site 2 (c) Site 31

(d) Autocorr. German wind index (e) Autocorr. site 2 (f) Autocorr. site 31

Fig. 2: Upper row block: Time series plots of the German wind index and the wind index at sites 2 and 31. Lower row block: Sample autocorrelation of the German wind index and the wind index at sites 2 and 31.

2.2 Wind power futures

The German wind index displayed in Fig. 2(a) acts as the underlying for German wind power futures (WPF) traded on the European Energy Exchange and Nasdaq. To clarify the WPF payoff, let us consider a contract with delivery period $[T_1, T_2]$, where $T_1 < T_2$. Further, let us denote by W_t^{DE} the realized wind index in Germany at day t. Then, a long position in a WPF contract yields the payoff

$$R^{WPF} = H\left(\frac{1}{T_2 - T_1 + 1}\sum_{t=T_1}^{T_2} W_t^{DE} - W_{t_0}\right)X, \tag{2}$$

where H is the number of wind production hours (wph) in the delivery period, W_{t_0} is the "market price" set at day $t_0 < T_1$, and X is the tick size set to 100 EUR at EEX. The price is quoted in EUR per wph, and a price of e.g. 30 EUR/wph translates to $W_{t_0} = 0.30$.

It follows from Eq. (2) that a long position in a WPF contract generates a profit in high wind scenarios, whereas a short position generates a profit in low wind scenarios. For a wind power generator, a short position in WPF contracts is therefore an interesting opportunity to mitigate the volumetric risk associated with wind power generation.

3 Modeling procedure

To model the joint behavior of the German wind index and a site-specific wind index, we employ ARMA–GARCH copula models; with 31 different wind sites, we end up with 31 bivariate models. The choice of modeling approach is motivated by the well-known flexibility of these models, and also by the copula being an excellent tool for modeling nonlinear dependence. Since ARMA–GARCH copula models are common in the financial and econometric literature, we shall not provide a comprehensive introduction here, and refer the interested reader to e.g. Patton (2006, 2013), Dias and Embrechts (2009), and Pircalabu et al. (2017).

In the following subsections, we consider 1) univariate ARMA–GARCH type models for the marginal behavior of the German wind index and all site-specific wind indexes, and 2) bivariate copulas for the dependence between the German wind index and each site-specific wind index. For readability, we only report the results for four chosen sites, namely site 2, 19, 20, and 31 cf. Fig. 1(a). The four sites are selected as to emphasize the spatial effects.

3.1 Marginal models

As already illustrated in Fig. 2, the wind index data is bounded between zero and one. Since ARMA-GARCH models are not suited for modeling this type of bounded data, we first apply the logit transform to the wind indexes.[1] Then, to capture the yearly seasonality revealed in Fig. 2, we employ a seasonal function of the form,

$$f_t = a_1 + c_1 \sin(2\pi t/365) + c_2 \cos(2\pi t/365), \tag{3}$$

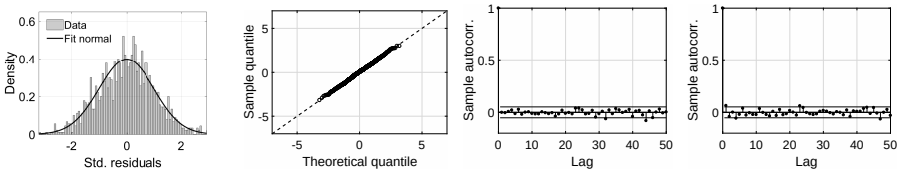
which is fitted to each logit-transformed wind index.

After removing the yearly seasonality from the data, we filter each marginal series through ARMA–GARCH models. To determine the optimal order of

[1] The logit transform is given by $\text{logit}(x) = \log(x) - \log(1-x)$.

the models, we employ the Bayesian Information Criterion (BIC), where we consider ARMA(p,q)–GARCH(h,k) models of order $p, q = 0, .., 5$ and $(h, k) = \{(0,0), (0,1), (1,1)\}$. Table 1 presents the optimal models and corresponding parameter estimates with standard errors in parenthesis, revealing that the preferred specifications vary across the chosen sites.

Considering the goodness-of-fit of the proposed models in Table 1, let us denote by η^{DE} the standardized residuals corresponding to the German wind index, and η^{i} the standardized residuals corresponding to the wind index at site i. Diagnostic plots are displayed in Fig. 3 for the German index, and in Fig. 4 for four selected sites. The autocorrelation functions for the standardized residuals and the standardized residuals squared show almost no correlation, hence indicating appropriate model selection. Regarding the distributional assumptions, the normal distribution provides a nice fit to the standardized residuals corresponding to the German wind index, as it appears from Figs. 3(a) and 3(b). For the site-specific wind indexes however, the normal distribution is less appropriate. As a result, we relax the normality assumption and find that a skew t distribution provides a much better fit, as shown in Fig. 4. Parameter estimates for the fitted skew t distributions together with standard errors are reported in the last row block of Table 1. Although results concerning four selected sites are displayed in this section, we stress that very similar results are obtained for the remaining series.



(a) Histogram (b) QQ plot (c) Acf of $\hat{\eta}_t^{DE}$ (d) Acf of squared $\hat{\eta}_t^{DE}$

Fig. 3: Model check for the German wind index.

3.2 Copula models

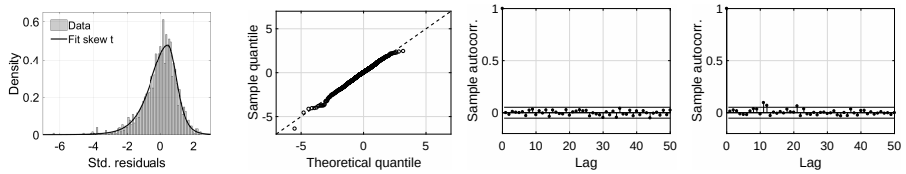
Next, we turn to dependence modeling with copulas, which are esentially distributions with uniform margins. Let us consider the random vector $\boldsymbol{Y}_t = (Y_{t,1}, Y_{t,2})$, with joint conditional distribution $F(\cdot|\mathcal{F}_{t-1})$. Then, according to Sklar's theorem [Sklar (1959)] for conditional distributions, we can decompose $F(\cdot|\mathcal{F}_{t-1})$ as

$$F(y_{t,1}, y_{t,2}|\mathcal{F}_{t-1}) = C(F_1(y_{t,1}|\mathcal{F}_{t-1}), F_2(y_{t,2}|\mathcal{F}_{t-1})|\mathcal{F}_{t-1}), \tag{4}$$

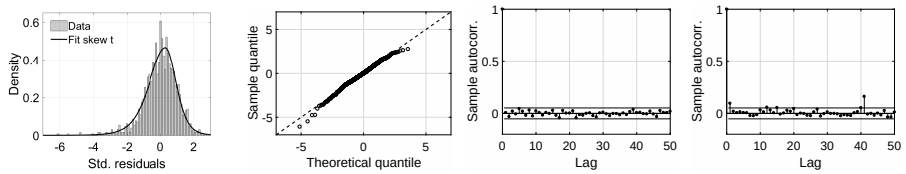
where C denotes the copula function, $F_1(\cdot|\mathcal{F}_{t-1})$ and $F_2(\cdot|\mathcal{F}_{t-1})$ are the marginal distribution functions of $Y_{t,1}$ and $Y_{t,2}$, respectively, and $\mathcal{F}_{t-1}$ is the filtration.

	Logit W^{DE}	**Logit $W^{\text{Site 2}}$**	**Logit $W^{\text{Site 19}}$**	**Logit $W^{\text{Site 20}}$**	**Logit $W^{\text{Site 31}}$**
Model	ARMA(1,1)	ARMA(1,1)–ARCH(1)	AR(1)	AR(2)–ARCH(1)	AR(3)–GARCH(1,1)
			Conditional mean		
AR1 $\hat{\phi}_1$	0.444 (0.039)	0.349 (0.042)	0.519 (0.018)	0.615 (0.036)	0.504 (0.025)
AR2 $\hat{\phi}_2$	-	-	-	-0.086 (0.029)	-0.053 (0.030)
AR3 $\hat{\phi}_2$	-	-	-	-	0.120 (0.026)
MA1 $\hat{\theta}_1$	0.253 (0.041)	0.157 (0.047)	-	-	-
Variance $\hat{\sigma^2}$	0.727 (0.028)	-	1.284 (0.031)	-	-
			Conditional variance		
Constant $\hat{\omega}$	-	1.590 (0.055)	-	1.131 (0.047)	0.046 (0.018)
ARCH $\hat{\alpha}$	-	0.126 (0.029)	-	0.311 (0.031)	0.031 (0.008)
GARCH $\hat{\beta}$	-	-	-	-	0.947 (0.015)
			Marginal distribution		
DoF $\hat{\nu}$	-	6.447 (0.931)	6.734 (1.006)	6.563 (1.063)	8.111 (1.680)
Skewness $\hat{\lambda}$	-	-0.321 (0.032)	-0.215 (0.033)	-0.322 (0.036)	-0.278 (0.033)

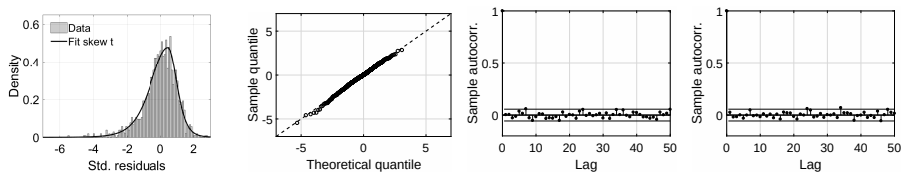
Table 1: Type and order of marginal models, parameter estimates, and corresponding standard errors in parenthesis.



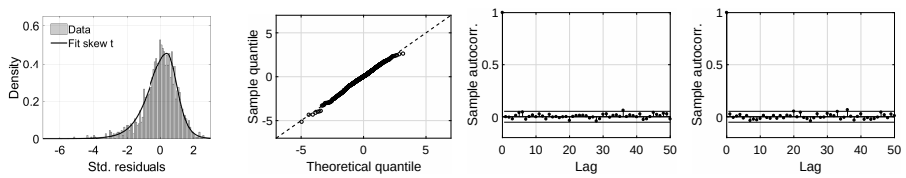
(a) Histogram (site 2) (b) QQ plot (site 2) (c) Acf of $\hat{\eta}_t^i$ (i=2) (d) Acf of sq. $\hat{\eta}_t^i$ (i=2)



(e) Histogram (site 19) (f) QQ plot (site 19) (g) Acf of $\hat{\eta}_t^i$ (i=19) (h) Acf of sq. $\hat{\eta}_t^i$ (i=19)



(i) Histogram (site 20) (j) QQ plot (site 20) (k) Acf of $\hat{\eta}_t^i$ (i=20) (l) Acf of sq. $\hat{\eta}_t^i$ (i=20)



(m) Histogram (site 31) (n) QQ plot (site 31) (o) Acf of $\hat{\eta}_t^i$ (i=31) (p) Acf of sq. $\hat{\eta}_t^i$ (i=31)

Fig. 4: Model check for four selected sites: 2, 19, 20, and 31.

Having established marginal models in the previous section, we are now left with the construction of the copula C. To this end, we apply the probability integral transform to the standardized residuals to obtain

$$\hat{u}_t^{DE} = F^{DE}(\hat{\eta}_t^{DE}|\mathcal{F}_{t-1}),$$
$$\hat{u}_t^{i} = F^{i}(\hat{\eta}_t^{i}|\mathcal{F}_{t-1}),$$

for $i = 1,\dots,31$. F^i and F^{DE} are distribution functions corresponding to each wind site and the German wind index, respectively. Fig. 5 presents the estimated uniforms for the four selected sites plotted against $\hat{u}_t^{DE}$, showing variations in the strength of the dependence, which we already expected considering the results in Fig. 1(b). However, Fig. 5 also indicates that the shapes of the dependence structure could possibly differ, and thus the preferred copula specifications might vary across the wind sites. Specifically, notice that Figs. 5(b) and 5(d) exhibit a slight asymmetric behavior with more concentration in the north east corner compared to the other corners. This pattern is not as clear in Figs. 5(a) and 5(c).

The copula models we consider are the Gaussian, Gumbel, rotated Gumbel (RGumbel), Clayton, rotated Clayton, Frank, symmetrized Joe-Clayton, and Student t^2. These are fitted to all 31 pairs and the Akaike Information Criterion (AIC) is used as model selection criterion. To introduce further flexibility, we also consider copula mixtures as in Rodriguez (2007) and Dias and Embrechts (2009). A copula mixture is given by

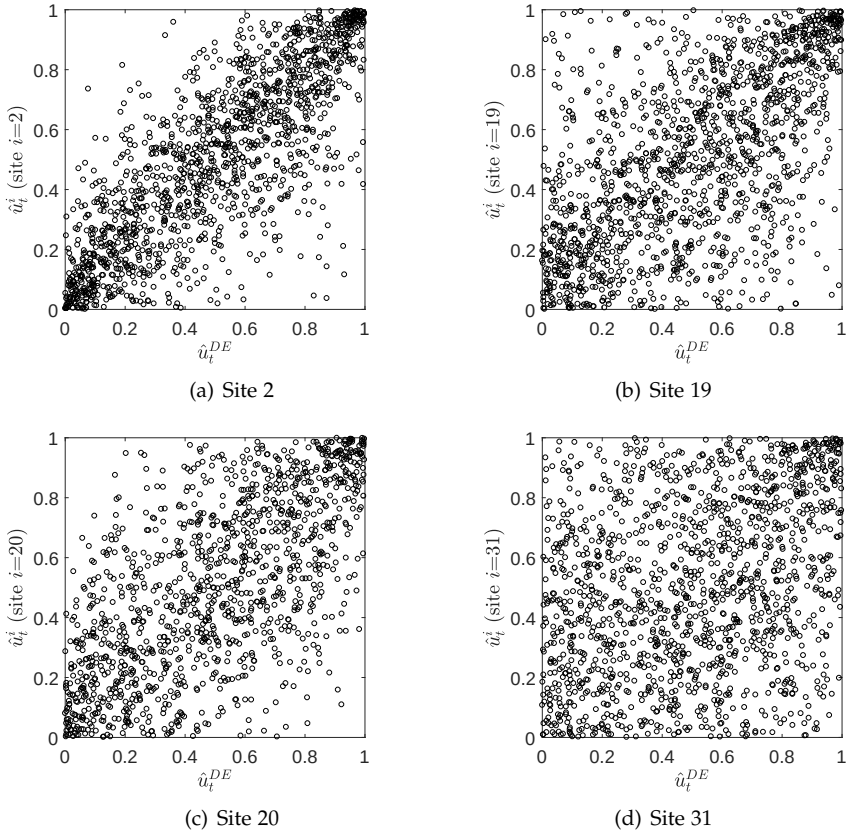
$$c^m(\cdot|\mathcal{F}_{t-1},\boldsymbol{\theta}^a,\boldsymbol{\theta}^b,\lambda) = \lambda c^a(\cdot|\mathcal{F}_{t-1},\boldsymbol{\theta}^a) + (1-\lambda)c^b(\cdot|\mathcal{F}_{t-1},\boldsymbol{\theta}^b), \tag{5}$$

where $c^a(\cdot|\mathcal{F}_{t-1},\boldsymbol{\theta}^a)$ is the copula density for copula a having parameter vector $\boldsymbol{\theta}^a$, and $c^b(\cdot|\mathcal{F}_{t-1},\boldsymbol{\theta}^b)$ is the copula density for copula b with parameter vector $\boldsymbol{\theta}^b$. The mixing parameter $0 < \lambda < 1$ controls the proportion of each copula.

In addition to the already mentioned copulas, we employ the following mixtures: Gumbel/RGumbel, Gaussian/Gumbel, Gaussian/RGumbel, Frank/Gumbel, and Frank/RGumbel. For each site, the optimal copula model is reported in Table 2. Moreover, we include in Table 2 the implied Kendall's τ, upper tail dependence, and lower tail dependence.[3] The optimal copula differs across the 31 wind sites, with Gaussian/Gumbel and Frank/Gumbel being preferred in most cases. These copula mixtures are both characterized by an asymmetric dependence structure and upper tail dependence. According to the results in Table 2, only two models corresponding to

[2]Rotation refers to a 180 degree rotation of data.

[3]Recall that tail dependence measures the dependence in extreme events. For positively related data, the upper tail dependence is defined as $\tau^U = \lim_{q\to 1^-} P(u_{1,t} > q|u_{2,t} > q)$, where q denotes the quantile and the u's represent standard uniform variables. The lower tail dependence is defined analogously.



(a) Site 2 (b) Site 19

(c) Site 20 (d) Site 31

Fig. 5: Empirical copula density plots for four selected sites, with site ID 2, 19, 20, and 31.

northern sites imply lower tail dependence, while upper tail dependence is a common feature across all models. The upper tail dependence corresponds to scenarios with high German wind index and high site-specific wind index. For the wind power generator, this feature weakens the power of WPF as hedging instruments, which we shall elaborate in the next section.

To test whether or not the optimal copulas in Table 2 are well-specified, we perform the Kolmogorov-Smirnov (K-S) and the Cramer-von Mises (CvM) goodness-of-fit tests (see Berg (2009), Genest et al. (2009) and Patton (2013) for a detailed description). The resulting *p*-values are reported in the last two columns of Table 2 and are based on 999 simulations, indicating that the null hypothesis of a well-specified copula cannot be rejected.

ID	Optimal copula	Kendall's τ	Upper tail dep.	Lower tail dep.	K-S	CvM
1	Mix Gaussian/Gumbel	0.582 (0.011)	0.246 (0.060)	-	0.893	0.886
2	Mix Gumbel/RGumbel	0.606 (0.008)	0.358 (0.038)	0.328 (0.039)	0.512	0.527
3	Mix Gaussian/Gumbel	0.534 (0.010)	0.004 (0.044)	-	0.864	0.989
4	Student t	0.642 (0.010)	0.215 (0.123)	0.215 (0.123)	0.968	0.998
5	Mix Gaussian/Gumbel	0.603 (0.009)	0.074 (0.059)	-	0.809	0.812
6	Mix Gaussian/Gumbel	0.516 (0.010)	0.189 (0.057)	-	0.902	0.955
7	Mix Gaussian/Gumbel	0.576 (0.009)	0.056 (0.050)	-	0.857	0.818
8	Mix Gaussian/Gumbel	0.609 (0.008)	0.082 (0.054)	-	0.881	0.877
9	Mix Gaussian/Gumbel	0.647 (0.008)	0.333 (0.069)	-	0.783	0.793
10	Mix Frank/Gumbel	0.453 (0.012)	0.211 (0.048)	-	0.851	0.755
11	Mix Frank/Gumbel	0.492 (0.011)	0.146 (0.047)	-	0.900	0.765
12	Mix Gaussian/Gumbel	0.659 (0.007)	0.239 (0.066)	-	0.683	0.789
13	Mix Gaussian/Gumbel	0.613 (0.009)	0.261 (0.068)	-	0.590	0.642
14	Mix Frank/Gumbel	0.554 (0.011)	0.373 (0.048)	-	0.683	0.823
15	Mix Frank/Gumbel	0.570 (0.010)	0.376 (0.046)	-	0.510	0.582
16	Mix Frank/Gumbel	0.513 (0.011)	0.494 (0.043)	-	0.716	0.797
17	Mix Frank/Gumbel	0.479 (0.012)	0.395 (0.043)	-	0.358	0.533
18	Mix Frank/Gumbel	0.570 (0.010)	0.246 (0.050)	-	0.876	0.783
19	Mix Frank/Gumbel	0.437 (0.013)	0.363 (0.045)	-	0.765	0.703
20	Mix Frank/Gumbel	0.470 (0.013)	0.190 (0.052)	-	0.958	0.911
21	Mix Gaussian/Gumbel	0.385 (0.015)	0.223 (0.058)	-	0.982	0.972
22	Mix Frank/Gumbel	0.506 (0.011)	0.302 (0.048)	-	0.796	0.757
23	Mix Frank/Gumbel	0.510 (0.013)	0.302 (0.048)	-	0.742	0.864
24	Mix Frank/Gumbel	0.416 (0.014)	0.196 (0.045)	-	0.504	0.664
25	Mix Gaussian/Gumbel	0.383 (0.014)	0.271 (0.068)	-	0.800	0.880
26	Mix Frank/Gumbel	0.437 (0.013)	0.250 (0.048)	-	0.773	0.781
27	Mix Gaussian/Gumbel	0.349 (0.013)	0.260 (0.059)	-	0.763	0.769
28	Mix Frank/Gumbel	0.390 (0.014)	0.298 (0.046)	-	0.457	0.558
29	Mix Frank/Gumbel	0.359 (0.015)	0.186 (0.046)	-	0.696	0.776
30	Mix Frank/Gumbel	0.361 (0.014)	0.282 (0.046)	-	0.766	0.758
31	Gumbel	0.215 (0.015)	0.277 (0.018)	-	0.756	0.778

Table 2: Optimal copula, Kendall's τ, upper and lower tail dependence implied by the optimal copula, and p-values resulting from performing goodness-of-fit tests. Standard errors are reported in parenthesis. The symbol "-" indicates that the tail dependence imposed by the chosen copula is zero. Kendall's τ for a convex combination of copula a and copula b, individually imposing a Kendall's τ of τ^a and τ^b, respectively, is $\lambda\tau^a + (1-\lambda)\tau^b$. Likewise, the upper (lower) tail dependence is a convex combination of the individual upper (lower) tail dependence coefficients. See Christensen et al. (2017) for more information.

4 Spatial hedging benefits of wind power futures

In this section, the models presented in Sec. 3 are employed to asses the hedging effectiveness of WPF for the different wind sites. As in Sec. 3, we limit ourselves to presenting results for four chosen sites – if otherwise, it will be clear from the context. In order to perform the hedging exercise, we assume the following: Wind power producers receive a fixed price of 30 EUR per produced MWh regardless of geographical location, and the installed capacity of the wind turbines at each wind site is 100 MW. Thus, from Eq. (1) we get that the daily wind power generation at site i and day t is

$$P_{t,i} = 100 \cdot W_{t,i} \cdot h,$$

and hence the wind power producer's unhedged profit at site i, over the period $[T_1, T_2]$, is given by

$$R_i^U = 30 \sum_{t=T_1}^{T_2} P_{t,i}. \tag{6}$$

When hedging the volume risk associated with wind power generation using a position $\gamma_i \in \mathbb{Z}$ in WPF – recall Eq. (2) – we get the following hedged profit

$$R_i^H = R_i^U + \gamma_i R^{WPF}. \tag{7}$$

To determine γ_i, we use 20,000 Monte Carlo simulations from the proposed models, each spanning a year from 1 January 2016. Since the end of the in-sample data is 31 December 2015, it follows that the simulations are performed out-of-sample. We shall restrict our attention to WPF with monthly delivery, and divide therefore each simulated path into 12 parts corresponding to the length of each month. This way, we are able to asses monthly differences in the hedging effectiveness of WPF across a calender year.

Regarding pricing of WPF, i.e., estimating W_{t_0} in Eq. (2) for each month, we note that t_0 is fixed at 31 December 2015 regardless of which monthly delivery we consider. Assuming a zero market price of risk, we obtain W_{t_0} as a simple average of Monte Carlo simulations covering the specified delivery period. We stress that this assumption implies that the expected value of the profit distribution will not be affected by changing γ_i. Thus, we concentrate on a minimum variance hedge, and determine γ_i for each month by

$$\min_{\gamma_i \in \mathbb{Z}} \mathbb{V}\mathrm{ar}[R_i^H]. \tag{8}$$

4.1 Empirical results

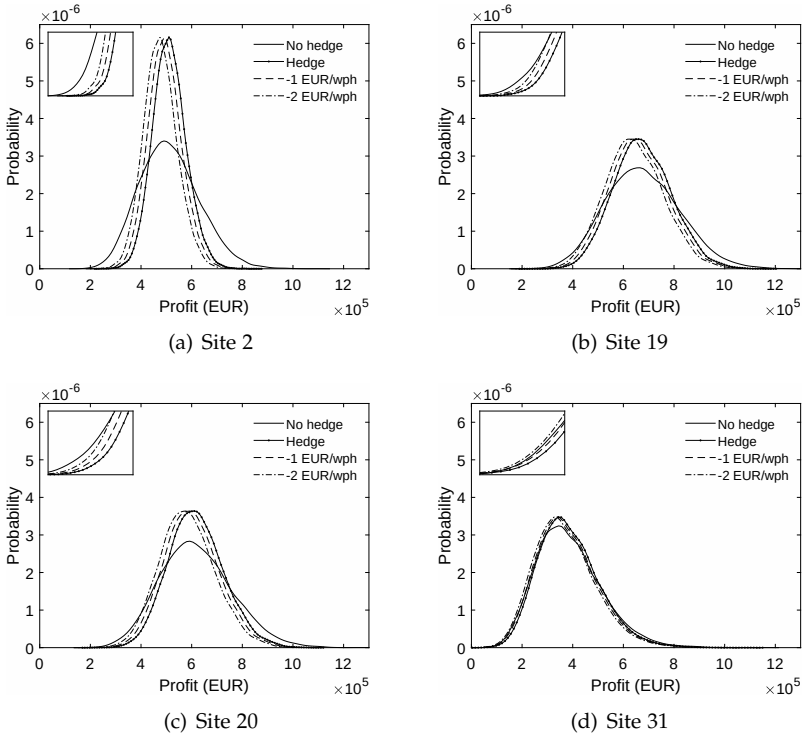
Using the specifications above, we estimate the unhedged profit distributions for all sites and all months of the year. In Fig. 6, we present the results for January and sites 2, 19, 20, and 31. Note that we are able to compare the profits, since the wind generation capacity is set at 100 MW for all sites. The unhedged profit distributions have different means, with site 31 having the lowest value. Regarding the hedged profit distributions, which are also included in Fig. 6, we observe that WPF are most beneficial for site 2, followed by sites 19 and 20, which achieve similar benefits. Hedging benefits for site 31 are lowest, since this site is least correlated to the German wind index cf. Table 2. Although Fig. 6 only illustrates the case for January, we find similar results for the remaining months. In fact, the average difference between the maximum and minimum variance reductions throughout the year for all sites is only 2.3%, and all range between 1% and 5%.

Motivated by the earlier studies of Gersema and Wozabal (2017) and Benth and Pircalabu (2017), who argue for the existence of a negative risk premium in the German wind power futures market, we consider next the effect of including such a quantity in our analysis. Cf. Eq. (2), a risk premium of e.g. -1 EUR/wph translates to a reduction in W_{t_0} of 1 %, and implies that WPF are sold at a discount. To indicate the effect of a negative risk premium on the hedged profit distribution, we plot in Fig. 6 the results of assuming a risk premium of -1 EUR/wph and -2 EUR/wph. We note that the hedge quantities remain unchanged, since hedging decisions are based on a minimum variance hedge, thus disregarding the cost of the hedge portfolio. In the hedging exercise considered here, this implies that a negative risk premium shifts the hedged profit distribution to the left compared to the case of a zero risk premium, as it appears from Fig. 6.[4]

To study further the implications of a negative risk premium, we consider the reduction in the mean of the hedged profit distribution when taking into account a risk premium of -2 EUR/wph, as opposed to none at all. The results are plotted in Fig. 7(a) on a monthly basis, and can be explained by the seasonality in the marginal models in combination with a static copula. On one hand, the seasonal behavior embedded in the site-specific wind index causes the profit corresponding to the naked position R_i^U to be lowest during summer and highest during winter. On the other hand, the optimal hedge quantity γ^* in WPF is highly influenced by the dependence – which is constant throughout the year – thus translating to a fairly constant γ^* as illustrated in Fig. 7(b).

Clearly, since wind power generators seek to cover their exposure to low wind scenarios, they would be more inclined to use WPF as hedging instruments during the summer months. According to the results in Fig. 7(a) however, the reduction in the mean of the hedged profit distribution is approximately double as high during summer compared to winter, which is highly unfavorable. We stress that the results in Fig. 7(a) only hold under the assumption of a constant risk premium across the months of a calendar year, which cannot be the case in practice. Nevertheless, since the empirical investigation carried out in Benth and Pircalabu (2017) does not suggest any clear seasonal pattern in the market price of risk associated with German WPF, we argue that our results are relevant from a practical perspective. Furthermore, the analysis above would only be rendered superfluous in a situation where the risk premium was shaped as to counteract the yearly seasonality in the wind index.

[4]In practice, the hedger is not indifferent to the cost of the hedging portfolio. The objective in Eq. (8) can easily be extended to e.g. a situation where the variance of the profit distribution is minimized while also minimizing the cost of the hedging portfolio, based on some preference regarding the trade-off between variance and cost reduction. This is however outside the scope of the present paper, and shall not be pursued further.



(a) Site 2 (b) Site 19

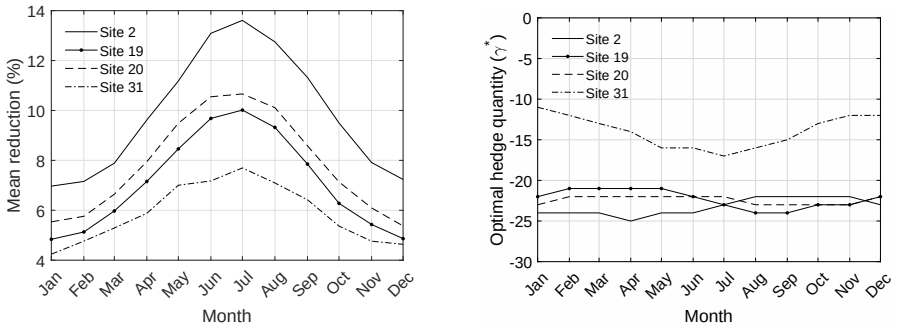
(c) Site 20 (d) Site 31

Fig. 6: Unhedged and hedged profit distributions for wind sites 2, 19, 20 and 31 cf. Fig. 1(a). The results are obtained by employing the marginal models determined in Sec. 3.1 and the optimal copulas cf. Table 2. The hedged distributions are computed with a risk premium corresponding to 0 EUR/wph, -1 EUR/wph and -2 EUR/wph. The optimal hedge quantities are obtained by solving Eq. (8), and do not vary with the risk premium.

On a different note, we mention in passing that the yearly seasonality associated with wind power production is also important in the context of liquidity in the German WPF market. According to Christensen et al. (2017), hedging benefits for conventional power producers are highest during winter. Hence, a large part of the buyer side is incentivized to take a position in WPF during winter, whereas the seller side is incentivized to do so during summer.

Next, we consider the variance reduction achieved by hedging with WPF. In Fig. 8(a), we present the estimated average variance reductions throughout the year. As expected, the variance reductions and geographical location of wind turbines are related, with the reductions decreasing as we move from north to south Germany. The results vary substantially across the wind sites, with the highest numbers being above 70%, and the lowest number being 13%.

Also relevant to consider in this context is the effect of variance reduction



(a) Percentage reduction in mean (b) Positions in WPF

Fig. 7: Percentage reduction in mean of the hedged profit distribution when including a risk premium of -2 EUR/wph as opposed to no risk premium, and the optimal number of WPF contracts for each month throughout the year. The results are displayed for wind sites 2, 19, 20, and 31 cf. Fig. 1(a).

on the tails of the hedged profit distribution. Under the assumption of a zero risk premium, we wish to quantify whether the right tail is reduced by more than the left tail, i.e., are we reducing the probability of a very high income by more than that of a very low income? To answer this question, we let $q^H_{j,i}$ denote the jth quantile of the demeaned hedged profit distribution for site i,

$$q^H_{j,i} = F^{-1}(R^H_i - \mathbb{E}[R^H_i], j).$$

Similarly, $q^U_{j,i}$ denotes the jth quantile of the demeaned unhedged profit distribution. The percentage change of the jth quantile from performing the WPF hedge can then be measured by

$$T^l_i(j) = \frac{q^U_{j,i} - q^H_{j,i}}{q^U_{j,i}}.$$

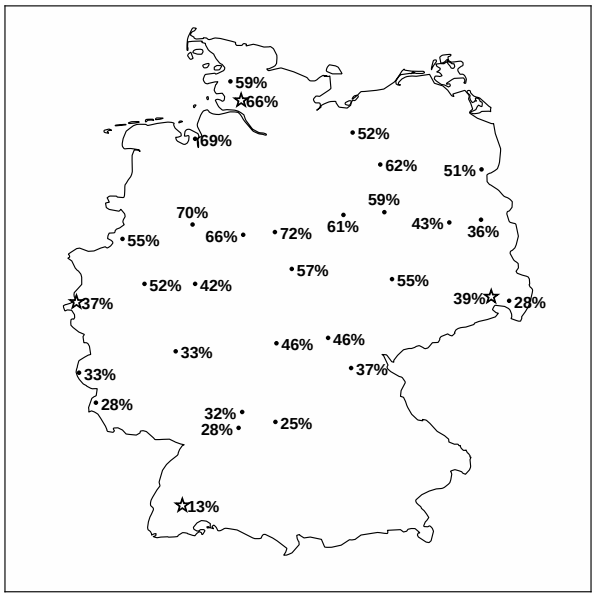
To measure the difference between the change in the tails, we define the variable,

$$T_i(j) = |T^l_i(j)| - |T^r_i(j)|,$$

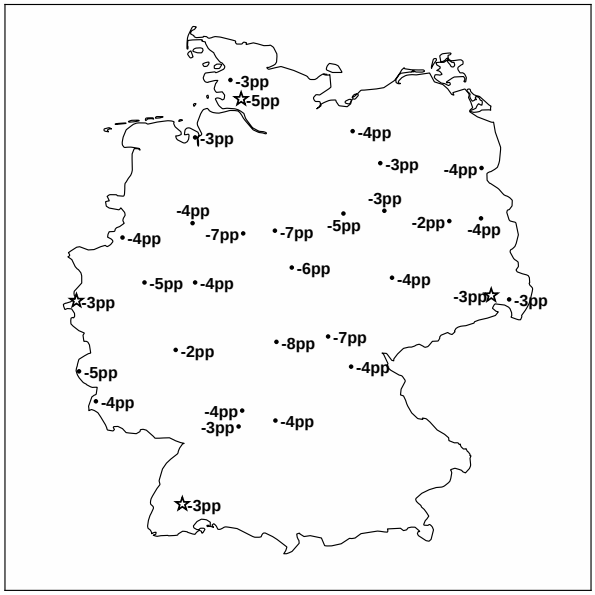
where

$$T^r_i(j) = \frac{q^H_{1-j,i} - q^U_{1-j,i}}{q^U_{1-j,i}}.$$

The quantity $T^l_i(j)$ defined above will be positive if we shrink the left tail (jth quantile) when performing the WPF hedge, which is the case for all sites.



(a) Average variance reductions



(b) $T_i(0.05)$

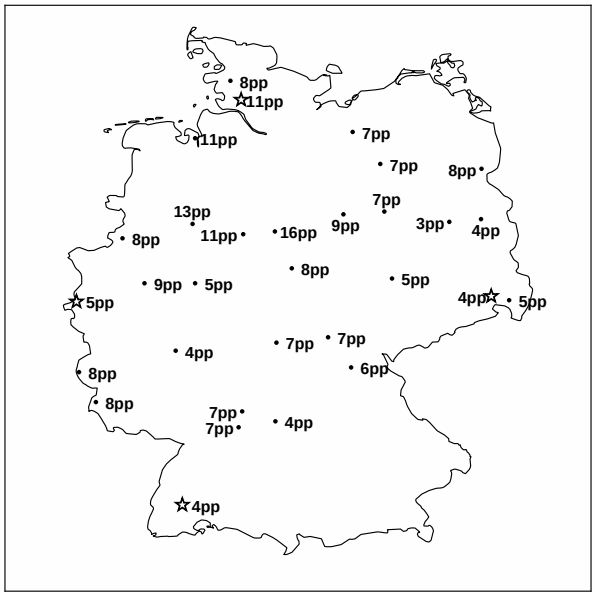
Fig. 8: Average variance reduction for each wind site using the optimal copula cf. Table 2, and the difference between the change in the tails of the profit distribution measured by $T_i(0.05)$. The stars indicate the location of the sites with ID 2, 19, 20, and 31.

Further, $T_i^r(j)$ will be negative if we shrink the right tail ($1-j$th quantile) when performing the WPF hedge, which is also the case for all sites. Hence, a negative value of $T_i(j)$ means that the right tail has shrunk more than the left tail. We note that $T_i(j)$ is measured in percentage points (pp). Estimates for $T_i(0.05)$ are reported in Fig. 8(b), revealing negative values for all sites, with -8pp being the lowest. Interestingly, the magnitude of the average variance reduction does not seem to influence the magnitude of $T_i(0.05)$, cf. Figs. 8(a) - 8(b). For example, the lowest value of $T_i(0.05)$ corresponds to a variance reduction of 46%, which is significantly lower than the highest variance reduction of 72%.

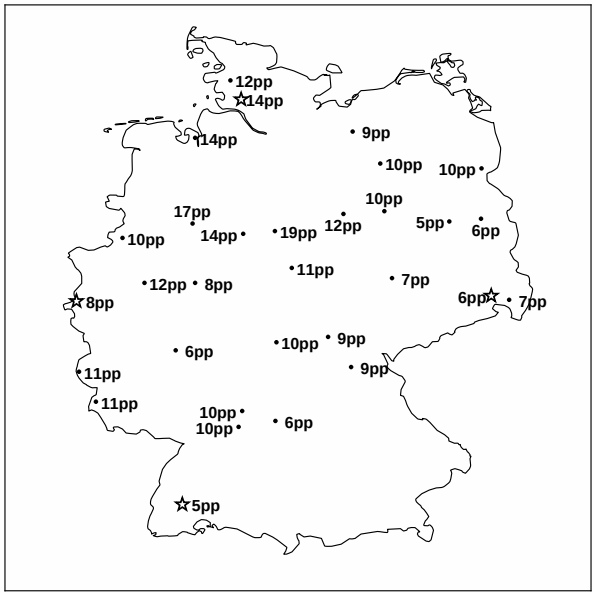
Since a zero risk premium is imposed concerning the results in Fig. 8(b), the negative values of $T_i(0.05)$ obtained for all sites can be explained by 1) the marginal specifications and 2) the presence of asymmetry and tail dependence in the copula. On one hand, the marginal distributions of wind indexes are skewed, with a heavy tail to the right. On the other hand, cf. Table 2, evidence for asymmetry and upper tail dependence is found for most wind sites, entailing that site-specific wind indexes are related to the German index in extreme high wind scenarios. Since wind power producers seek to cover their exposure to low wind scenarios, lower tail dependence would be much preferred to the upper tail dependence that we find in the data. It is however difficult to separate the effects related to the margins from those related to the dependence structure without further analysis. In the next section, we return to this issue. Lastly, we stress that the results in Fig. 8(b) are obtained assuming a zero risk premium. In the presence of a negative risk premium, the negative effects illustrated in Fig. 8(b) will clearly be magnified.

4.2 Comparison study

To highlight the effects of employing a copula that allows for asymmetry and upper tail dependence, we consider in this section a comparison study based on a naive copula. Specifically, we rerun the computations performed in Section 4.1 with the Frank copula, while keeping the marginal models fixed. We employ the Frank copula as a naive alternative since Frank is represented in the optimal copula mixtures in 16 out of 31 cases cf. Table 2. In addition, the Frank copula is symmetric and imposes zero tail dependence, whereas the optimal copulas are all asymmetric (except for site 4) and imply tail dependence (upper tail dependence for the most part). Further, by performing the K-S and CvM goodness-of-fit tests, we cannot reject the Frank copula for any of the 31 sites at a 5% significance level. In fact, if excluding the copula mixtures from our analysis, Frank would become the preferred copula specification for many wind sites based on AIC. In the absence of copula mixtures, a hedger or risk manager could therefore easily have employed a Frank copula to model dependence.



(a) Average difference



(b) Maximum difference

Fig. 9: Average and maximum difference in variance reductions implied by the optimal copula cf. Table 2 and the Frank copula (optimal minus naive). The maximum differences refer to the largest monthly difference. The stars indicate the location of the sites with ID 2, 19, 20, and 31.

In Fig. 9(a), we report the average difference in the percentage variance reduction across the months of the year, when going from the optimal copula cf. Table 2 to the Frank copula. We identify increases in the average variance reduction for all sites when using the optimal copula. We complement these findings with Fig. 9(b), where we report the maximum monthly difference between percentage variance reduction implied by the optimal copula and Frank. The maximum difference reaches as high as 19pp, revealing that using a less optimal model can have a substantial impact. In a nutshell, we would simply undermine the variance reduction strength of WPF with the Frank copula.

Finally, recall Fig. 8(b), where it was difficult to separate the effects from the marginals and the dependence structure. Performing similar calculations as in the case of Fig. 8(b) with the Frank copula yields the estimates of $T_i(0.05)$ given in Fig. 10. Compared to Fig. 8(b), we find that introducing a different dependence structure accounts for approximately 2-3pp of the difference between left and right tail percentage change. Hence, using the Frank copula would lead to an underestimation of the difference between the change in tails implied by hedging with WPF.

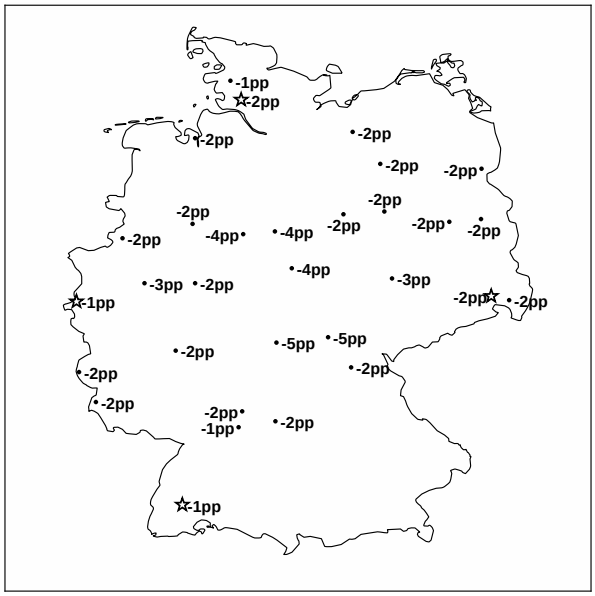


Fig. 10: Estimates of $T_i(0.05)$ when performing the hedge with WPF using the naive approach, i.e, the Frank copula. The stars indicate the location of the sites with ID 2, 19, 20, and 31.

5 Conclusion

In this paper, we analyze the hedging effectiveness of wind power futures for wind power generators with 31 different locations in Germany. We propose ARMA–GARCH copula models for each pair of German wind index and site-specific wind index. The ARMA–GARCH type models capture the marginal behavior of the variables rather well, and a detailed study of the different dependence structures reveals a general tendency of asymmetric dependence and upper tail dependence, thus justifying the use of copulas.

Based on Monte Carlo simulations from the proposed models, the benefits of wind power futures are quantified through a minimum variance hedge, which we apply for each wind site. The variance reductions vary from 13% to 72%, with the lowest values corresponding to wind farms located in the south of Germany. Motivated by earlier findings in the literature that argue for the existence of a negative risk premium in the German wind power futures market, we extend our analysis to include this feature. Compared to the case of a zero risk premium, the hedged profit distribution is shifted to the left, thus weakening the benefits of wind power futures in the context of our paper. Considering how the wind power futures hedge affects the tails of the profit distributions, our results indicate a higher decrease in the upside potential compared to the decrease in the downside risk. While this is expected due to the presence of skewness in the unhedged profit distribution, the asymmetric dependence plays an important role as well.

Lastly, we highlight the importance of capturing asymmetry and upper tail dependence by performing a comparison study, where the hedging exercise is carried out assuming a naive model for the dependence structure. With the naive specification, the variance reduction resulting from the wind power futures hedge would be undermined. Further, approximately half of the difference in reduction between the tails the hedged profit distribution is caused by the asymmetric dependence.

Acknowledgments

The authors would like to thank Esben Høg and the Quantitative Analytics team at Neas Energy for providing valuable comments and suggestions.

Funding

Troels Sønderby Christensen is supported by the Innovation Fund Denmark under Grant 5189-00117B. Anca Pircalabu is supported by the Innovation Fund Denmark under Grant 4135-00082B.

References

M. D. Aepli, R. Füss, T. E. S. Henriksen, and F. Paraschiv. Modeling the multivariate dynamic dependence structure of commodity futures portfolios. *Journal of Commodity Markets*, 6:66 – 87, 2017.

K. Avdulaj and J. Barunikl. Are benefits from oil–stocks diversification gone? New evidence from a dynamic copula and high frequency data. *Energy Economics*, 51:31 – 44, 2015.

F. E. Benth and P. C. Kettler. Dynamic copula models for the spark spread. *Quantitative Finance*, 11(3):407 – 421, 2011.

F. E. Benth and A. Pircalabu. A non-Gaussian Ornstein-Uhlenbeck model for pricing wind power futures. *Working paper*, 2017.

D. Berg. Copula goodness-of-fit testing: An overview and power comparison. *European Journal of Finance*, 15(7-8):675 – 701, 2009.

T. S. Christensen, A. Pircalabu, and E. Høg. A seasonal copula mixture for hedging the clean spark spread with wind power futures. *Working paper*, 2017.

A. Dias and P. Embrechts. Testing for structural changes in exchange rates' dependence beyond linear correlation. *The European Journal of Finance*, 15 (7–8):619 – 637, 2009.

C. Elberg and S. Hagspiel. Spatial dependencies of wind power and interrelations with spot price dynamics. *European Journal of Operational Research*, 241(1):260 – 272, 2015.

C. Genest, B. Remillard, and D. Beaudoin. Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and Economics*, 44(2): 199 – 213, 2009.

G. Gersema and D. Wozabal. An equilibrium pricing model for wind power futures. *Energy Economics*, 65:64 – 74, 2017.

O. Grothe and J. Schnieders. Spatial dependence in wind and optimal wind power allocation: A copula-based analysis. *Energy Policy*, 39(9):4742 – 4754, 2011.

B.-Y. Liu, Q Ji, and Y. Fan. Dynamic return-volatility dependence and risk measure of CoVaR in the oil market: A time-varying mixed copula model. *Energy Economics*, 68:53 – 65, 2017.

A. J. Patton. Modelling asymmetric exchange rate dependence. *International Economic Review*, 47(2):527 – 556, 2006.

A. J. Patton. Copula methods for forecasting multivariate time series. In Graham Elliott and Allan Timmermann, editors, *Handbook of Economic Forecasting*, volume 2B, pages 899 – 960. Elsevier B.V., 2013.

A. Pircalabu and J. Jung. A mixed C-vine copula model for hedging price and volumetric risk in wind power trading. *Quantitative Finance*, 17(10): 1583 – 1600, 2017.

A. Pircalabu, T. Hvolby, J. Jung, and E. Høg. Joint price and volumetric risk in wind power trading: A copula approach. *Energy Economics*, 62:139 – 154, 2017.

J. C. Rodriguez. Measuring financial contagion: A copula approach. *Journal of Empirical Finance*, 14(3):401 – 423, 2007.

A. Sklar. Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de L'Université de Paris*, 8:229 – 231, 1959.

ISSN (online): 2446-1636
ISBN (online): 978-87-7210-123-1

AALBORG UNIVERSITY PRESS