Hain, Martin; Kargus, Tobias; Schermeyer, Hans; Uhrig-Homburg, Marliese; Fichtner, Wolf

Working Paper

An electricity price modeling framework for renewable-dominant markets

Working Paper Series in Production and Energy, No. 66

Provided in Cooperation with:
Karlsruhe Institute of Technology (KIT), Institute for Industrial Production (IIP)

Suggested Citation: Hain, Martin; Kargus, Tobias; Schermeyer, Hans; Uhrig-Homburg, Marliese; Fichtner, Wolf (2022) : An electricity price modeling framework for renewable-dominant markets, Working Paper Series in Production and Energy, No. 66, Karlsruhe Institute of Technology (KIT), Institute for Industrial Production (IIP), Karlsruhe, https://doi.org/10.5445/IR/1000151367

This Version is available at:
https://hdl.handle.net/10419/267803

Karlsruhe Institute of Technology

An Electricity Price Modeling Framework for Renewable-Dominant Markets

Martin Hain, Tobias Kargus, Hans Schermeyer,
Marliese Uhrig-Homburg, Wolf Fichtner

No. 66 | OCTOBER 2022

WORKING PAPER SERIES IN PRODUCTION AND ENERGY

An Electricity Price Modeling Framework for Renewable-Dominant Markets

Martin Hain[a], Tobias Kargus[a], Hans Schermeyer[b], Marliese Uhrig-Homburg[a], Wolf Fichtner[b]

[a]Chair of Financial Engineering and Derivatives, Institute for Finance (FBV), Karlsruhe Institute of Technology (KIT)
Blücherstr. 17, Building 09.21, 76185 Karlsruhe, Germany
Phone: +49 721 608-48183, Email: derivate@fbv.kit.edu

[b]Chair of Energy Economics, Institute for Industrial Production (IIP), Karlsruhe Institute of Technology (KIT)
Hertzstr. 16, Building 06.33, 76187 Karlsruhe, Germany
Phone: +49 721 608-44460, Email: info@iip.kit.edu

Abstract

Renewables introduce new weather-induced patterns and risks for market participants active in the energy commodity sector. We present a flexible framework for power spot prices that is capable of incorporating a weather model for the joint distribution of local weather conditions. This not only allows us to make use of a long history of local weather data in the calibration procedure but also makes it possible to assess how changes in the renewable generation portfolio impact the characteristics of future wholesale spot prices. Empirical tests demonstrate the model's capability to reproduce salient features of market variables. We furthermore show why our model offers unique benefits for market players compared to existing approaches.

An Electricity Price Modeling Framework for Renewable-Dominant Markets*

Martin Hain†, Tobias Kargus‡, Hans Schermeyer§,
Marliese Uhrig-Homburg¶, and Wolf Fichtner‖

This version: September 2022

Abstract

Renewables introduce new weather-induced patterns and risks for market participants active in the energy commodity sector. We present a flexible framework for power spot prices that is capable of incorporating a weather model for the joint distribution of local weather conditions. This not only allows us to make use of a long history of local weather data in the calibration procedure but also makes it possible to assess how changes in the renewable generation portfolio impact the characteristics of future wholesale spot prices. Empirical tests demonstrate the model's capability to reproduce salient features of market variables. We furthermore show why our model offers unique benefits for market players compared to existing approaches.

*The paper is a revised and extended version of the earlier paper *Hain, Martin; Schermeyer, Hans; Uhrig-Homburg, Marliese; Fichtner, Wolf (2017): An Electricity Price Modeling Framework for Renewable-Dominant Markets. Karlsruhe (Working Paper Series in Production and Energy, 23).*

†Martin Hain, Chair of Financial Engineering and Derivatives, Karlsruhe Institute of Technology (KIT), P.O. Box 6980, D-76049 Karlsruhe, Germany, Email: derivate@fbv.kit.edu.

‡Tobias Kargus, Chair of Financial Engineering and Derivatives, Karlsruhe Institute of Technology (KIT), P.O. Box 6980, D-76049 Karlsruhe, Germany, Phone: +49 721 608 48191, Email: tobias.kargus@kit.edu.

§Hans Schermeyer, Institute for Industrial Production (IIP), Karlsruhe Institute of Technology (KIT), P.O. Box 6980, D-76049 Karlsruhe, Germany, Email: info@iip.kit.edu.

¶Marliese Uhrig-Homburg, Chair of Financial Engineering and Derivatives, Karlsruhe Institute of Technology (KIT), P.O. Box 6980, D-76049 Karlsruhe, Germany, Phone: +49 721 608 48183, Email: marliese.uhrig-homburg@kit.edu.

‖Wolf Fichtner, Institute for Industrial Production (IIP), Karlsruhe Institute of Technology (KIT), P.O. Box 6980, D-76049 Karlsruhe, Germany, Phone: +49 721 608 44460, Email: wolf.fichtner@kit.edu.

1 Introduction

State-of-the-art energy commodity models rely on well-established reduced-form approaches from the fixed-income literature. However, the transition to a low-carbon economy gives birth to renewable-dominant electricity markets and weather risk advances to a major risk factor. Meteorological studies provide concise and reliable information on weather data and have put forth powerful tools for modeling and understanding weather dynamics. Our contribution is to combine these insights and develop a new electricity price model that uses local weather conditions and the spatial distribution of renewable assets to model dynamics in wholesale electricity prices.

There are two main ingredients in the model. First, we rely on a suitable approach to model electricity prices. Instead of using reduced-form approaches that refrain from explicitly modeling fundamental supply and demand factors we take advantage of hybrid structural models. They allow us to incorporate the drivers of the supply side and consequently permit power prices to directly depend on weather conditions. Second, we model the temporal as well as spatial distribution of wind speed and solar irradiation, map these local weather conditions to electricity production, and incorporate the renewable electricity supply in our hybrid structural model. We show that this two-step specification has clear advantages in times of deep structural changes, which potentially render most parts of the historical price data obsolete.

The major strength of our stochastic price model is its ability to guide investment decisions, the assessment of hedging strategies, or policy decisions in a rapidly changing market environment. A natural first application is to study the quantitative impact of the local distribution of renewable generation capacities on the risk characteristics of power prices. Such an analysis should not only provide important insights on the potential future impact of the fast growing amount of renewable generation technology from a regulatory point of view but might be of interest for many other stakeholders in the electricity sector from owners of conventional power plants to potential investors in wind parks or solar farms at different locations.

Generally, stochastic price models allow market participants to calculate risk measures, to price certain risks and assets, or to deduct reasonable hedging strategies. By now, power prices require different modeling techniques than other energy commodities such as crude oil or natural gas for which established reduced-form approaches from the fixed-income literature have been shown to work quite satisfactory (e.g. Trolle and Schwartz (2009), Brooks

and Prokopczuk (2013), or Hain et al. (2018)). First, electricity markets are usually much more local in nature due to the fact that storing the commodity is mostly impossible and thus, power has to be produced to match local demand exactly in every instant. Paired with mostly price-inelastic demand this makes wholesale prices very sensitive to shocks in fundamentals such as unexpected weather changes or supply disruptions. These characteristics have prompted researchers to look into pricing models that consider the interaction of major supply- and demand factors and wholesale prices (e.g. Barlow (2002) or Howison and Coulon (2009)).

Second, once we add renewables to the equation things become even more complicated. Electricity generation from wind and solar power is itself highly sensitive to local weather conditions. Additionally and in strong contrast to conventional power plants, the spatial distribution of installed renewable capacity has a considerable impact on the characteristics of market-wide renewable power production. Compare, for example, a scenario in which all available renewable capacity is clustered within one single location with a more diversified scheme. Conditional on the joint distribution of local weather variables production in the former case is potentially much more volatile. This special "localness"-characteristic of wind and solar power complicates projections with regard to how wholesale prices react to capacity additions at different locations. And with renewable generation becoming more and more economically feasible, its share to total power production will almost surely continue to rise.

Our flexible approach is able to deal with the above challenges. Given that energy systems worldwide are entering and proceeding a phase of transformation, our approach may prove its benefits in a variety of electricity markets around the world. In this paper we detail the implementations for the German market: Germany has grown and continues to be one of the leaders of power generation out of renewable energy sources among large industrial nations (EurObserv'ER (2019)), its electricity market currently has one of the largest share of wind and solar power (International Renewable Energy Agency (2018)), and Germany has set very ambitious targets to further cut emissions drastically. So there is still much change ahead.

First and foremost, our research contributes to the field of electricity price modeling by proposing a methodology capable of incorporating renewables. While several approaches incorporate renewables as exogenous variables for modeling wholesale power spot prices (e.g. Cludius et al. (2014), Kallabis et al. (2016), or Lehna et al. (2021)) studies that explicitly model the stochasticity from renewable generation and incorporate it within a power price modeling framework are rather scarce. Keles et al. (2013) use a regression-based approach to incorporate global wind power generation within a regime-switching model for

electricity spot prices. Perhaps closest to our approach is the residual demand framework by Wagner (2014), which considers weather-driven demand and both market-wide wind and solar power production. Since direct modeling of wind and solar power completely ignores the spatial distribution and dynamic variations thereof, the residual demand framework of Wagner (2014) cannot distinguish between the potentially different impact of capacity additions at distinct locations. We therefore choose to model the local constituents forming the market-wide renewable generation. Intuitively, this can be seen as applying the basic idea behind structural price models for electricity markets recursively by additionally modeling the driving forces behind local renewable production which in turn drive wholesale electricity prices – an approach which we label the "Second-Layer Hybrid Structural model" (SLHS model in short). A major advantage of our SLHS modeling approach is the disentangling of the various drivers of wholesale electricity prices (demand, installed capacity, wind speed, solar irradiation) making it possible to calibrate part of our model to a rich history of weather data (over 28 years of local hourly weather variables). This is advantageous because there is only a relatively short history of renewable generation available and the potential impact of changes in the spatial distribution is captured naturally in our case while existing approaches can at best account for an absolute increase in installed capacity and might be misleading if the spatial distribution changes once again. Empirical tests furthermore demonstrate that our methodology is well capable of reproducing salient features in the time series of renewable production and wholesale day-ahead spot prices in Germany.

Our work is also related to the strand of literature analyzing the impact of renewable generation on wholesale electricity prices. Early work mostly considers how the average price level is affected (e.g. Jacobsen and Zvingilaite (2010) or Paraschiv et al. (2014)) whereas in following studies additional focus is put on price volatility as well (e.g. Jónsson et al. (2010), Ketterer (2014), or Wozabal et al. (2016)). All of the above studies rely on historical observations of aggregate renewable generation and spot prices entirely. However, as pointed out above, we cannot be sure how an increase of renewable capacity at different locations translates into the volatility of wholesale electricity prices. Our model could therefore help in fostering the understanding of this largely unexplored characteristic and may be used as a smart extrapolation tool to assess such impacts quantitatively.

Finally, our study is linked to the literature dealing with the assessment of site potential when faced with the difficult decision of choosing an optimal location for new physical assets. While Ritter et al. (2015) and Pieralli et al. (2015) outline potential issues with using idealized production curves for assessing wind power production potential, Grothe and Müsgens (2013) and Ritter and Deckert (2017) discuss spatial differences of revenues from wind parks. All of

the above studies remain silent with regard to the associated revenue uncertainty of specific sites. Also, as these studies rely on historical bootstrapping methods they are not well suited to validate how alterations to the renewable generation portfolio (or other important market variables) translate into the risk characteristics of different sites. In contrast, our methodology could help investors to form a better understanding of potential risks and returns associated with renewable energy projects without being purely backward-looking.

The structure of the subsequent sections is as follows. Section 2 introduces our general modeling framework, whereas section 3 gives details on how our approach allows us to incorporate a model for the joint distribution of weather variables. Section 4 then looks at other model components on the demand and supply side while section 5 discusses possible model extensions. Section 6 entails an analysis of the model's capability of reproducing salient features of important market variables and furthermore contains examples of how market participants can benefit from using our model. Section 7 concludes.

2 A Residual Demand Approach with Local Information

This section entails a detailed discussion of our chosen modeling framework. We briefly discuss how our model connects with various existing streams of modeling approaches and then show why and how we seek to introduce local information with regard to weather conditions and installed capacity into the framework. Due to the wide variety of modeling approaches, we focus on an aggregated overview of the literature and refer to Weron and Ziel (2019), Lago et al. (2021) and Lehna et al. (2021) for detailed comparative studies.

2.1 Hybrid Structural Price Modeling

State-of-the-art modeling approaches for electricity prices can broadly be separated into three categories. In the first category are approaches which heavily borrow from reduced-form models from equity- or interest-rate markets, sometimes also referred to as quantitative or stochastic approaches. Early studies use low-dimensional stochastic processes to capture patterns such as mean-reversion and seasonality (Lucia and Schwartz (2002)). In order to capture pronounced price spikes, unmatched in other commodity markets, Deng (2000), Cartea and Figueroa (2005), Geman and Roncoroni (2006), Seifert and Uhrig-Homburg (2007), or Hambly et al. (2009) consider variations of jump processes. Although these models generally share nice properties such as closed-form pricing formulas for derivatives, they are

often difficult to use in the rapidly changing electricity markets as their input contains only observed price data. For instance, changes in fundamentals (e.g. new large consumers (demand-side) or producers (new wind capacity)) can render historical price data completely useless in extreme cases. Also, since the liquidity of option contracts on power exchanges is usually very thin there is generally no reliable forward-looking information with regard to higher-moment price risk to look at in order to re-calibrate a reduced-form model.[1]

The second category consists of so-called structural models, sometimes referred to as fundamental models. In these approaches wholesale market clearing prices result from a problem in which demand must be satisfied under certain side restrictions such as transmission constraints (Eydeland and Wolyniec (2003)). The approach requires very detailed information on technical peculiarities of power plants or environmental constraints of the whole power market in question (fundamental input). A significant drawback is the fact that such models generally only make predictions on expected price levels and do not allow inference on higher moment price risks. This disqualifies them as a viable tool for risk management purposes such as hedging. The modeling and prediction of these fundamental inputs use diverse statistical or deep learning techniques, while still side restrictions and the functional associations of the fundamental input are central.

The third and recently fast growing category is composed by machine learning and advanced statistical techniques where neither stochastic processes as in reduced-form models nor functional associations between used exogenous variables as in structural models are the defining characteristics. Prominent approaches from advanced statistical modeling are models using the least absolute shrinkage and selection operator (LASSO) or the Lasso Estimated Autoregressive (LEAR) models (e.g. Uniejewski et al. (2016)). Considering machine learning, Recurrent Neuronal Networks (RNN), Nonlinear Autoregressive Neural Networks (NARX-NN), Long-Short Term Memory (LSTM), Gated Recurrent Unit (GRU) or Convolutional Neural Networks (CNN) are used (e.g. Lago et al. (2018)). Based on their flexibility with regard to non-linearities, high computational complexity and sensitivity to starting values, these approaches are rather used for technically constrained, short-term forecasts than for mid-term investment and hedging decisions (Weron and Ziel (2019)). In comparison however, structural models still reveal more fundamentally how factors drive market price. We therefore take statistically modeling techniques into account, but suggest an approach based on an overall structural relationship.

[1] Although future contracts tend to be very liquid in European power markets, they usually carry not much information with regard to volatility risk.

Hybrid structural models lie somewhere in between the reduced form and structural models – basically resulting from a trade-off between analytical tractability and the degree of granularity with which specific market characteristics are captured.[2] As opposed to reduced-form approaches that try to grasp how prices move this model class looks beyond prices and asks why prices move in the first place (Eydeland and Wolyniec (2003)). Key fundamental factors driving prices consist of market-wide demand or supply-side variables related to the cost or availability of generation capacities. For instance, a more volatile demand process usually results in larger swings in wholesale spot prices. In electricity markets dominated by renewable generation, such as Norway, Spain, or Germany, weather variables can play a major role for the characteristics of available power generation over time. The stochasticity of such factors can then be captured by well-established reduced-form modeling approaches of the financial literature. Existing hybrid structural models range from slightly altered reduced-form models (e.g. Eydeland and Wolyniec (2003) or Cartea et al. (2009)) to more involved modeling frameworks varying in the number of fundamental factors considered and the kind of information being used for calibration (e.g. Burger et al. (2003), Howison and Coulon (2009), Aid et al. (2013), Füss et al. (2015), or Ziel and Steinert (2016)). Nevertheless, studies taking renewable generation into account are still scarce. Keles et al. (2013) propose a methodology to incorporate wind power and solar generation in a reduced-form model whereas Cludius et al. (2014) include these as an exogenous variable in a structural model. Lehna et al. (2021) use weather data as a proxy for renewable energies and an external regressor. Wagner (2014) develops a residual demand framework and explicitly models the uncertainty of solar and wind generation as additional fundamental factors. Note however, that the above approaches neglect the spatial distribution of the renewable generation portfolio completely.

As hybrid structural models seem to be best suited to match our purposes we pursue a residual demand approach similar to Wagner (2014) in which renewable electricity production is fed into the system with priority. The basic intuition behind such a model lies in the fact that conventional generators satisfy the corresponding (inelastic) demand in hour t which has been adjusted by the uncertain amount of renewable generation re_t in the system.[3] Figure 1 shows the relationship between spot prices during peak and offpeak hours with both demand

[2] As combinations of different layers and algorithms in machine learning approaches or in general combinations between the different modeling categories are referred to as hybrid models, we specifically use the term hybrid structural model.

[3] Because the feed-in of wind/solar is basically free of marginal cost, we assume their generation bids to be accepted on the wholesale market independent from wholesale market prices (which is true, unless very negative prices occur). Please also note that we use load and demand synonymously, following Kiesel and Kusterman (2016).

d_t as well as residual demand $\hat{d}_t = d_t - re_t$. As expected, the figures show that given a level of raw demand d_t there is still considerable variation in wholesale spot prices s_t. This variation (along the y-axis) is clearly reduced if we instead consider residual demand $\hat{d}_t$. Also note that for offpeak hours there are many cases of very low and even negative spot prices whereas demand was not even exceptionally low. A direct comparison with the residual demand - spot price relationship shows that these low prices were in fact caused by exceptionally large production levels from renewables. This shows that in a market with a significant presence of renewables such as Germany a modeling approach for wholesale electricity prices should account for the stochasticity from renewables. This leads us to the first building block of the modeling framework.

Model Component 2.1 *The model for hourly wholesale day-ahead spot prices s_t is*

$$s_t = f_t(\hat{d}_t) + \sigma_t \tag{1}$$

$$\hat{d}_t = d_t - re_t \tag{2}$$

where

f_t corresponds to the supply curve function, and
σ_t is a residual volatility process.

The supply function f_t thus maps the current inelastic demand to a respective hourly day-ahead spot price s_t. This curve basically results from an auction which orders the generators according to their bids. Generators offer their generating capacities at marginal costs and market-wide inelastic residual demand $\hat{d}_t$ determines the intersection with the supply curve and with it the resulting market clearing price. σ_t is an error term accounting for randomness unexplained by the structural modeling framework such as capacity outages or transmission issues.

2.2 Hybrid Structural Price Modeling

Why is localizing the renewable generation important? In a rapidly changing market environment market participants need to be capable of assessing the impact of capacity additions at some given location. Existing approaches may run into difficulties if there has not been any installed renewable capacity at the given location before or a historical track record of (local) renewable production is not available, which is usually the norm. Deducing the im-

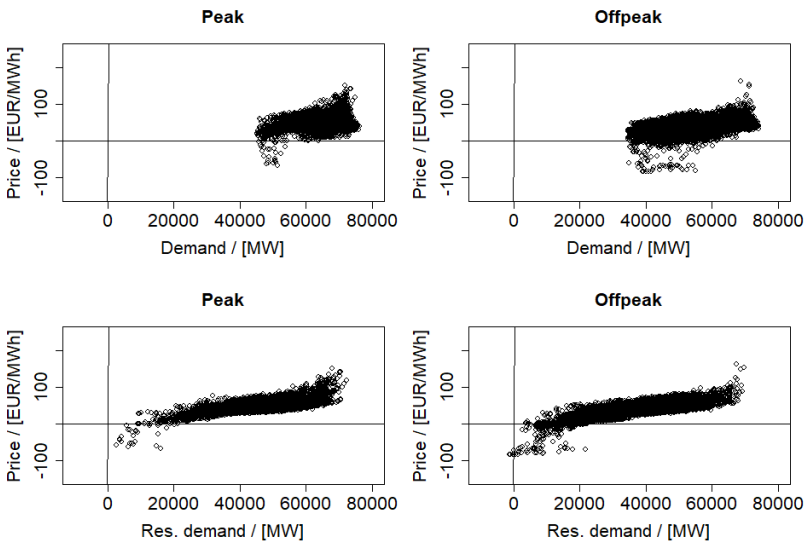

Figure 1: Empirical relationship between demand, residual demand, and spot prices

The figure depicts the relationship between demand and spot prices (top two figures) as well as between residual demand and spot prices for peak and offpeak hours from 01 January 2017 to 31 December 2018. Peakload hours correspond to all hours within 8 am and 8 pm from Monday to Friday whereas offpeak hours correspond to the remaining ones.

pact from the historical (aggregate) renewable generation process can thus be fatally flawed (out-of-sample problem). Alterations to the renewable portfolio might also introduce difficulties if one tries to calibrate a model that solely considers the aggregate generation of the renewable technology. Fitting a single stochastic process to the aggregate wind (solar) power generation then might at least necessitate time-dependent parameters to accommodate for any changes happening over time (in-sample problem). To resolve these problems we do not model the aggregate generation of a specific renewable technology directly but rather focus on its local constituents. This results in the following description of the second main model component:

Model Component 2.2 *The model for the hourly renewable generation process* re_t *is*

$$re_t = \sum_{u \in U} re_t^u \tag{3}$$

$$re_t^u = \sum_{k \in K} g^{u,k}(y_t^{u,k}) \tag{4}$$

where

K is the set of different locations covering the market area,
U corresponds to the set of different renewable technologies in the market,
re_t^u is the technology specific aggregated generation,
$g^{u,k}$ is the production curve mapping weather conditions to output (in MWh), and
$y_t^{u,k}$ is the location- (k) and technology-specific (u) weather variable.

The approach thus incorporates the modeling of the distribution of local weather conditions and their corresponding mapping $g^{u,k}$ to local and with it market-wide generation of a certain renewable technology. In other words, we recursively adopt the basic idea behind hybrid structural modeling approaches on a "deeper" layer of model structure by asking "what drives the drivers of wholesale market prices?". We therefore label the approach "Second-Layer Hybrid Structural model" (SLHS model). Thus, instead of modeling the resulting process of renewable generation of various technologies directly, we instead choose to model the underlying drivers of renewable generation, the weather, and use suitable transformations to map from local weather conditions to aggregate output. Below we show how to estimate these mapping functions from local weather conditions. This allows us to incorporate a long history of weather data.[4]

3 Making Weather Data Useful: Modeling Details and Estimation Strategy

Our model philosophy hinges on the idea of a "smart disentanglement" of major contributors of electricity production from renewable energy sources. These consist of local weather conditions on the one hand and the amount of local renewable generation technology on the other hand. Ongoing research in the field of meteorology has broadened our possibility to better understand the dynamics of key weather variables. Most importantly, one can resort to publicly available comprehensive databases covering highly detailed weather information. Similarly, information with regard to the geographical location of renewable power plants across market areas is usually publicly available with reasonable temporal frequency as well.

Figure 2 showcases publicly available market-wide renewable production data for wind (top) and solar (bottom) for the case of the German power market from 2015 to 2018. Clearly, time series properties are very different for the two technologies considered which proves us right to model power production from wind and solar separately. It is also apparent that renewable

[4] In Appendix A an overview of the different data used in each model component is displayed.

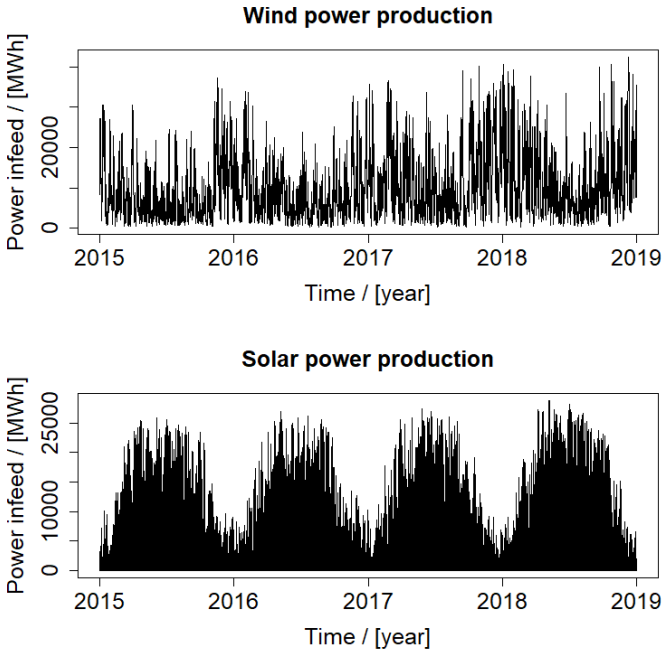

Figure 2: Market-wide renewable production in Germany

The figure shows market-wide renewable generation for wind power (top) as well as solar power (bottom) from 01 January 2015 to 31 December 2018.

production has experienced considerable growth over the last few years indicated by a clear trend in average yearly production volumes. The question is now why one should refrain from modeling these aggregates directly. First, one is left with a relatively short period of data to calibrate model parameters. Second, such a modeling approach is completely blind for any spatial variations in the renewable generation portfolio, potentially necessitating time-dependent parameters. And third, one basically skips a large amount of weather data which could otherwise be of use in the calibration procedure.

In order to bypass these issues and to make use of weather data with high spatial resolution, this section showcases our new empirically-driven approach to estimate the mapping function $g^{u,k}(.)$ from aggregated production data and corresponding weather variables.

3.1 Weather Data

Obtaining a time series of wind speed and solar irradiation with a sufficiently high spatial as well as temporal resolution is usually very difficult. For once, meteorological stations tend to be situated near airports and thus one is usually left with large gaps in geographical

coverage (Rose and Apt (2015)). We therefore opt for reanalysis data that are based on a mix of meteorological observations as well as model-based interpolations in space and time offering a rich history (usually up to several decades) and very high spatial granularity.[5]

For our analysis we use a historical time series of solar irradiation as well as wind speed at 120 m above ground supplied by Anemos (2019). The data is generated through downscaling of reanalysis data from the NASA program Modern-Era Retrospective Analysis for Research and Applications (MERRA) applying the PSU/NCAR (2019). It offers a temporal resolution of 10 minutes (spanning from 1990 to 2018) and a spatial resolution of 20 km x 20 km.[6] We divide the German market area into a grid of 100 km $\times$ 100 km areas resulting in K = 38 weather cells (see Figure 3) for which we compute hourly averages. Our choice of the spatial resolution is mostly motivated by the fact that we want to limit the computational burden in the estimation and scenario generation of our weather model with numerous trajectories later on. Furthermore, several production relevant data is anyway only available on a balance zone level. The choice of the temporal resolution appears reasonable, as most other variables of interest (e.g. demand or day-ahead spot prices) are of hourly frequency as well. Note that our weather grid also entails two offshore regions in the North Sea as well as the Baltic allowing the incorporation of wind power generation at offshore sites.

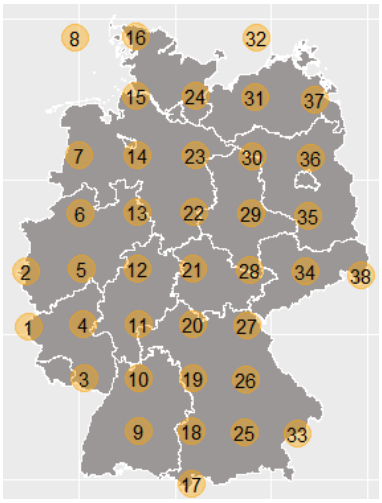

Figure 3: Weather cells of the German market area

The figure shows the used grid of weather cells that covers the German market area.

[5] Although Rose and Apt (2015) have expressed some concerns and demonstrated the presence of a small bias for reanalysis data in the US we have yet to find any better alternative that allows for such high granularity, both in the temporal as well as in the spatial dimension.

[6] A detailed comparison between the model's irradiation data and measurements from weather stations is done by Schermeyer et al. (2014).

3.2 The Renewable Generation Portfolio and Production Curves

The missing link to incorporate the large panel data set of weather variables into our modeling framework now lies in the specification and estimation of the production curve functions $g^{u,k}(.)$. One approach is to set these functions exogenously by making use of production curves. However, such published relationships usually only hold under idealized conditions and numerous studies document substantial bias and deviations from empirical production curves published by manufacturers (e.g. Pieralli et al. (2015) or Ritter et al. (2015)). Furthermore, using such power curves would furthermore require the consideration of wind direction and angle of solar irradiation which would add another layer of complexity to our weather modeling framework.[7] As a result, we opt for a more empirically-driven approach by looking at observable aggregates of renewable generation across several weather cells. We then use these aggregates to calibrate a representative production curve $g^{u,k}(.)$ for the respective candidate weather cells by using the times series of suitably weighted local weather conditions as input variables. Before going into the details of how the production curves are estimated, however, we discuss two key ingredients in our calibration procedure: (1) the dynamics of the spatial configuration of the renewable generation portfolio and (2) the time series of observable aggregates of renewable generation:

(1) Capturing developments in the spatial configuration of renewable power plants over time is an integral part of our modeling approach. EnergyMap (2016) and Bundesnetzagentur (2019a) track the exact amount and geographical locations of renewable generation capacities on a monthly basis. Since there are approx. 2 million renewable energy plants in Germany Bundesnetzagentur (2021), we have to simplify this diversity and aggregate the installed capacity. We allocate the capacity of wind and solar generators to one of our $K = 38$ weather cells according to monthly values from EnergyMap (2016) and Bundesnetzagentur (2019a). To arrive at hourly values we then linearly interpolate between adjacent annually values for installed capacity.[8] For simplicity, we furthermore refrain from accounting for any renewable generation located in Austria as the amount of renewable capacities is very small compared to Germany.[9] This results in a multivariate hourly time series of installed wind

[7] We actually tested a simplified approach using power curves of turbines most often sold and installed in the German market area. However, as expected we ended up with a substantial positive bias in generated output when comparing model-implied with observed values.

[8] Due to the fact that the extraction of geographical locations from EnergyMap (2016) and Bundesnetzagentur (2019a) is relatively time consuming, we started off with a yearly updated dataset for each technology and use linear interpolation in between adjectent years. This should capture potential rapid growth happening throughout the year for the most part.

[9] Note that the German and Austrian power grid actually belonged to one common market area together with Luxembourg until October 2018.

and solar power from 01 January 2017 to 31 December 2018 for each of our weather cells.

(2) Ideally, one would like to use a time series of renewable generation for every location considered to estimate a corresponding production curve. Unfortunately, such data does not exist for most markets and the German one is no exception in this regard.[10] What we do observe is the aggregate renewable generation for four so-called "balancing-areas" though. There are four of such areas which, taken together, form the complete German market area and which incorporate all 38 weather cells in our modeling framework. Figure 4 shows these four balancing-areas. Data is provided by ENTSO-E (2022) and leads to hourly wind and power generation for each of these four balancing-areas from 01 January 2017 to 31 December 2018.[11]

Figure 4: Balancing-areas in the German power market

The figure visualizes the separation of the German market area into four distinct balancing areas (source: Netzentwicklungsplan (2022)): 50Hertz, TenneT, Amprion, as well as TransnetBW.

To finally estimate representative production curves for the renewable power plants we have to incorporate local weather data. We consequently make use of a panel set of hourly wind

[10] Note that there are studies making use of proprietary data sets for specific wind parks (e.g. Ritter et al. (2015)). Nevertheless, even if we had access to several of such data sets this would not be sufficient for our purpose as we require information regarding generated power from wind and solar for each of our weather cells which are unlikely to exist.

[11] Note that we expect to incur some error by allocating renewable generation assets to the nearest weather cell in terms of geographical distance as this does not guarantee that it matches with the affiliation to a respective balancing area.

speed and solar irradiation for each of the $k \in K$ weather cells for the respective time frame. In order to estimate a representative curve for each balancing area we have to specify how the set of input variables (weather) is aggregated when held against the output variable (aggregate renewable production). A straightforward approach could consist in averaging respective weather variables included.

However, this neglects the fact that the amount of installed renewable generation capacity of each weather cell contained within the respective balancing-area is not necessarily identical. The relative contribution to renewable production in a given balancing area tends to be larger for cells with a larger amount of renewable power plants. Obviously, this renders the weather conditions of these cells more important, too. Furthermore, the local amount of installed capacity might change throughout the year and thus result in sudden changes of renewable production within certain weather cells. To cover the above effects we assign weights normalized by the overall installed capacity of a certain technology in a certain balancing-area $TR_n, n = 1, .., 4$ to form representative input variables:

$$\omega_t^{u,k} = \frac{cap_t^{u,k}}{\sum_{k \in TR_n} cap_t^{u,k}}$$

$cap_t^{u,k}$ is the installed capacity (in MW) of technology u in location k at hour t. Using these weights we compute a representative (balancing-area-specific) weather variable for hour t in balancing-area TR_n:

$$z_t^{u,n} = \sum_{k \in TR_n} y_t^{u,k} \omega_t^{u,k}$$

Finally, these variables can be used to estimate the relationship $g^{u,n}(.)$ between balancing area related renewable generation $re_t^{u,n}$ of technology u and $z_t^{u,n}$.[12] Motivated by technology

[12]For the case of wind speed we additionally account for the fact that measurements are taken from 120 m whereas hub heights of turbines might vary. We therefore extrapolate to the average hub height of wind turbines in each weather cell by using the power law which is common practice in literature (e.g. see Brown et al. (1984)):

$$y' = y \left(\frac{z}{h}\right)^{\alpha}$$

where $y'(y)$ corresponds to the wind speed at hubheight z (height of measurement h), and α being the shear coefficient ($\alpha = 0.085$).

specific shapes of production curves indicated by Figure 5, we use a logistic function for the case of wind and a second-order polynomial for solar.[13] We then estimate the respective parameter vector $\Theta^{u,n}$ of technology u of region n by means of the following minimization:

$$\min_{\Theta^{u,n},} \sum_{t \in T} \left(\hat{re}_t^{u,n} - \hat{g}^{u,n} \left(z_t^{u,n}, \Theta^{u,n} \right) \right)^2$$

where both $g(.)^{u,n}$ as well as re_t have been normalized by the sum of total installed capacity $\sum_{k \in TR_n} cap_t^{u,k}$, resulting in hourly efficiency rates $\hat{re}$ and $\hat{g}$. Using these estimated production shapes then allows us to infer local production conditional on local weather conditions and installed capacities.

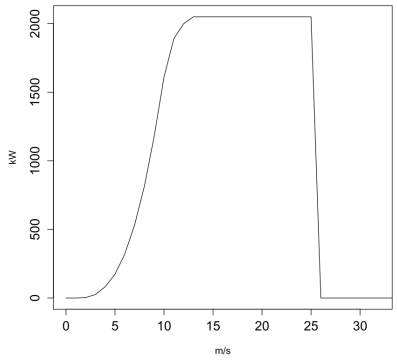

Figure 5: Power curve of exemplary wind turbine

The graph visualizes the dependency of power production and contemporaneous wind speed for the Enercon E82 wind turbine (capacity 2.05 MW, source: Enercon (2022)).

Figure 6 highlights the estimated production curves for wind and solar of the TenneT balancing-area with regard to the normalized generation per 1 MW installed capacity, justifying our choice of a logistic and second-order polynomial function.[14] As indicated by the QQ-plots in Figure 7 our model seems to be quite capable of capturing the distributional properties of balancing-area specific generation from wind. However, our methodology somehow fails to capture the very high peaks in solar production for some balancing-areas. Nevertheless, we obtain correlations of over 94 % for model-implied and observed renewable production (for each of the four balancing-areas), giving indication for the soundness of the model.

[13] Although the power curve of solar is mostly linear in irradiation levels, there is a slight decrease of efficiency for higher levels of irradiation. Solar-panel efficiency is partly reduced by higher temperatures which happens to be correlated with overall irradiation levels.

[14] As expected, the second-order polynomial turns out to be significant and negative for the case of all solar production curves, resulting in deviations from the otherwise linear relationship for large levels of solar irradiation. More details with regard to the parameterization of power curves and estimation results can be found in Appendix B.

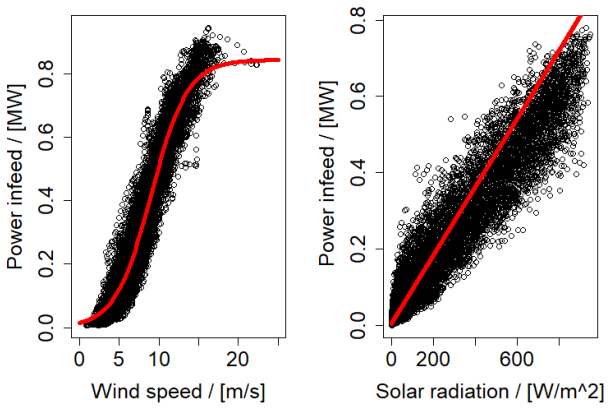

Figure 6: Empirical power curves

The two graphs visualize the dependency between TenneT balancing areas' weather variable $z_t^{u,n}$ and renewable generation for wind (left) as well as solar (right) for 2018. The red line corresponds to the estimated power curves.

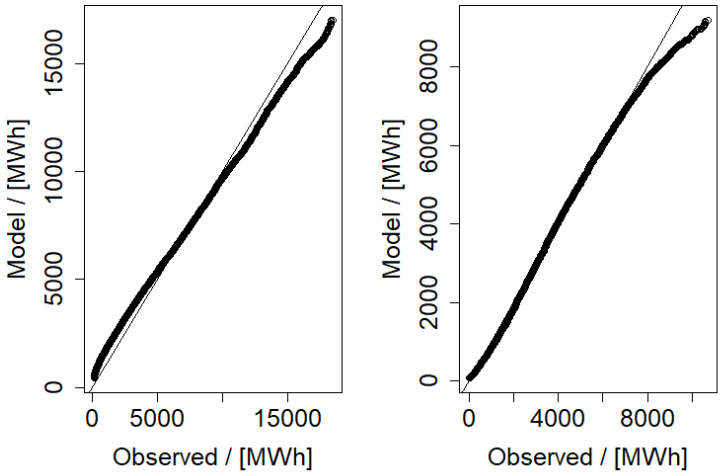

Figure 7: QQ-plots of renewable generation

The two graph depict QQ-plots of observed vs. model-implied renewable generation for wind (left) and solar (right) in the TenneT balancing-area from 01 January 2017 to 31 December 2018.

Overall, our approach to model renewable generation production yields good results in terms of capturing time series properties of observable aggregates. More importantly, a separate modeling of the renewable generation portfolio and production curves from local weather risk makes our approach very flexible. For instance, we can calibrate weather models to a large history of more than two decades of local weather data. Once calibrated, we can also easily change the configuration of the renewable generation portfolio in order to assess its potential impact on market-wide renewable power production and with it on wholesale electricity spot prices.

3.3 Modeling Weather Risk

The last section proposed a methodology that is capable of using local weather conditions as input variables in a framework for market-wide renewable power generation. This effectively enables the researcher to model and calibrate weather risk separately by means of state-of-the art econometric approaches.

Literature with regard to the modeling of the temporal dimension of wind speed and solar irradiation primarily makes use of autoregressive approaches and various extensions of it (e.g. Brown et al. (1984), Mora-Lopez and Sidrach de Cardona (1998), Caporin and Preś (2012) or Alexandridis and Zapranis (2013)). There are furthermore studies focusing on the spatial distribution only. For instance, both Papaefthymiou and Kurowicka (2009) or Hagspiel et al. (2012) make use of copulas in order to capture the local dependencies of wind speed at different locations in Germany. Approaches dealing with both dimensions are less numerous. Morales et al. (2010) and Papavasiliou and Oren (2013) both make use of vector autoregressive (VAR) models to capture the wind speed dynamics at various locations in the United States whereas Grothe and Schnieders (2011) combine autoregressive models with pair-copula constructions (PCC) in a similar setting for German wind speed data. We follow the latter stream of literature and choose a VAR structure for the joint distribution of the multivariate times series of wind speed and solar irradiation.[15] In what follows, we give a short overview of the general idea behind the modeling approach. We then look at stylized statistical features of wind speed and solar irradiation in our data set for Germany and discuss adjustments in the model specifications necessary to capture weather-specific characteristics (e.g. time-dependent volatility) for each case.

[15] Although pair-wise copula constructions offer more flexibility to capture heterogeneity and asymmetries in dependence structures, their estimation and simulation of scenarios is more demanding from a computational point of view. As a result, we stick to the simpler VAR structure and leave an incorporation of copula theory in this regard for further research.

3.3.1 A General Multivariate Weather Model

Modeling weather data necessitates a description of serial and spatial correlation among different locations along with univariate peculiarities such as non-normality. In order to account for seasonal patterns we first remove region-specific trends $\mu_t^{u,k}$ which yields de-trended data $\overline{y}_t^{u,k} := y_t^{u,k} - \mu_t^{u,k}$. To capture non-normality one undertakes a transformation by using the empirical distribution function of the de-trended time series resulting in approximately normally distributed weather variables $\hat{y}_t^{u,k}$:

$$\hat{y}_t^{u,k} = \phi^{-1}\left[F^{u,k}\left[\overline{y}_t^{u,k}\right]\right] \tag{5}$$

where ϕ^{-1} is the inverse cumulative distribution function of a standard normal random variable, and $F^{u,k}$ corresponds to the (empirical cumulative distribution function of the original untransformed (but detrended) time series $\overline{y}_t^{u,k}$. Note that $F^{u,k}$ should correspond closely to the distribution function of the "true" data-generating process since we can resort to a large data set of a long history spanning 28 years of hourly observations.[16] Since the transformation preserves the covariance structure of the weather variables (e.g. see Liu and Der Kiureghian (1986)), we can model their corresponding joint distribution by means of a VAR-model of order P:

$$\hat{Y}_t^u = \sum_{p=1}^{P} \Psi^p \hat{Y}_{t-p}^u + U_t$$

with $\hat{Y}_t^u$ corresponding to a vector of observations of $\hat{y}_t^{u,k}$'s of all regions Ψ^p are coefficient matrices of dimension K x K, and U_t is an error-term following a multivariate normal distribution with a mean vector of zeros and covariance matrix $\sum^u$. The optimal leg-length P is decided upon by using the AIC, remaining appearance of autocorrelation and cross-correlation terms, whereas coefficients along with covariance matrix $\sum^u$ are estimated by maximum likelihood. Henceforth, we are capable of capturing both site-specific pecularities of the marginal distribution and on top of that using the normally distributed transformed data to calibrate a VAR model that allows us to capture serial and spatial correlation.

[16] Some studies impose parametric restrictions on the transform (for example, both Brown et al. (1984) and Morales et al. (2010) use the Weibull distribution for wind speed). However, since we have access to a long history of data it makes more sense to directly make use of the observed empirical distribution function offering more flexibility in terms of capturing site-specific peculiarities. This also allows us to capture potential seasonality in higher moments by making the empirical distribution function time-dependent.

3.3.2 Pecularities of Wind Speed and Solar Irradiation Dynamics

As expected, an analysis of weather variables in Germany reveals distinct seasonal patterns. Wind speed exhibits highly non-normal behavior with volatility being larger during autumn and winter seasons. Skewness and kurtosis varies across the year as well. Furthermore, there is much cross-sectional variation. For instance, wind speed in northern regions is larger on average and more volatile. To capture these aspects we augment the general specification by introducing a time-dependent, site-specific volatility function $\sigma_t^{w,k}$ used as an additional normalization factor besides the trend-function $\mu_t^{w,k}$. We furthermore allow the empirical distribution function $F_t^{w,k}$ to vary across seasons in order to account for time-variation in skewness and kurtosis.

Solar irradiation is behaving quite differently. First, seasonal patterns are much more pronounced: there is no sunshine throughout the night. Furthermore, average solar irradiation levels are ten times as high during summer season. Cross-sectional differences are much smaller if compared to the case of wind though. The absence of sunshine during night complicates the modeling of hourly irradiation levels by means of a VAR approach. Cloud formations have a considerable impact on the resulting solar irradiation and one is incapable of observing these during night time. Consequently, we are blind for any uncertainty that might affect the weather variables during the early morning hours. We therefore deviate from Morales et al. (2010) and loosely follow Wagner (2014) by instead modeling the daily maximum irradiation level $\tilde{y}_t^{s,k}$ and capture any intraday variation by means of a deterministic pattern function.

An inspection of QQ-plots as well as autocorrelation- and crosscorrelation-functions reveals that the augmented modeling approaches are successful at explaining the dynamics of the considered weather variables. More details with regard to stylized characteristics of our weather data, model specifications, and goodness-of-fit tests can be found in Appendix C.

4 Conventional Supply and Demand Factors

4.1 Supply Function

A key ingredient for the model framework is how current market conditions such as demand, renewable generation, and production costs from conventional generation assets are trans-

lated into the market clearing spot price in the day-ahead market. In language of our model framework, this essentially means what kind of structure we impose on the supply function f_t in (1). Existing modeling attempts for f_t range from relatively simple approaches (e.g. Barlow (2002), Burger et al. (2003), and Wagner (2014)) to more involved dynamic frameworks which also account for shifts in the supply function due to variations in different fuel prices (e.g. Howison and Coulon (2009), Coulon et al. (2013)). Since the former approaches have been shown to work quite well and our primary goal lies in a quantitative assessment of weather risks which are mostly unrelated to global fuel prices, we generally follow this stream of literature. $f_t(.)$ is therefore captured by means of a time-dependent deterministic function.

As can be seen in Figure 1 spot prices are non-linear in residual demand as such as they drop disproportionally if $\hat{d}_t$ is low enough and vice versa. The relationship seems to be considerably different when one compares peak to offpeak hours. Additionally, we notice seasonal patterns with regard to the shape, especially within the range of negative day-ahead prices and low residual demand. The noticeable seasonal differences in the shape of the supply curve are supposingly caused by the influence of seasonal fuel prices and seasonal patterns of intra-European import/export of electricity. Addressing mainly the lower part of the supply function, we noticed corresponding remarkable lower coal prices in the first half of the year than in the second half of the year.[17] Imports to the German market zone and exports from the German market zone display seasonal patterns as the German net position (from ENTSO-E (2022)) is mainly negative during summer months and mainly positive during winter months linked to corresponding renewable infeed. As we assume in our modeling framework all renewable infeed produced in Germany to influence directly the residual demand, exported renewable infeed is therefore reducing the residual demand in our model more than according to the actual German merit order, mapping German day-ahead prices therefore with lower residual demands and changing the slope in the lower part according to seasonal pattern. Additionally, the heterogeneous compensation for renewables in 2017 and 2018 as wells as the influence of seasonality in transmission and outage issues could explain further deviations. To capture the behavior, we use the following parametric specification and estimate monthly supply functions $f_{t,m}(.)$ for peak and for offpeak hours:

$$f_{t,m}(\hat{d}_t) = \min(s_{max}, \max(s_{min}, c_{t,m}(\hat{d}_t))$$

[17] For the inspection of coal prices we used the Rotterdam RB Index and the import prices of hard coal from GmbH (2022).

The minimum and maximum wholesale price $s_{min} = -500$ EUR/MWh and $s_{max} = 3000$ EUR/MWh are set exogenously by the EEX. For $c_{t,m}(\hat{d}_t)$ we use the functional form of monthly, cubic smoothing splines for peak and offpeak hours.[18] Smoothing splines are a flexible estimation technique within the category of non-parametric regression models where we can capture a functional relationship with polynoms (splines) while estimating the parameters of this functional shape by minimizing a penalized least square criterion with the smoothing hyperparameter λ (smoothing). The applicable methodology was proposed by Hastie and Tibshirani (1993) and selectively used in the context of electricity prices (e.g. Taylor and Majithia (2000), Sigauke (2017)). The non-negative tuning parameter λ controls and displays the trade off between the smoothness of the functional form and the fit of the curve. We observed in some months (e.g. May and October) a heavier need of the fitting of more complex steps in the merit order, while in other months (e.g. January and July) the supply curve was less complex and had an easily identifiable bid-stack of the underlying offering power plants. Formally, given the respective smoothing parameter λ the used cubic smoothing spline estimator $c_{\lambda,m}$ is the minimizer of a penalized least squares functional:

$$\sum_{t_m}(s_{t_m} - c_{t,m}(\hat{d}_{t_m}))^2 + \lambda_m \int (\frac{d^2 c_{t,m}(\hat{d}_{t_m})}{d\hat{d}_{t_m}^2})dt_m$$

where integration is over all peak ($t_m \in T_m^{peak}$) and offpeak hours ($t_m \in T_m^{offpeak}$) of month $m \in (1, .., 12)$. As the tuning method, we choose ordinary cross-validation ("leave-one-out") after comparing the performance with general cross validation, maximum likelihood and information criteria methods following Berry and Helwig (2021).[19]

Figure 8 shows our estimates for offpeak hours in January and peak hours in October whereas Table 1 provides information about the tuning parameters and fitting criterions. We observe considerable differences in the shape for different time frames. It seems that in contrast to earlier studies differences between peak and offpeak hours as well as between months have increased considerably (e.g. Burger et al. (2003)) and at low residual demand levels the functional relationship is significantly steeper. Additional information on the monthly, cubic smoothing splines can be found in Appendix D.

Since our model does not explicitly account for all fundamental factors (e.g. power plant

[18] Similiar to this methodology is the approach from Jenkin et al. (2018) using quarterly linear and cubic regressions.

[19] The other tuning mehtods from Berry and Helwig (2021) showed in the monthly cubic smoothing splines in one month or more a lower adjusted R^2.

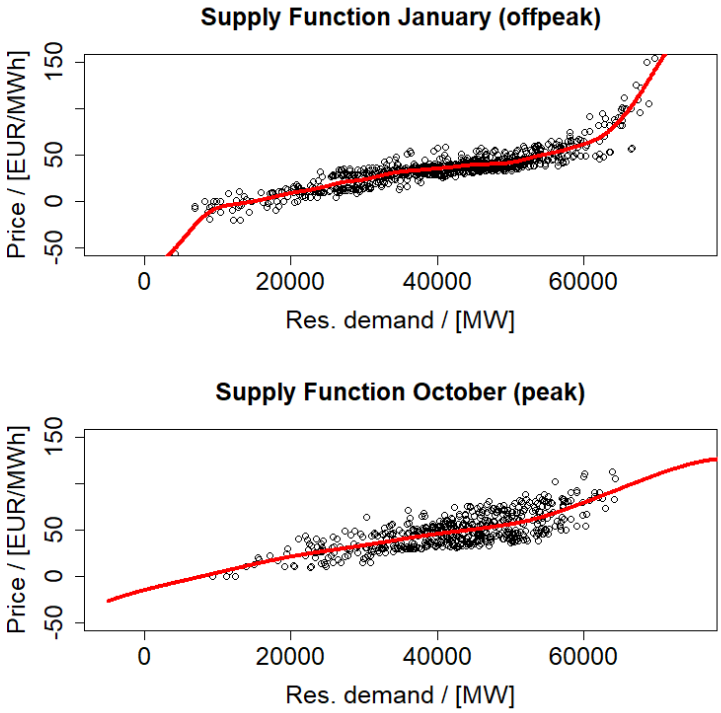

Figure 8: Supply Curve Estimation

The figure shows the estimated supply curves for offpeak hours in January and peak hours in October (in red).

	df	λ	Adj. R^2
Jan. (offpeak)	18.65	$1.15 \cdot 10^{-7}$	0.88
Oct. (peak)	6.22	$1.51 \cdot 10^{-5}$	0.56

Table 1: Tuning parameters and fitting criterions of the exemplary supply curves

The table shows tuning parameters λ and derived degrees of freedom as well as the adjusted R^2 as a fitting criterion for the smoothing splines as estimated supply curves of offpeak hours in January and peak hours in October.

outages, transmission restrictions) or other aspects (e.g. market psychology) an inspection of the residual process $\sigma_t := f(d_t) - s_t$ seems warranted. Very similar to Burger et al. (2003) we find that this process is mostly unrelated to fundamentals such as weather variables or demand.[20] In order to make use of the model in risk management applications such as hedging or Value-at-Risk calculations it makes sense to capture this additional uncertainty as well. For instance, if weather-related risk such as wind power generation at a specific location is to be hedged with spot price derivatives, neglecting this additional (independent) source of risk would potentially bias hedging efficiencies. We therefore follow the rather practical approach by Burger et al. (2003) and model the residuals σ_t by means of a parsimonious time series model. The inspection of the σ_t reveals both serial correlation at various lags as well as heteroscedasticity prompting us to choose an ARIMA process with GARCH noise.[21]

4.2 Demand

Demand can be seen as the major driver of price changes. It exhibits pronounced seasonal patterns on a yearly, weekly, as well as intra-daily basis with additional holiday-effects.[22] Modeling demand has been investigated in numerous studies (e.g. Weron (2006), Burger et al. (2003), Howison and Coulon (2009) or Wagner (2014)) and is quite well understood. It has been shown that besides the need for a flexible trend function in order to capture the pronounced seasonality patterns demand also requires to capture serial dependencies justifying the use of ARIMA-type modeling approaches. We once again follow Burger et al. (2003) and use an ARIMA-type model to capture auto-correlation and a deterministic trend function with indicator variables for months, hours, weekends, holidays and winter season based on hourly demand data from EEX (2022). As in the aforementioned studies, the model is capable of capturing the characteristics of temperature-driven market-wide demand.

5 Discussion of Model Extensions

Our parsimonious structural modeling framework obviously carries its limitations. We will therefore shortly look at its most severe drawbacks, how one could approach these, and in which cases these might be of minor importance.

[20] All time series correlations are below 20 % (in absolute terms).

[21] A more detailed discussion of the model for the residual volatility process σ_t can be found in Appendix E.

[22] Please refer to Appendix F for a more detailed discussion of the modeling details and statistical properties of market-wide demand in Germany.

Obviously, the deterministic supply function could be modeled in a more involved fashion. First, the data-driven cubic smoothing spline estimation leads to several supply functions that are not monotonous as the underlying hourly merit order curve would economically suggest.[23]. Therefore, further monotony conditions should be implemented to maintain the logical bid-stack order. Second, if variations in the supply function are modeled with additional, explicit model components, the more data-driven methodology of using monthly supply curves could be changed to one common supply curve. Variations in the shape of the curve are overall caused by three factors: (1) import/export of electricity, (2) power plant availability, and (3) power generation costs.

In our modeling framework we assume all fundamental factors to influence only the German market environment. With the growing integration of a European electricity market, not all German renewable generation is satisfying only German demand and other European power plants with their respective marginal costs can be price-setting during a period of supranational price convergence. As the market coupling of multiple market zones is a complex modeling attempt, a first step of more accurately measuring the price inelastic residual demand for a common supply curve methodology could be corrected with the German net position Δ: $\hat{d} = d - re_t + \Delta$. Without detailed consideration of individual interconnected relationships, the residual demand would therefore be higher in case of import ($\Delta > 0$) and lower in the case of exports ($\Delta < 0$). Technically, modeling the net position could be performed comparable to demand based on an SARIMA model.

Unexpected technical issues can result in sudden shutdowns of conventional generation assets resulting in potential changes in the shape of the supply curve. Note that such impacts might be very different conditional on which kind of asset is affected by the outage. For instance, an outage of a large nuclear baseload plant with low marginal costs usually results in a shift of the whole curve to the left, affecting the price formation for all hours within a day. In contrast, an outage of a flexible smaller natural gas power plant is more likely to make the supply curve steeper in the right part of the supply function, thereby having a stronger impact on prices during peak hours. Given detailed data on "default rates" of the generation fleet of conventional assets one could model the aggregate plant availability over time and construct the supply function based on available assets at a given point in time. This would then render the supply function stochastic.

As power plant owners usually offer to produce electricity for prices close to their marginal costs, fuel prices are an important driver of time variation and seasonality in the supply

[23] Figure 29 and Figure 30 in Appendix D visualize all the estimated monthly supply curves.

function. Note that different parts of the supply function are driven by different fuel costs as well. At the moment, we capture any seasonal monthly variation by a different estimated shape of the supply function and any further uncovered variation by a residual volatility process σ_t. The residual volatility process is assumed to be stationary and therefore does not cover any long-term risks such as permanent changes in fuel prices. A geopolitical event (e.g. pandemic, military conflict) could shift part of the supply function up- or downwards and result in a long-term average price shift. In a similar way, fluctuating CO2 emission prices could have an impact on the supply function as well.[24]

The above extensions for a common supply function necessitate a lot of highly detailed data on the conventional generation portfolio which might not always be available. Although long-term risks in terms of changes in fuel prices are not captured, the residual volatility component should at least be capable of covering short-term deviations attributable to outages to some extent. We therefore regard the limitations primarily relevant for market participants concerned with long-term prospects of market dynamics. An ad-hoc approach to accommodate for a current permanent shift in the supply function could be partly captured by an additional parameter in the model component (1):

$$s_t = f_t(\hat{d}_t) + \sigma_t + \delta$$

If permanent, such a shock should be reflected in traded forward-looking instruments. As a result, δ could then be estimated by means of long-maturity electricity futures, somewhat similar to the practice of yield curve fitting for spot-price models of the fixed-income literature. The price of such a contract is its expected value under the risk neutral measure:

$$\begin{aligned} f_t^e &= \mathbb{E}_t^{\mathbb{Q}} \left[\sum_{t \in T} s_t \right] \\ &= \mathbb{E}_t^{\mathbb{Q}} \left[\sum_{t \in T} f_t(\hat{d}_t) + \sigma_t + \delta \right] \\ &= \mathbb{E}_t^{\mathbb{Q}} \left[\sum_{t \in T} f_t(\hat{d}_t) \right] + \hat{\delta} \end{aligned}$$

The second equality follows from the fact that the residual volatility process $\sum_{t \in T} \sigma_t$ is zero in expectation and independent from other sources of considered risk factors. Using a

[24] As CO2 emissions vary across electricity generation technologies, the impact of changes in CO2 prices is not the same for different parts of the supply function.

calibrated version of the model could in turn be used to match observed futures prices by adjusting $\hat{\delta} = \sum_{t \in T} \delta$ accordingly. Nevertheless, care has to be taken in the estimation of δ. First, futures contracts incorporate risk premia which need to be estimated beforehand (e.g. based on historical fundamental data and quotes on futures). Second, one has to make sure that the shift in average prices is not attributable to forward-looking changes in fundamentals on the demand- or supply-side.

Other overlooked aspects in the model are wind direction, air density, surface roughness and variations in irradiation angle. The direction a physical asset faces can have a considerable impact on the resulting power production both for wind as well as for solar. For wind power this is much less of a concern. This is due to the fact that so-called horizontal axis turbines, which represent the vast majority of commercial assets nowadays, can quickly adjust their rotor blades to variations in wind direction.[25]

In contrast, the majority of commercial solar power plants are unable to adjust their orientation towards incoming irradiation which changes throughout the day. Therefore, a solar power plant's individual configuration influences the resulting power generation. An extension of our model could incorporate information about each asset's configuration within a weather cell. This could be captured by adjusting the local capacity weights when estimating the balancing-area-specific production curve $g_t^{s,n}$ (see section 3.2) according to the respective local orientations of solar power plants and be of help to capture local renewable generation dynamics more realistically.[26] Furthermore, this would allow market participants to consider an additional strategic layer in their decision making process.

6 Empirial Analysis

Using the calibrated SLHS model, we now address its performance in terms of reproducing wholesale power price dynamics. We then look into how our model can be applied in practice and in which cases it provides unique benefits for different stakeholders in the electricity sector.

[25] Wind direction as well as air density and surface roughness can become important for the detailed analysis of the power production profile of a specific wind park though (e.g. the exact placement of turbines next to each other and/or the consideration of obstacles such as hills). Given the spatial granularity of our weather model (weather cells of 100 km $\times$ 100 km size) such a level of detail is beyond the scope of this study.

[26] We actually attribute some of the unexplained variation by our solar production curves (see section 3.2 Figure 6) to the neglection of solar panel orientation.

In the following analysis, we make use of simulations to compare statistical properties such as observed and model-implied moments or quantiles. We found $N = 1000$ simulation scenarios to be reasonably robust in terms of the sampling error. To obtain one trajectory of wholesale power prices we have to simulate all state variables: (1) weather variables, (2) market-wide demand, and (3) residual volatility. Our analysis also entails assessments of specific time frames in isolation (e.g. a specific month). If not stated otherwise, we create simulation scenarios in these cases by conditioning all state variables to equal their mean values at the beginning of the month.

6.1 Explaining Wholesale Power Prices

To underline the soundness of the proposed modeling framework we now shed light on its capability of reproducing salient statistical features of wholesale spot prices in the German electricity market. We start with a visual inspection of historical as well as simulated day-ahead spot prices. Figure 9 is an exemplary selected plot of s_t from 01 October 2018 to 31 December 2018. It can be seen that the stylized features of observed day-ahead prices seem to be captured quite well. For instance, negative price spikes tend to occur more frequently during winter holidays when the demand from large industrial consumers is missing.[27]

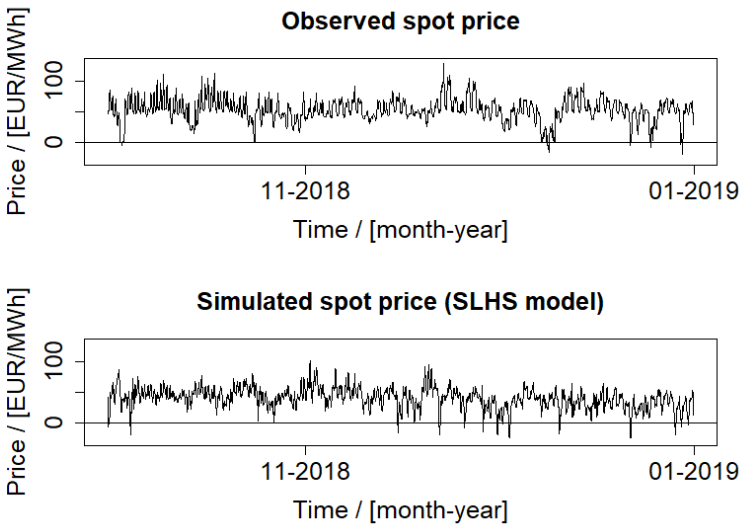

Figure 9: Observed vs. simulated price trajectories
The figure depicts observed (top) and simulated (bottom) day-ahead spot prices for the last quarter of 2018.

[27] Further details considering the model's capability of reproducing a comparable time series can be found in Appendix H Figure 38.

Next, we look at QQ-plots of model-implied and observed spot prices. As a benchmark, we choose a standard reduced-form price model for electricity price dynamics used in Benth et al. (2013) and calibrate it to the same time series of wholesale spot prices used for the SLHS model. The model features a flexible trend function, a mean-reversion component, as well as a spike process to capture heavy tails in the empirical distribution.[28] Since reduced-form approaches are calibrated to the observed price distribution directly they serve well as a benchmark for our modeling framework. Figure 10 highlights the QQ-plots between the observed price distribution and the corresponding modeling approaches for the time between 01 January 2017 to 31 December 2018. Both seem to be capable of capturing the majority of the empirical distribution although the extreme tails are not captured perfectly. The SLHS model has certain problems with estimating very negative spot prices, while the reduced form model overestimates moderately positive spot prices. This demonstrates that the SLHS model can compete with existing reduced-form approaches.[29]

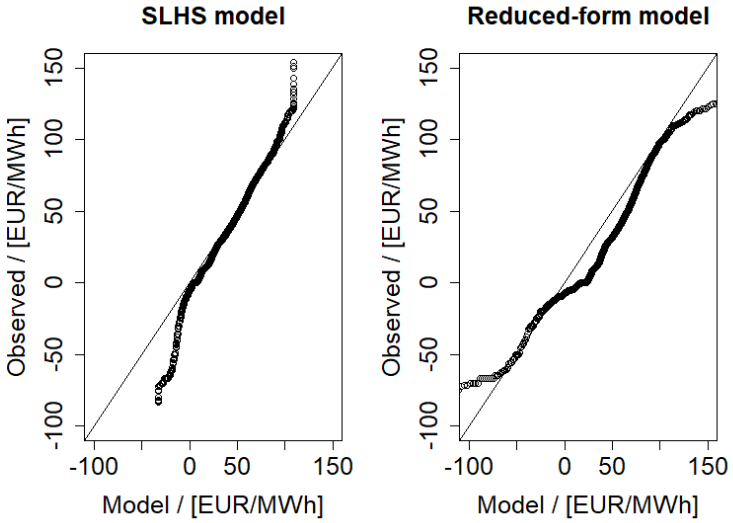

Figure 10: QQ-plot between observed and simulated day-ahead prices
The figure shows QQ-plots between observed and simulated spot prices using the SLHS model approach (left) as well as using the reduced-form model (right) between 01 January 2017 and 31 December 2018.

Next, we have a look at how the model fares with capturing key properties of price volatility. Volatility is a key factor for important managerial decisions or risk management applications. For example, the scheduling of electricity production of highly flexible power plants is a complex path-dependent dynamic optimization problem. The asset derives much of its value

[28] Please refer to Appendix G for a description of model specification and calibration.

[29] Please note that we used winsorizing at the 99.9 % / 0.1 % quantile for the simulated price distribution (1000 trajectories) of the SLHS model.

from price volatility. Consequently, the decision to ramp up such an asset should then be based on the most relevant observable state variables influencing future volatility. As a result, we require a model in which the dependence between volatility and other variables of interest is captured accordingly.

To demonstrate the SLHS model's capability in this regard we look at the relationship between residual demand $\hat{d}_t$ and price volatility. To set the corresponding variables into perspective, we calculate daily average values for both variables of interest and inspect the relationship visually (see Figure 11).[30] Although the pattern is not captured perfectly, the model correctly predicts higher volatility levels for higher residual demand levels. We also observe a reversal of this effect for very low levels of residual demand both for historical data as well as model-implied. This pattern can be attributed to the fact that for very low and high levels of residual demand $\hat{d}_t$ the supply function is relatively steep (see Figure 8), potentially causing larger price swings. Note that this effect is present in other electricity markets as well (e.g. see Eydeland and Wolyniec (2003)) and called "inverse leverage effect".

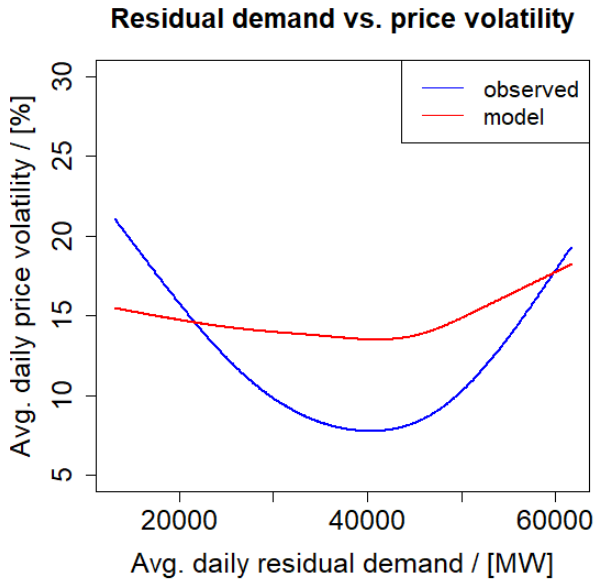

Figure 11: Relationship between residual demand and price volatility

The figure shows the relationship between average daily residual demand $\hat{d}_t$ and average daily price volatility for historical data (blue line) and model-implied (red line).

Finally, we test the model's capability to describe the joint distribution of the modeled (aggregate) drivers and the wholesale spot prices - a property that becomes very important

[30] We actually depict splines fitted to the empirically observed relationships to facilitate the presentation. Further details considering the used splines can be found in Appendix H Table 6.

for hedging practices. For instance, the owner of a wind park might be interested in using price-based derivative instruments (e.g. electricity futures) to reduce the volatility of his future cash flows. In order to make an assessment of the potential of such strategies he necessitates a model that correctly predicts the relationship between prices and wind power production. Table 2 depicts the time series correlations of renewable generation from solar, wind, demand, and residual demand with spot prices. As can be seen, the model-implied correlations are very close to what we actually observe.[31] For instance, the link between residual demand and spot prices is stronger than for the case of raw demand and wind seems to have a stronger impact on spot prices than solar. Additional evaluation of the model performance can be found in Appendix H.

	model	data
ρ_{s_t,re_t^w}	-0.46	-0.46
ρ_{s_t,re_t^s}	-0.09	-0.04
ρ_{s_t,d_t}	0.45	0.47
$\rho_{s_t,\hat{d}_t}$	0.79	0.77

Table 2: Correlations between spot prices and fundamental factors

This table compares correlations for several fundamental factors (demand d_t, residual demand $\hat{d}_t$, wind generation re_t^w as well as solar generation re_t^s) with wholesale spot prices s_t for the model (left column) as well as the historical data (right column) from 01 January 2017 to 31 December 2018. The model-implied correlation is calculated as the mean of the pearson correlations at each hour of all simulated trajectories.

6.2 Managing Market Risks with the SLHS Model

Market participants in modern power markets are facing increasingly complex weather-dependent uncertainties. This section highlights how the SLHS model can help to better understand the associated risks.

6.2.1 Setup of Risk Analysis

We take the renewable portfolio configuration from the start of 2017 as our base scenario and hold the installed capacity constant at first. This assumes approx. 55 (37) GW of installed

[31] We are well aware that correlations can be misleading for non-linear relationships. Nevertheless, the similarity of both empirical and model-implied moments is striking.

wind (solar) power.[32] We make use of simulations to deduce corresponding measures of interest and set the number of scenarios to N = 1000. One trajectory of wholesale power prices necessitates the simulation of all state variables: (1) weather variables, (2) market-wide demand, and (3) residual volatility. In some cases, we analyze different time frames (e.g. months or years) in isolation. If not stated otherwise, we then compute the corresponding measure by conditioning the state variables to equal their mean values at the beginning of the respective time interval.

6.2.2 The Impact of Renewable Generation

As outlined in section 3 the renewable generation portfolio is in constant change. Market participants need to understand how capacity additions impact their current commercial operations. This subsection will showcase two examples of important market players and how these might be affected.

Our first example consists of a merchant power plant which is not tied to any customer needs or long-term power purchase agreements. If one abstracts from technical restrictions as well as other fixed costs a conventional power plant can be seen as a strip of call options written on wholesale spot prices. A stylized profit margin of such a physical asset is given as follows:

$$rv_T^{pp}(c) = \sum_{t \in T} \max(s_t - c)^+ \tag{6}$$

where T corresponds to the set of all hours within the respective time interval and c to the variable cost required to produce the equivalent of 1 MWh of electricity. The variable costs are determined by the type of power plant considered.[33] We focus on flexible peaking power plants that usually burn natural gas in order to produce electricity. Using a corresponding efficiency rate and fuel costs results in variable costs c of about 60 EUR/MWh.[34] Given the

[32] Please note that this differs from section 6.1 where we reflect on our model's capability of reproducing the observed electricity spot prices. By holding the installed capacity constant, we neglect the further real expansion of installed renewable capacity during 2017 and accept a moderate deviation to the observed spot prices.

[33] In order for Eq.(6) to hold the power plant needs to be flexible enough such that it can be switched on and off with very short notice. Since coal and lignite power plants often times require several hours or even days to ramp up and down we focus on more flexible gas power plants for this example. Additionally, we assume that the peaking power plant is not the price-setting power plant in the day-ahead auction, as otherwise the call options would never have a positive value.

[34] Efficiency rates of such physical assets usually range between about 20 - 30% (see Eydeland and Wolyniec (2003)), thus for every MWh of electricity one necessitates the equivalent of 3 to 5 MWh of naturals gas. Prices of natural gas in Europe have not changed considerably during the years 2017 and 2018 and averaged

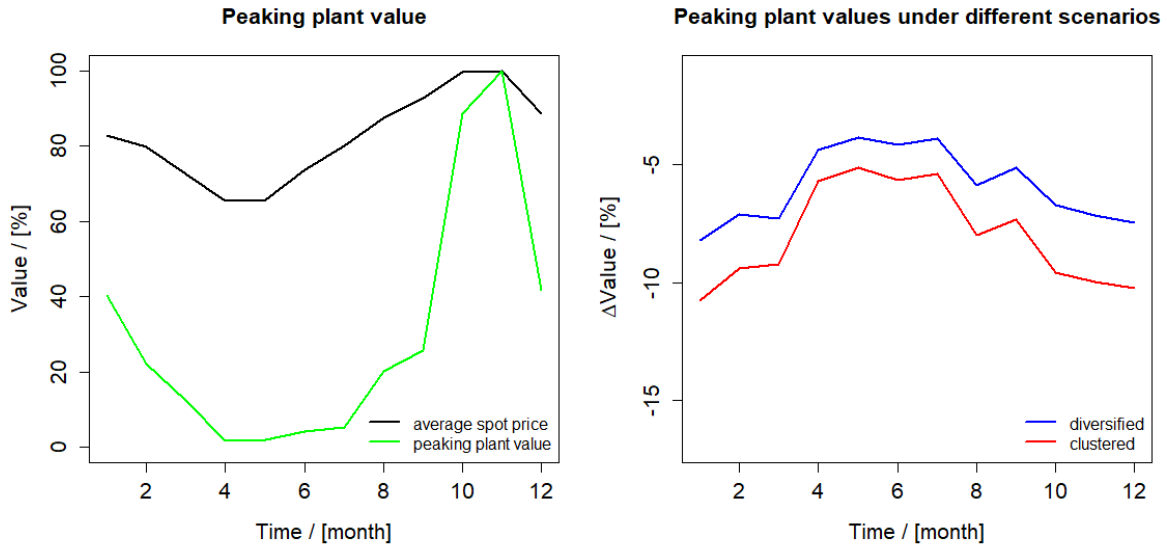

Figure 12: Peaking power plant

The left graph shows values of v_{t_0} of the peaking power plant as well as the average (observed) spot price for different months over the year (2017), in which both time series have been normalized by their maximum value. The right graph shows the percentage reduction in the peaking plant's value in different scenarios over the year.

fact that c is almost twice the price of average spot prices demonstrates that such an asset can thus be basically regarded as a strip of deep out-of-the money call option contracts.

In order to value the basket of option contracts one needs to calculate their expected value under the pricing measure $\mathbb{Q}$:

$$v_{t_0}(c) = \mathbb{E}^{\mathbb{Q}}_{t_0}[rv^{pp}_T(c)] \tag{7}$$

We ignore risk premia and compute the above expectation under the physical measure $\mathbb{P}$. Given the absence of a closed-form solution for (7) we resort to simulations and assume a risk free rate of $r_f = 0.2$ % to calculate a discounted value at t_0 of future profits. Option values are heavily driven by higher order price risks and our SLHS model allows us to price these risks accordingly.

The left graph of Figure 12 shows v_{t_0} as well as the average electricity price level for all months during a year. Clearly, there are seasonal patterns in both cases. The average price level is larger during autumn and winter season due to higher demand levels (see section

at about 20 EUR/MWh which is why we assume them to be constant for our analysis.

4.2) and lower supply in terms of solar power production (see section 3.3.1). The lower level in December is comparable to what we actually observe in 2017. However, in contrast to the average price level, the value of the peaking power plant rises approx. tenfold in November. This aspect can be explained by the non-linear shape of the supply function f_t. During autumn and winter, the intersection of supply and demand is much more likely to take place in a steeper part of f_t (see section 4.1 Figure 8). The chances of positive price spikes therefore increases disproportionally. This increase of (positive) jump risk then has a large impact on the value of the portfolio of deep out-of-the money call option contracts (the peaking plant v_{t_0}).[35] This explains the much larger discrepancies between values of v_{t_0} during autumn/winter and spring/summer season compared to average price levels.[36]

Now suppose there is renewable capacity added to the supply-side instantly at the beginning of the period. Given the heterogeneity in dependencies of local weather variables as well as the non-linearity of local production curves $g^{u,k}$, it is a non-trivial task to quantify how such capacity additions influence aggregate renewable generation. Existing approaches (Wagner (2014)) can at best account for a proportional growth of renewable capacities at all locations such that the relative contribution of all local constituents remains exactly the same. Of course, this is highly unrealistic. Our approach on the other hand captures these local aspects. The following example demonstrates their further usage.

To outline the varying impact of new renewable capacity additions we consider two scenarios in which a total amount of 3.9 GW wind capacity is installed (2.2 GW wind offshore, 1.7 GW wind onshore). This corresponds to the average annual growth rate needed to achieve the targets of renewable installed capacity for wind of the German government from 2020 to 2030.[37] We consider two schemes to allocate those across the K = 38 weather cells. The first one corresponds to an equally distributed scheme where each location is allocated an equal amount of wind power (Scenario Diversified). For the second scenario we assume that all capacity is clustered in the weather cells with the highest average corresponding wind speed (Scenario Clustered). We choose to illustrate an expansion of wind capacity because based on the weather variables the PV generation is not as focused on single weather cells as the

[35] In relation to the price of futures contracts ("average price level") or prices of at-the-money options, out-of-the money options prices increase disproportionally in value if jump risk increases, since large jumps are usually the only events that causes these contracts to end up in-the-money.

[36] Note that some existing structural models (e.g. Wagner (2014)) would have allowed to make a similar assessment. Keep in mind though, that we make use of a much longer time series of weather data, whereas e.g. Wagner (2014) is restricted to the few years of renewable production data currently available. This lack of data might result in less reliable parameter estimates for the market-wide renewable processes.

[37] § 4 sec. 1 EEG is codifying the target of 71 GW for wind onshore and § 1 sec. 2 WindSeeG 40 GW for wind offshore.

wind generation. However, all individual expansion scenarios in both technologies using our assumed weather cell grid can be modeled and quantitative implications derived.

Using our SLHS model, we assess changes in the peaking plant value due to changes in the renewable generation portfolio. We find that in both cases, price volatility rises but as can be seen from Table 3 there is a considerable negative impact on the value of the power plant ranging from -6.87 % to -9.49 %. The reason for this is the fact that in both scenarios, the probability of higher price states is decreased and with it the profit margin of the peaking plant. The right graph in Figure 12 again visualizes the impact on the monthly (average) values of the physical asset. As can be seen, the influence is felt the most during winter season. More importantly however, there is a difference between the two scenarios. Due to the fact that Scenario B situates new wind power capacities in a more windy weather cell there is an even larger discount on the value of our basket of out-of-the-money options compared to the equally weighted capacity addition case. For the yearly value of an average sized peaking power plant (e.g. 500 MW) this results in a economically sizable difference in value of 500 MW $\times$ (1.12-1.09) EUR/MWh $\times$ 8760 h = 131,400 EUR. This shows that changes in the spatial distribution of the wind power generation portfolio have an economically important impact on conventional physical assets. Most importantly, this impact could not have been analyzed quantitatively by the help of existing reduced-form (e.g. Benth et al. (2013)) or structural models (e.g. Wagner (2014)).

	base	scenario A	scenario B
$v_{t_0}(c = 60)$	1.20	1.12	1.09
% *increase*	-	-6.87 %	-9.49 %

Table 3: Valuation of peaking power plant under different scenarios

This table summarizes annual values and relative changes of a stylized natural gas-fired power plant (1 MW) under three different scenarios. The base scenario corresponds to the market environment of 2017 whereas scenario A and B correspond to an equally weighted expansion and a clustered expansion respectively. The second line is the relative change in the respective scenario compared to the base case.

Next, we look at the risk electricity suppliers face. In most electricity markets the majority of consumers enter some kind of load-serving contract in which the exact quantity of power is left unspecified while a fixed price for every consumed unit is set ex-ante. While very appealing for risk-averse consumers, this can potentially result in very risky non-linear exposures for the other party. These so-called Load-Serving Entities (LSEs) usually manage a large portfolio of such customers. A simplified version of the associated revenue stream arising from such a commitment is given as follows:

$$rv_T^{lse}(p) = \sum_{t \in T} \hat{q}_t (p - s_t) \tag{8}$$

where $\hat{q}_t$ corresponds to the consumer-specific demand in hour t and p is the contracted fixed price charged from the LSE for every quantity of electricity consumed. For electricity markets, quantity and price variables are usually positively related, resulting in non-linear payoff patterns. As $\hat{q}_t$ is not traded in the marketplace, contracts of this type are rather difficult to hedge.[38] To visualize the problem, we now look at a simplified yearly load-serving contract for the German power market using our calibrated price model. For simplicity, we ignore the idiosyncratic part of the consumer demand and furthermore assume it to be perfectly correlated with market-wide power demand d_t. We furthermore set the fixed price p such that the contract has a value of zero at initiation ($\mathbb{E}_{t_0}^{\mathbb{Q}} rv_T^{lse}(p) = 0$). Given the concave payoff structure of the load-serving contract and assuming again $r_f = 0.2$ %, the fixed price is larger than the average electricity spot price ($\mathbb{E}_{t_0}^{\mathbb{Q}} s_t = 44.27 < p = 45.50$ EUR/MWh).

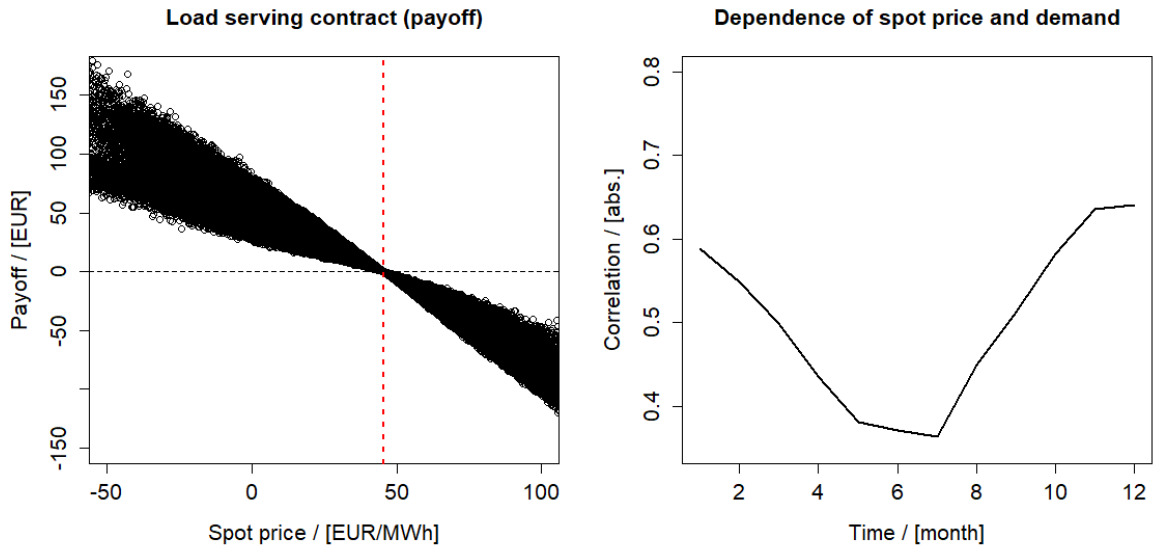

Figure 13: Load-serving contract

The left graph shows the (hourly) payoff from the load-serving contract ($\hat{q}_t(p - s_t)$) relative to the prevailing spot price s_t. The red vertical line corresponds to the contracted price p. The right graph shows the correlation between spot prices s_t and market-wide demand d_t in every month of the year. Both graphs use simulated spot prices and demand from the calibrated SLHS model.

The left graph in Figure 13 shows the relationship between hourly payoffs and the spot price

[38] In other commodity markets, the supplying company can at least build buffer stocks by storing an adequate amount of the underlying good physically. Unfortunately, electricity is non-storable making the problem even more difficult.

in the load-serving contract. The non-linear pattern is clearly visible: consumers tend to ask less for lower price states and ask for more in higher price states.[39]

Before analyzing the impact of potential changes in the renewable generation portfolio, it is insightful to shed some light on how renewables affect the LSE's business. The right graph in Figure 13 shows the monthly correlations of spot prices s_t and market-wide demand d_t (which is assumed to be perfectly correlated with the above customer demand $\hat{q}_t$).[40] As can be seen, there is a pronounced seasonal pattern with a significant drop of dependencies arising throughout the summer months. This shows how the increasing presence of renewables introduces new sources of risks into wholesale market prices. As shown in section 3.3 solar irradiation is at extremely low levels during the winter season making wind power a much more important contributor to market-wide power generation. During summer, solar power production rises about ten-fold whereas wind production drops by about 30% - 50%. Overall, aggregated market-wide renewable generation is much higher and more volatile during summer. As a result, the link between market-wide demand d_t and the resulting spot price s_t is decreased. This has important implications for the LSE. First, if the dependencies between spot prices and the uncertain quantity variable change so does the amount of the non-linear exposure in the payoff function, potentially even decreasing the risk of extreme losses for the LSE. On the other hand, a weaker link between wholesale prices and demand means that the hedging efficiency of electricity futures is lower in some periods during the year. LSEs are consequently more and more affected by the growing share of weather-dependent electricity production.

Suppose now, it is publicly known that new renewable capacities are added to the generation portfolio next year (as was the case for the peaking power plant example). The consumer now wants to re-assess whether the fixed price p he is charged is still "fair" in this regard. To do so he can compute p such that $\mathbb{E}_{t_0}^{\mathbb{Q}} rv_T^{lse}(p) = 0$ holds. The fixed price under scenario A (diversified wind power expansion) is larger than under scenario B (clustered wind power expansion). We obtain $p \approx 44.67\ EUR/MWh$ for Scenario Diversified and $p \approx 44.34\ EUR/MWh$ for Scenario Clustered. So, if one is not charged accordingly this can easily result in large discrepancies in the resulting electricity costs. Even for just a medium-sized commercial business (e.g. 150 MW per hour) the difference is economically

[39] Although not central to our analysis, this demonstrates the hedging dilemma the LSE faces. A linear hedge using futures contracts will consequently not be capable of completely protecting the company from adverse payoffs in high- and low price states. Unfortunately, non-linear instruments, such as options, are not liquid. As a result, LSEs usually adjust the fixed price to compensate for the unhedgable risks they bear.

[40] To arrive at a representative correlation coefficient, we take the average correlation across all $N = 1000$ simulated scenarios of a month.

significant with $150\ MW \cdot (44.67 - 44.34)\frac{EUR}{MWh} \cdot 8760\,h = 433,620.00\,EUR$, which is even larger than for the case of the power plant outlined in the last section as well.

Overall, the above two examples demonstrate the impact of the steady increasing amount of renewables on commercial activities of market participants. It is shown that even if one is not directly invested in local renewable production assets it can be important to incorporate the spatial distribution of renewable generation capacities within a power price model.

7 Conclusion

The intermittent nature of renewable electricity production has changed the landscape of many liberalized power markets. New weather related risks arise for the stakeholders. Aggregate (market-wide) renewable production is itself determined by the sum of local (renewable) production facilities across a given market area. As local output is highly sensitive to local weather conditions, the spatial distribution of installed renewable generation capacities is therefore a potentially interesting characteristic that should be captured by a meaningful risk management tool. We propose a flexible modeling framework for wholesale power prices capable of incorporating the local aspect of renewables for the case of the German power market. We recursively adopt the basic idea behind hybrid structural models by not only looking at the drivers of wholesale prices (e.g. temperature driven demand, aggregate solar or wind production) but also at the (local) drivers of renewable production itself. In a nutshell, we disentangle the modeling of weather conditions, the mapping to (local) renewable production, and the amount of (local) installed capacity. This is advantageous since it allows us to calibrate part of our model to a rich history of weather data instead of having to rely on the (relatively) short time frame of renewable production. Our results show that the SLHS model is well capable of reproducing the statistical properties in the time series of renewable production and wholesale power prices and can consequently be regarded as a valid tool for risk management purposes. We furthermore outline how market participants can make use of our flexible modeling framework. It is shown that changes in the renewable generation portfolio have a considerable impact on the commercial business activities of market participants such as power plant owners as well as suppliers. More importantly, and in contrast to existing approaches, the SLHS model is capable of quantifying the distinct impact of changes in the renewable generation portfolio on wholesale market prices. This makes it clear that market participants, facing the challenges in renewable-dominant power markets, require a modeling approach like ours. Given the fact our model captures local risks

it lends itself well to manage production risks of renewable energy projects. Local weather conditions across distant locations can be very different. Consequently, we also expect a similar degree of heterogeneity for local renewable production. A methodology that is only capable of capturing the aggregate market-wide renewable generation is thus insufficient to assess idiosyncratic production risks of specific locations. For instance, investors might want to weight risk and reward for new renewable energy projects. Also, to manage production risks, producers might be interested in the potential of risk transfer by the use of derivative instruments whose payoff is often times tied to market-wide aggregate variables, such as wholesale spot prices or weather-related indices. A meaningful assessment consequently necessitates a model which incorporates the joint distribution of local weather conditions and important market variables such as spot prices. The SLHS model is a suitable tool to address these challenges.

References

Aid, R., L. Campi, and N. Langrené (2013). A Structural Risk-Neutral Model for Pricing and Hedging Power Derivatives. *Mathematical Finance 23*, 387–438.

Alexandridis, A. and A. Zapranis (2013). Wind Derivatives: Modeling and Pricing. *Computational Economics 41*, 299–326.

Anemos (2019). *Wind Atlas and Production Index Europe: Commercial Supplier of Weather Data for Wind Power Applications*. Anemos Gesellschaft für Umweltmeteorologie mbH.

Barlow, M. T. (2002). A Diffusion Model for Electricity Prices. *Mathematical Finance 12*, 287–298.

Benth, F., J. Kallsen, and T. Mayer-Brandis (2007). A Non-Gaussian Ornstein-Uhlenbeck Process for Electricity Spot Price Modeling and Derivatives Pricing. *Applied Mathematical Finance 14*, 153–169.

Benth, F. E., R. Biegler-König, and R. Kiesel (2013). An Empirical Study of the Information Premium on Electricity Markets. *Energy Economics 36*, 55–77.

Berry, L. N. and N. E. Helwig (2021). Cross-Validation, Information Theory, or Maximum Likelihood? A Comparison of Tuning Methods for Penalized Splines. *Stats 4*, 701–724.

Bieger-König, R. (2013). *The Information Premium on Electricity Markets*. Dissertation, University of Duisburg-Essen.

Brooks, C. and M. Prokopczuk (2013). The Dynamics of Commodity Prices. *Quantitative Finance 13*, 527–542.

Brown, B. G., R. W. Katz, and A. H. Murphy (1984). Time Series Models to Simulate and Forecast Wind Speed and Wind Power. *Journal of Applied Meteorology and Climatology 23*, 1184–1195.

Bundesnetzagentur (2019a). EEG-Registerdaten. `https://www.bundesnetzagentur.de/DE/Fachthemen/ElektrizitaetundGas/ErneuerbareEnergien/ZahlenDatenInformationen/EEG_Registerdaten/start.html`.

Bundesnetzagentur (2019b). Marktstammdatenregister. `https://www.marktstammdatenregister.de/MaStR`.

Bundesnetzagentur (2021). EEG in Zahlen. `https://www.bundesnetzagentur.de/DE/Sachgebiete/ElektrizitaetundGas/Unternehmen_Institutionen/ErneuerbareEnergien/ZahlenDatenInformationen/start.html`.

Burger, M., B. Klar, A. Müller, and G. Schindlmayr (2003). A Spot Market Model for Pricing Derivatives in Electricity Markets. *Quantitative Finance 4*, 109–122.

Caporin, M. and J. Preś (2012). Modelling and Forecasting Wind Speed Intensity for Weather Risk Management. *Computational Statistics and Data Analysis 56*, 3459–3476.

Cartea, Á. and M. G. Figueroa (2005). Pricing in Electricity Markets: A Mean Reverting Jump Diffusion Model with Seasonality. *Applied Mathematical Finance 12*, 313–335.

Cartea, Á., M. G. Figueroa, and H. Geman (2009). Modelling Electricity Prices with Forward Looking Capacity Constraints. *Applied Mathematical Finance 16*, 103–122.

Cludius, J., H. Hermann, F. C. Matthes, and V. Graichen (2014). The Merit Order Effect of Wind and Photovoltaic Electricity Generation in Germany 2008-2016: Estimation and Distributional Implications. *Energy Economics 44*, 302–313.

Coulon, M., W. B. Powell, and R. Sircar (2013). A Model for Hedging Load and Price Risk in the Texas Electricity Market. *Energy Economics 40*, 976–988.

Deng, S. (2000). Stochastic Models of Energy Commodity Prices and Their Applications: Mean-Reversion with Jumps and Spikes. Working Paper.

EEX (2022). EEX-Transparency. `https://www.eex-transparency.com/`.

Enercon (2022). E-82 E2. `https://www.enercon.de/en/products/ep-2/e-82/`.

EnergyMap (2016). EEG-Anlagenregister. `http://www.energymap.info/`.

ENTSO-E (2022). ENTSOE-E Transparency Platform. `https://transparency.entsoe.eu/`.

EurObserv'ER (2019). The State of Renewable Energies in Europe.

Eydeland, A. and K. Wolyniec (2003). *Energy and Power Risk Management: New Developments in Modeling, Pricing, and Hedging*. Wiley Finance.

Füss, R., S. Mahringer, and M. Prokopczuk (2015). Electricity Derivatives Pricing with Forward-Looking Information. *Journal of Economic Dynamics and Control 58*, 34–57.

Geman, H. and A. Roncoroni (2006). Understanding the Fine Structure of Electricity Prices. *The Journal of Business 79*, 1225–1261.

GmbH, E. (2022). energate messenger+. `https://www.energate-messenger.de/`.

Grothe, O. and F. Müsgens (2013). The Influence of Spatial Effects on Wind Power Revenues under Direct Marketing Rules. *Energy Policy 58*, 237–247.

Grothe, O. and J. Schnieders (2011). Spatial Dependence in Wind and Optimal Wind Power Allocation: A Copula-Based Analysis. *Energy Policy 39*, 4742–4754.

Hagspiel, S., A. Papaemannouil, M. Schmid, and G. Andersson (2012). Copula-Based Modeling of Stochastic Wind Power in Europe and Implications for the Swiss Power Grid. *Applied Energy 96*, 33–44.

Hain, M., M. Uhrig-Homburg, and N. Unger (2018). Risk Factors and Their Associated Risk Premia: An Empirical Analysis of the Crude Oil Market. *Journal of Banking and Finance 95*, 44–63.

Hambly, B., S. Howison, and T. Kluge (2009). Modelling Spikes and Pricing Swing Options in Electricity Markets. *Quantitative Finance 9*, 937–949.

Hastie, T. and R. Tibshirani (1993). Varying-Coefficient Models. *Journal of the Royal Statistical Society 55*, 757–779.

Howison, S. and M. Coulon (2009). Stochastic Behaviour of the Electricity Bid Stack: From Fundamental Drivers to Power Prices. *The Journal of Energy Markets 2*, 29–69.

International Renewable Energy Agency (2018). Renewable Energy Balances. `https://www.irena.org/Statistics/View-Data-by-Topic/Renewable-Energy-Balances/Final-Renewable-Energy-Consumption`.

Jacobsen, H. K. and E. Zvingilaite (2010). Reducing the Market Impact of Large Shares of Intermittent Energy in Denmark. *Energy Policy 38*, 3403–3413.

Jenkin, T., A. Larson, M. Ruth, B. King, and P. Spitsen (2018). The Use of Statistically Based Rolling Supply Curves for Electricity Market Analysis: A Preliminary Look. *National Renewable Energy Laboratory*.

Jónsson, T., P. Pinson, and H. Madsen (2010). On the Market Impact of Wind Energy Forecasts. *Energy Economics 32*, 313–320.

Kallabis, T., C. Pape, and C. Weber (2016). The Plunge in German Electricity Futures Prices – Analysis Using a Parsimonious Fundamental Model. *Energy Policy 95*, 280–290.

Keles, D., M. Genoese, D. Möst, S. Ortlieb, and W. Fichtner (2013). A Combined Modeling Approach for Wind Power Feed-In and Electricity Spot Prices. *Energy Policy 59*, 213–225.

Ketterer, J. C. (2014). The Impact of Wind Power Generation on the Electricity Price in Germany. *Energy Economics 44*, 270–280.

Kiesel, R. and M. Kusterman (2016). Structural Models for Coupled Electricity Markets. *Journal of Commodity Markets 3*, 16–38.

Lago, J., F. De Ridder, and B. De Schutter (2018). Forecasting Spot Electricity Prices: Deep Learning Approaches and Empirical Comparison of Traditional Algorithms. *Applied Energy 221*, 386–405.

Lago, J., G. Marcjasz, B. De Schutter, and R. Weron (2021). Forecasting Day-Ahead Electricity Prices: A Review of State-of-the-Art Algorithms, Best practices and an Open-Access Benchmark. *Applied Energy 293*, 116983.

Lehna, M., F. Scheller, and H. Herwartz (2021). Forecasting Day-Ahead Electricity Prices: A Comparison of Time Series and Neural Network Models Taking External Regressors Into Account. *Energy Economics 106*, 105742.

Liu, P.-L. and A. Der Kiureghian (1986). Multivariate Distribution Models with Prescribed Marginals and Covariances. *Probabilistic Engineering Mechanics 1*, 105–112.

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

Meyer-Brandis, T. and P. Tankov (2008). Multi-Factor Jump-Diffusion Models of Electricity Prices. *International Journal of Theoretical and Applied Finance 11*, 503–528.

Mora-Lopez, L. and M. Sidrach de Cardona (1998). Multiplicative ARMA Models to Generate Hourly Series of Global Irradiation. *Solar Energy 63*, 283–291.

Morales, J. M., R. Minguez, and A. J. Conejo (2010). A Methodology to Generate Statistically Dependent Wind Speed Scenarios. *Applied Energy 87*, 843–855.

Netzentwicklungsplan (2022). Netzentwicklungsplan. `https://www.netzentwicklungsplan.de`.

Papaefthymiou, G. and D. Kurowicka (2009). Using Copulas for Modeling Stochastic Dependence in Power System Uncertainty Analysis. *IEEE Transactions on Power Systems 24*, 40–49.

Papavasiliou, A. and S. S. Oren (2013). Multiarea Stochastic Unit Commitment for High Wind Penetration in a Transmission Constrained Network. *Operations Research 61*, 578–592.

Paraschiv, F., D. Erni, and R. Pietsch (2014). The Impact of Renewable Energies on EEX Day-Ahead Electricity Prices. *Energy Policy 73*, 196–210.

Pieralli, S., M. Ritter, and M. Odening (2015). Efficiency of Wind Power Production and Its Determinants. *Energy 90*, 429–438.

PSU/NCAR (2019). MM5 Community Model Homepage: MM5 Modeling System Overview. `https://www2.mmm.ucar.edu/mm5/`.

Ritter, M. and L. Deckert (2017). Site Assessment, Turbine Selection, and Local Feed-In Tariffs through the Wind Energy Index. *Applied Energy 185*, 1087–1099.

Ritter, M., Z. Shen, B. L. Cabrera, M. Odening, and L. Deckert (2015). Designing an Index for Assessing Wind Energy Potential. *Renewable Energy 83*, 416–424.

Rose, S. and J. Apt (2015). What Can Reanalysis Data Tell Us about Wind Power? *Renewable Energy 83*, 963–969.

Schermeyer, H., V. Bertsch, and W. Fichtner (2014). Validation and Utilization of Numerical Weather Model Data in Energy Systems Analysis, Modelling and Forecasting of Decentralized Renewable Electricity Production. *7th International Scientific Conference on Energy and Climate Change*, 137–146.

Seifert, J. and M. Uhrig-Homburg (2007). Modelling Jumps in Electricity Prices: Theory and Empirical Evidence. *Review of Derivatives Research 10*, 59–85.

Sigauke, C. (2017). Forecasting Medium-Term Electricity Demand in a South African Electric Power Supply System. *Journal of Energy in Southern Africa 28*, 54–67.

Taylor, J. W. and S. Majithia (2000). Using Combined Forecasts with Changing Weights for Electricity Demand Profiling. *Journal of the Operational Research Society 51*, 72–82.

Trolle, A. B. and E. S. Schwartz (2009). Unspanned Stochastic Volatility and the Pricing of Commodity Derivatives. *The Review of Financial Studies 22*, 4423–4461.

Uniejewski, B., J. Nowotarski, and R. Weron (2016). Automated Variable Selection and Shrinkage for Day-Ahead Electricity Price Forecasting. *Energies 9*, 621.

Wagner, A. (2014). Residual Demand Modeling and Application to Electricity Pricing. *The Energy Journal 35*.

Weron, R. (2006). *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach*. Wiley Finance.

Weron, R. and F. Ziel (2019). *Electricity Price Forecasting*. U. Soytas and R. Sari (Editors).

Wozabal, D., C. Graf, and D. Hirschmann (2016). The Effect of Intermittent Renewables on the Electricity Price Variance. *OR Spectrum 38*, 687–709.

Ziel, F. and R. Steinert (2016). Electricity Price Forecasting Using Sale and Purchase Curves: The X-Model. *Energy Economics 59*, 435–454.

A Overview of Data

The following data is used for the respective model components:

Model Component	Time	Source	Note
Solar irradiation	01.01.1990 - 31.12.2018	Anemos (2019)	hourly, 6704 weather cells, reanalysis
Wind speed	01.01.1990 - 31.12.2018	Anemos (2019)	hourly, 6704 weather cells, reanalysis
Demand	01.01.2017 - 31.12.2018	ENTSO-E (2022)	15 min, hourly averages, balancing area, ex ante
Wind energy infeed	01.01.2017 - 31.12.2018	ENTSO-E (2022)	15 min, hourly averages, balancing area, ex ante
PV energy infeed	01.01.2017 - 31.12.2018	ENTSO-E (2022)	15 min, hourly averages, balancing area, ex ante
PV and wind capacity	31.12.2016 - 31.12.2018	Bundesnetzagentur (2019a), Bundesnetzagentur (2019b)	yearly, hourly interpolating
Day-ahead prices	01.01.2017 - 31.12.2018	EEX (2022)	hourly, market zones (change of DE-AT-LU in 2018)

Table 4: Overview of data time series for the SLHS model components

The table shows for each model component the time frame, data source as well as notes considering data granularity or specific circumstances.

B Estimation of Production Curves

Given observed patterns regarding dependencies between renewable power production and its corresponding weather variable we choose a 3-parameter logistic function for wind and a second-order polynomial for solar:

$$g^{w,k}(y_t^{w,k}) = \frac{\gamma_0^k}{1 + e^{-\gamma_1^k(y_t^{w,k} - \gamma_2^k)}}$$
$$g^{s,k}(y_t^{s,k}) = \pi_0^k + \pi_1^k y_t^{s,k} + \pi_2^k (y_t^{s,k})^2$$

Parameters are obtained by non-linear least squares (least-squares) for wind power (solar power). Table 5 shows estimated coefficients along with standard errors.

A. Wind power	50Hertz	TenneT	Amprion	TransnetBW
γ_0^k	$6.53 \cdot 10^{-1}$	$8.43 \cdot 10^{-1}$	$4.66 \cdot 10^{-1}$	$2.78 \cdot 10^{-1}$
	$(3.43 \cdot 10^{-3})^{***}$	$(3.04 \cdot 10^{-3})^{***}$	$(2.16 \cdot 10^{-3})^{***}$	$(1.57 \cdot 10^{-3})^{***}$
γ_1^k	$4.14 \cdot 10^{-1}$	$4.41 \cdot 10^{-1}$	$4.57 \cdot 10^{-1}$	$4.47 \cdot 10^{-1}$
	$(2.30 \cdot 10^{-3})^{***}$	$(2.30 \cdot 10^{-3})^{***}$	$(2.35 \cdot 10^{-3})^{***}$	$(2.70 \cdot 10^{-3})^{***}$
γ_2^k	$1.02 \cdot 10^{1}$	$9.07 \cdot 10^{0}$	$9.57 \cdot 10^{0}$	$9.67 \cdot 10^{0}$
	$(3.07 \cdot 10^{-2})^{***}$	$(2.13 \cdot 10^{-2})^{***}$	$(2.52 \cdot 10^{-2})^{***}$	$(3.10 \cdot 10^{-2})^{***}$
B. Solar power	50Hertz	TenneT	Amprion	TransnetBW
π_0^k	$7.06 \cdot 10^{-3}$	$5.56 \cdot 10^{-3}$	$2.89 \cdot 10^{-3}$	$2.21 \cdot 10^{-2}$
	$(4.94 \cdot 10^{-4})^{***}$	$(4.17 \cdot 10^{-4})^{***}$	$(2.58 \cdot 10^{-4})^{***}$	$(1.46 \cdot 10^{-3})^{***}$
π_1^k	$9.38 \cdot 10^{-4}$	$8.94 \cdot 10^{-4}$	$5.14 \cdot 10^{-4}$	$2.12 \cdot 10^{-3}$
	$(5.76 \cdot 10^{-6})^{***}$	$(4.74 \cdot 10^{-6})^{***}$	$(2.92 \cdot 10^{-6})^{***}$	$(1.61 \cdot 10^{-5})^{***}$
π_2^k	$-1.83 \cdot 10^{-7}$	$-2.18 \cdot 10^{-7}$	$-1.19 \cdot 10^{-7}$	$-4.16 \cdot 10^{-7}$
	$(8.33 \cdot 10^{-9})^{***}$	$(6.72 \cdot 10^{-9})^{***}$	$(4.15 \cdot 10^{-9})^{***}$	$(2.19 \cdot 10^{-8})^{***}$

Table 5: Parameter estimates of power curves

The table shows parameter estimates and standard errors (in parantheses) for production curves of wind power (Panel A) as well as for solar power (Panel B) in all four balancing-areas. Estimates are obtained by non-linear least-squares (least-squares) for wind power (solar power) from 01 January 2017 to 31 December 2018. Note that parameter estimates are very small for the case of solar due to the fact that the weather variable (solar irradiation) takes relatively large values (between 0 and 1000 W/m^2) compared to wind speed (between 0 and 20 m/s). ***, **, and * denotes statistical significance at the 0.1%, 1%, and 5% level.

C Empirical Analysis of Weather Model

C.1 Wind speed dynamics

Figure 14 visualizes the hourly wind speed in the southern part of Germany for 2005-2008 as well as for a week in June 2002. Clearly, wind speed is highly volatile and can change dramatically just within a few hours. Its levels also tend to be lower during summer than winter.[41] On top of that, volatility tends to be a lot higher during autumn and winter months. In a similar fashion, higher moments such as skewness and excess kurtosis also vary and seem to rise during autumn and winter season. For most locations, the wind speed distributions' shape thus seems to exhibit considerable seasonality. Wind speed furthermore exhibits considerable evidence for non-Gaussian behavior with a positive skewness and significant positive autocorrelation (see Figure 14). Unsurprisingly, these characteristics are not homogeneous across locations. Figure 15 depicts the time series of wind speed at two different weather cells and clearly demonstrates that although key characteristics such as trends in level and volatility prevail, differences do exist.

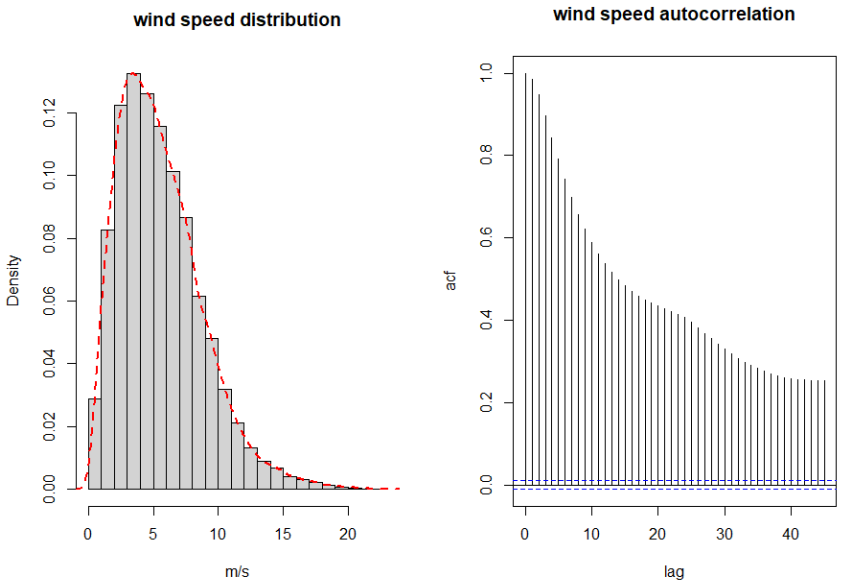

Figure 14: Wind speed distribution and autocorrelation function

The left graph shows the histogram of wind speed in a weather cell located in southern Germany (Baden-Wuerttemberg) along with the density of a normal distribution with identical mean and standard deviation (red dotted line). The right graph depicts the corresponding autocorrelation function.

[41] We also observe a weak day- and night pattern.

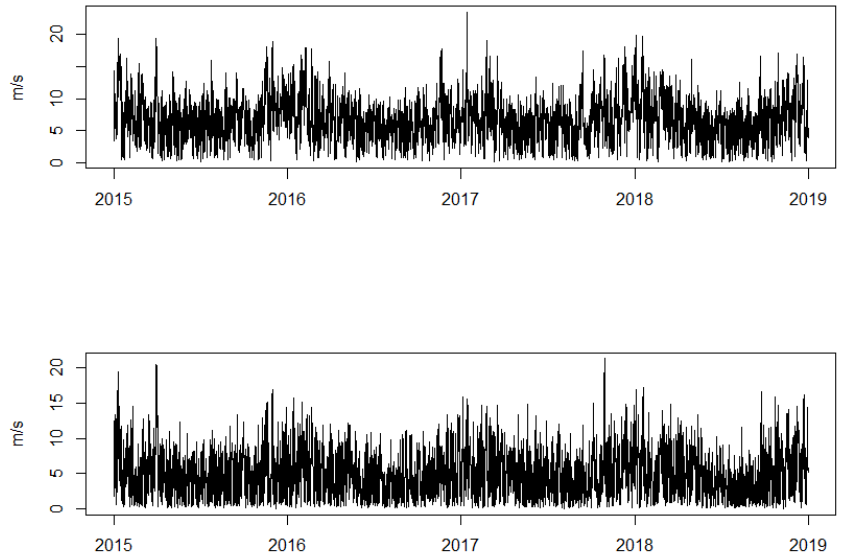

Figure 15: Time series of wind speed at different locations

The figure shows time series of wind speed for a weather cell in southern Germany (top) as well as for a weather cell located in the North Sea (bottom) for the years 2015-2018.

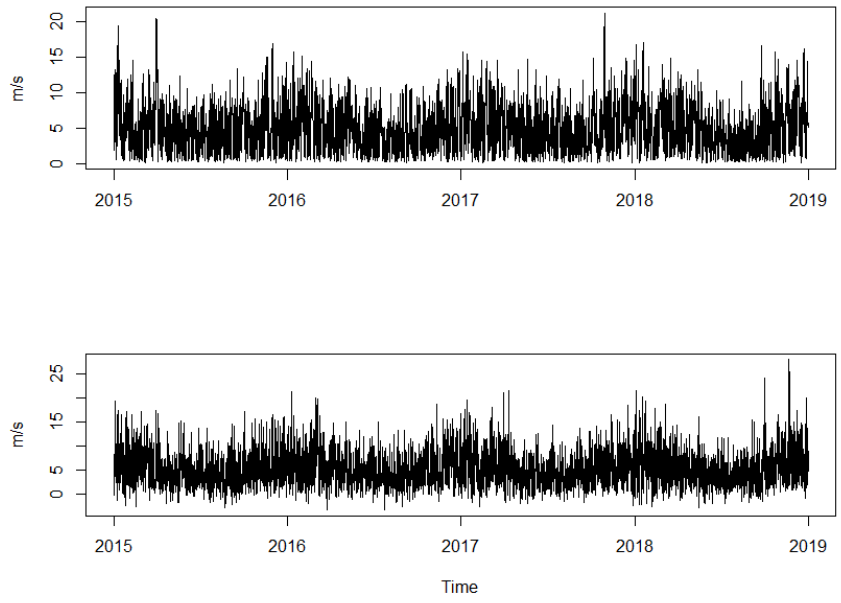

Figure 16: Observed vs. simulated wind speed

The figure shows observed (top) as well as simulated (bottom) wind speed for a weather cell located in southern Germany.

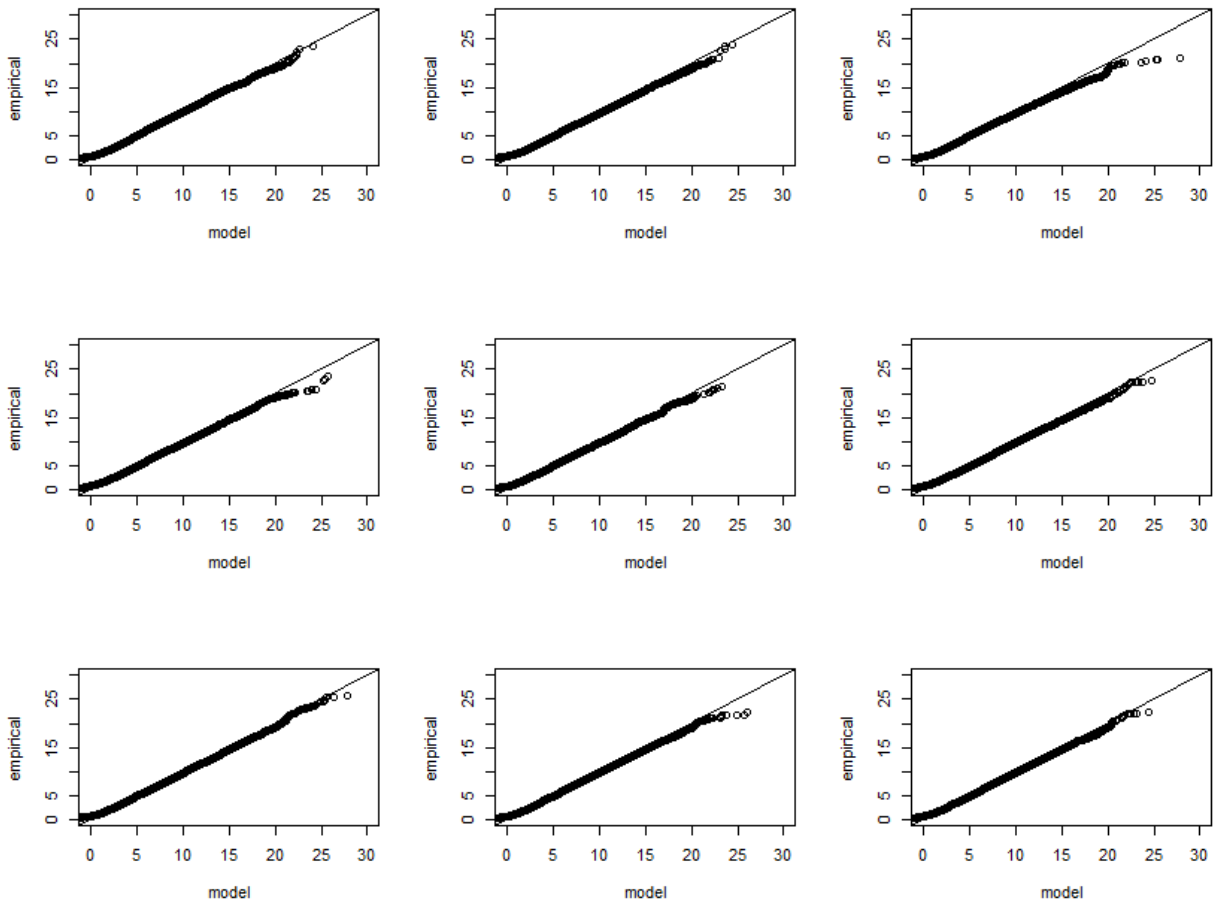

Figure 17: QQ-plots of simulated vs. observed wind speed

The figure depicts QQ-plots of simulated (x-axis) vs. observed (y-axis) wind speed of nine exemplary selected locations throughout the German market area.

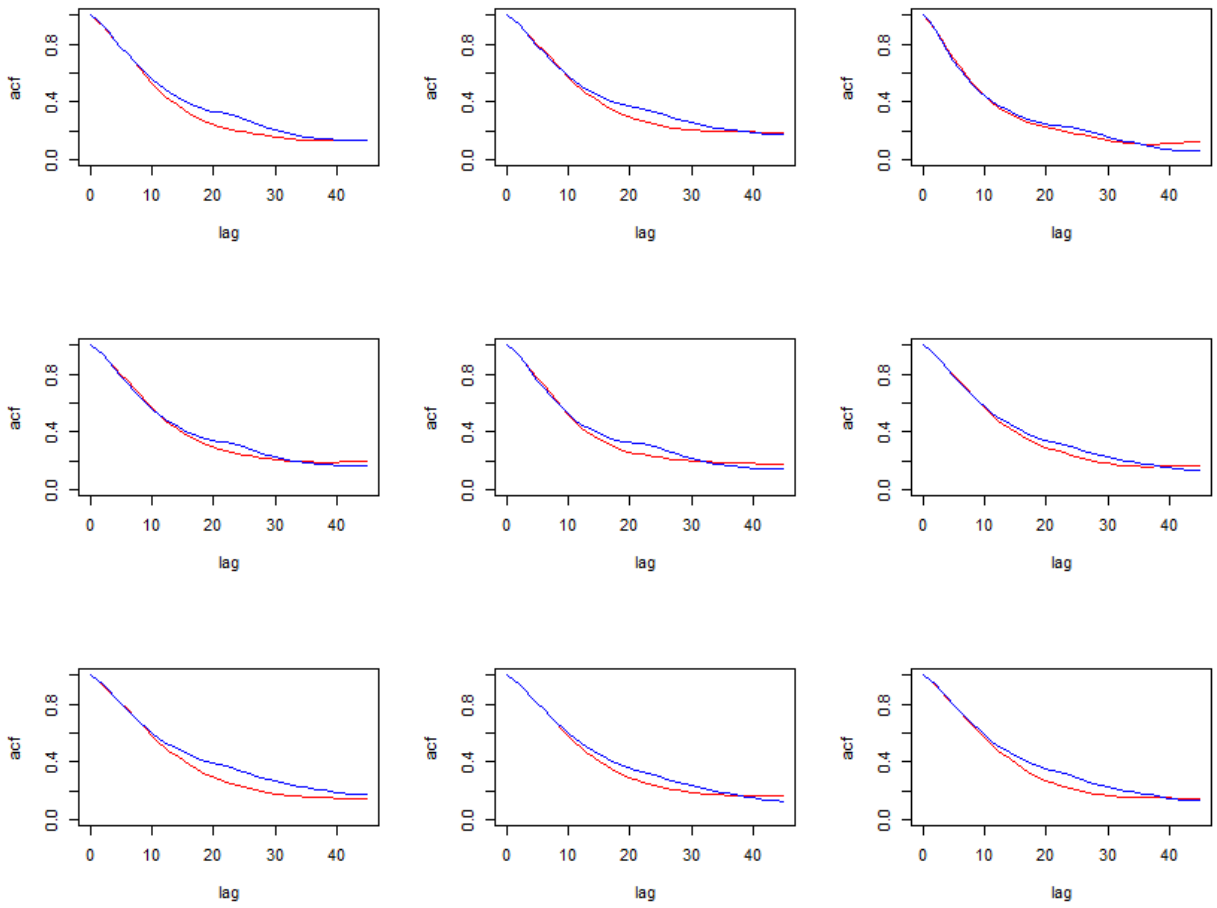

Figure 18: ACFs of observed and simulated wind speed

The figure visualizes the actual (blue) as well as the simulated (red) autocorrelation function of de-trended and transformded wind speed $\hat{y}_t^w$ for nine exemplary selected locations throughout Germany.

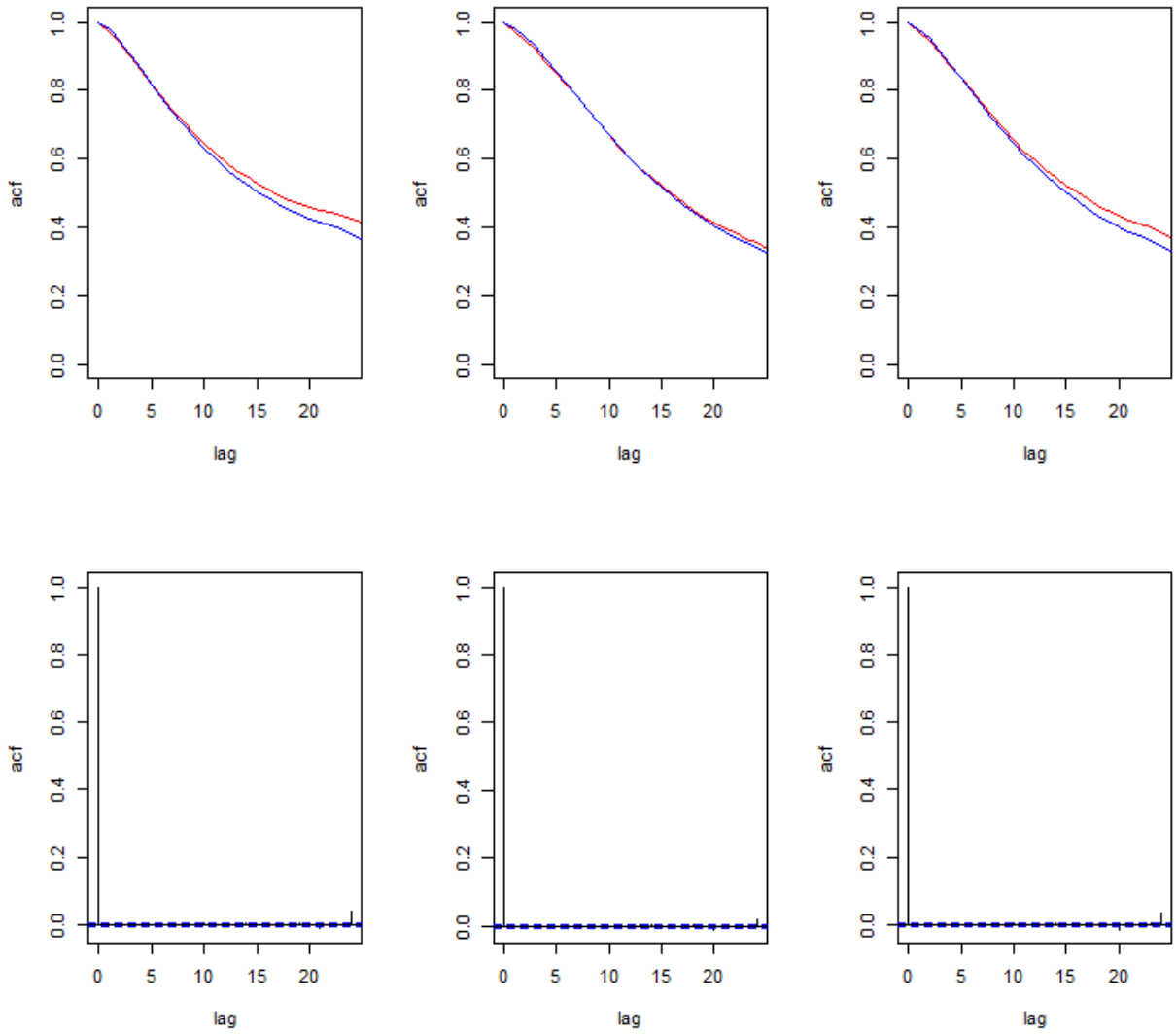

Figure 19: ACFs of observed and simulated wind speed

The top three graphs visualize the actual (blue) as well as the simulated (red) autocorrelation function of de-trended and transformed wind speed $\hat{y}_t^w$ for three exemplary locations in Germany. The bottom three graphs show the autocorrelation function of the corresponding model residuals with 5 % confidence bound.

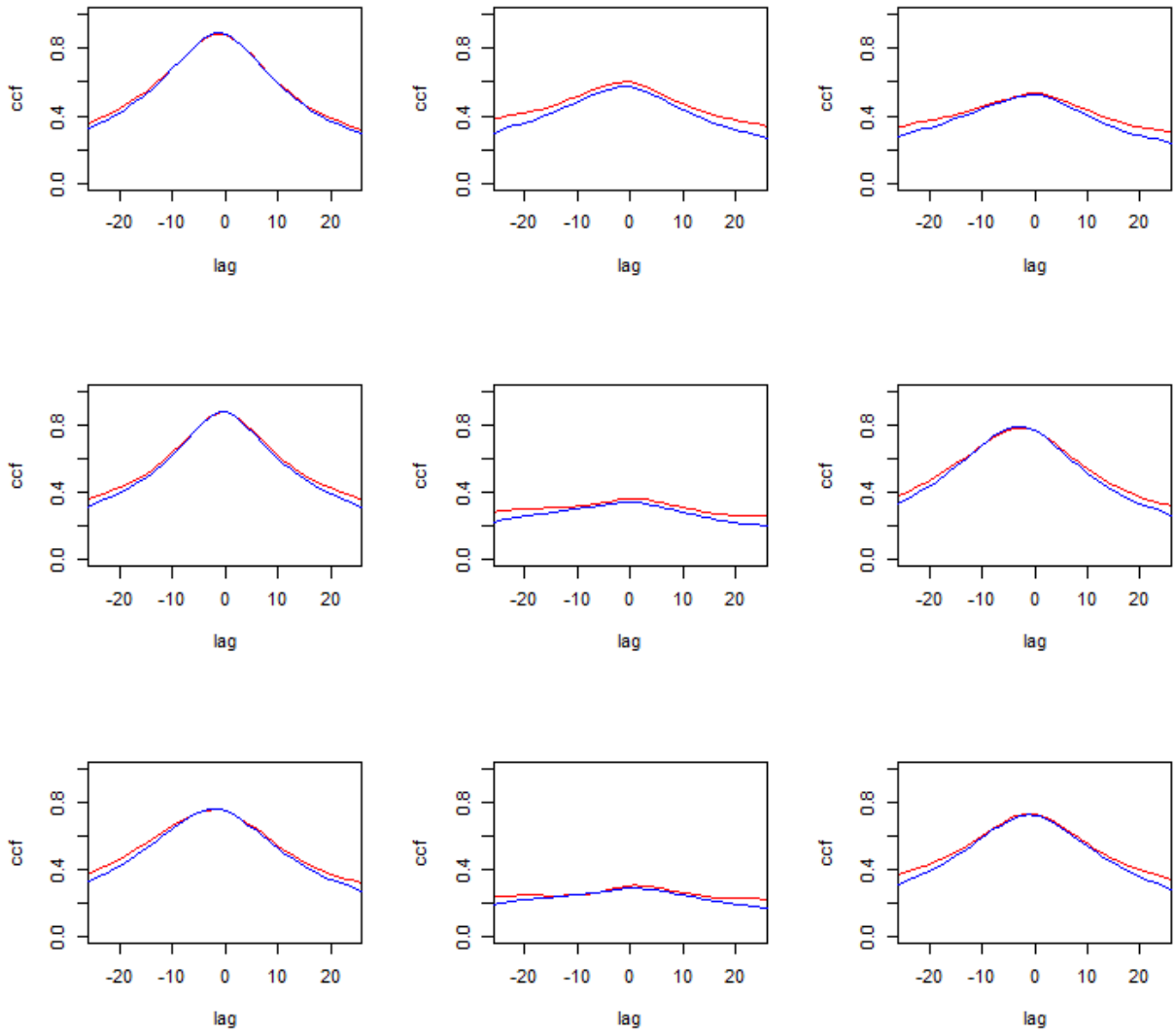

Figure 20: CCFs of observed and simulated wind speed

The figure visualizes the actual (blue) as well as the simulated (red) crosscorrelation function of de-trended and transformed wind speed $\hat{y}_t^w$ for nine pairs of exemplary selected locations throughout Germany.

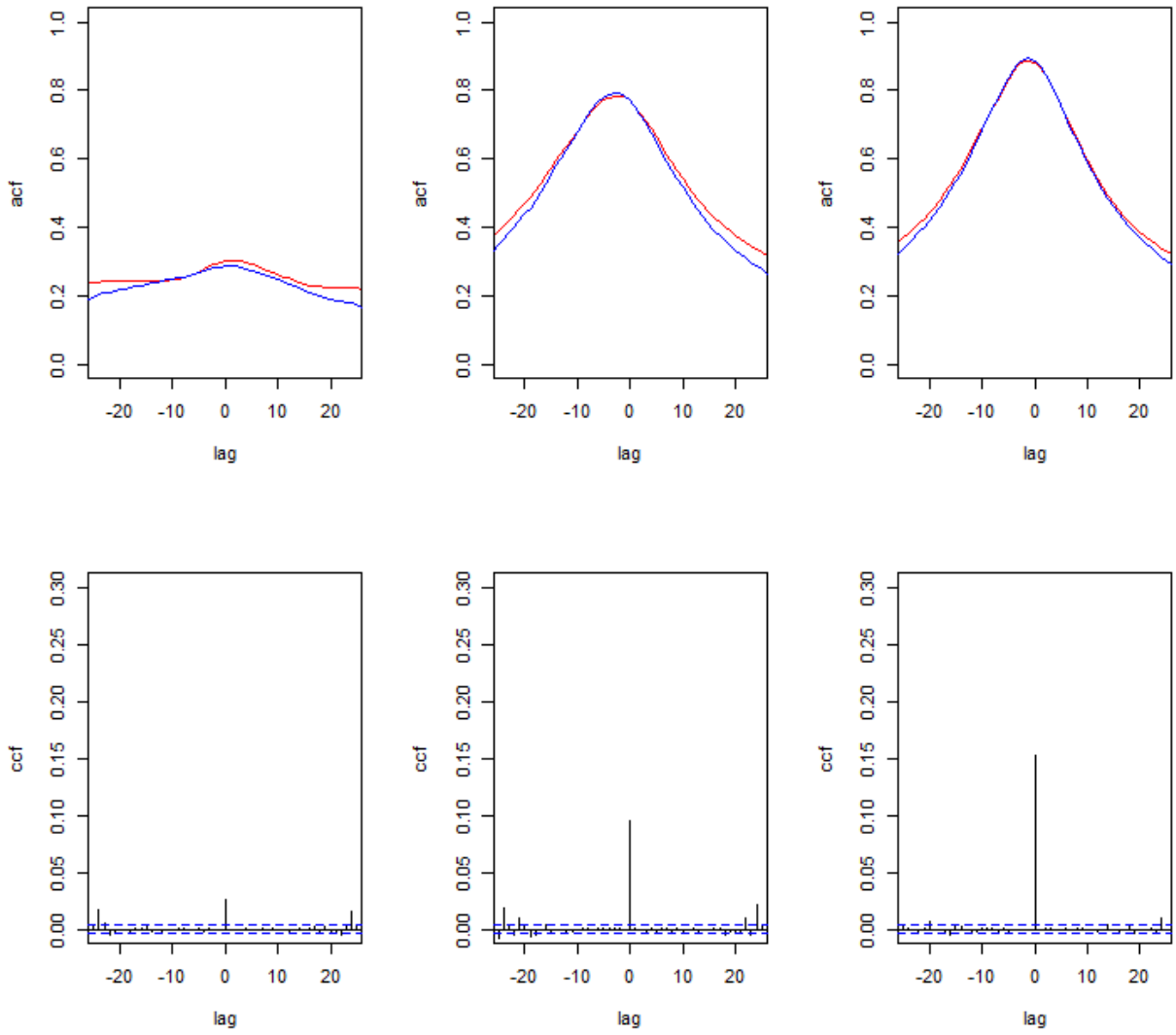

Figure 21: CCFs of wind speed and model residuals

The top three graphs visualize the actual (blue) as well as the simulated (red) crosscorrelation function of de-trended and transformed wind speed $\hat{y}_t^w$ between three pairs of exemplary locations in Germany. The bottom three graphs show the crosscorrelation function of the corresponding model residuals with 5 % confidence bounds.

In addition to the general specification outlined in Section 3.3.1, we therefore incorporate a time-dependent and location-specific volatility function $\sigma_t^{w,k}$ and distribution function $F_t^{w,k}$ to capture seasonal time-variation in the symmetry of the distribution. This is achieved by normalizing the wind speed times series both by its trend as well as by its (seasonal) standard deviation:

$$\overline{y}_t^{w,k} = (y_t^{w,k} - \mu_t^{w,k})(\sigma_t^{w,k})^{-1}$$

where $\sigma_t^{w,k}$ is estimated on a monthly basis. We then group $\overline{y}_t^{w,k}$ by quarters and estimate a corresponding distribution function. Since skewness and kurtosis are highly sensitive to outliers we choose a quarterly time frame to increase the number of observations in each case.

Dependencies between regions also vary and are largely determined by distance. For instance, exemplary adjacent regions located in the of south Germany (state of Baden-Wuerttemberg) exhibit a correlation of over 90% (after removing trends in levels and volatility) while the relationship to a third region situated in the north-eastern part of Germany is much lower (30 and 40 % respectively).

To demonstrate the overall fit of our modeling approach we compare empirical and (model-implied) simulated values. Figure 16 shows the trajectory of wind speed for a region in north-western Germany over the years 2010-2018 along with its simulated correspondent indicating that the approach seems to be capable of reproducing the prominent statistical characteristics. This can also be seen by inspecting QQ-plots of simulated vs. observed wind speed values (Figure 17). The model-implied autocorrelation functions also come very close to their observed counterparts (Figure 18). This is also apparent by having a look at model residuals which mostly void of any significant autocorrelation (Figure 19). Similarly, the modeling approach does a good job at capturing the cross-sectional dependencies between locations. We obtain a close fit to the observed crosscorrelation function (Figure 20) and model residuals show hardly any evidence of dependencies across locations (Figure 21). Overall, this results suggest that our approach is capable of capturing both location-specific characteristics and cross-sectional dependencies.

C.2 Solar Irradiation Dynamics

Solar irradiation shows a very pronounced intraday pattern (see left graph in Figure 22) due to the fact that there is no sunshine during night time. This pattern gradually changes throughout the year, that is, the sun rises earlier and sets later in the evening during the summer months. However, the peak is happening at noon regardless of the current season. The absolute level of daily peaks also changes seasonally indicated by the right graph in Figure 22. Apart from the above deterministic day- and night pattern, cloud formations can lead to unexpected drops in irradiation levels. Figure 23 shows hourly irradiation in the southern part of Germany for a week in April. Sudden strong decreases thus regularly happen. However, we noticed that most of the time, days tend to be either sunny or rather cloudy overall. Thus, consecutive extreme drops and rises throughout the same day are rather rare.

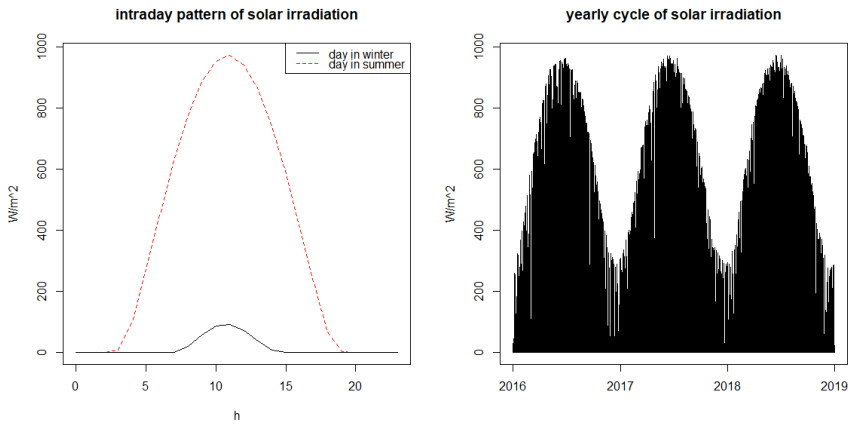

Figure 22: Intraday pattern and yearly cycle of solar irradiation

The top two graphs show the pronounced seasonality in the solar irradiation dynamics within a day (top left) as well as across seasons (top right). The weather cell is located in southern Germany (Baden Wuerttemberg).

Cloud formations thus have a considerable impact on the resulting levels of solar irradiation and one is incapable of observing these during night time. Consequently, we are blind for any larger cloud formations that have been built up in the hours just before sunrise and which heavily affect the weather variables during the early morning hours. We therefore adjust the general methodology and loosely follow Wagner (2014). Motivated by the fact that extreme intraday volatility is rather uncommon for our data set, we model daily maximum irradiation level $\tilde{y}_t^{s,k}$ and capture any variation throughout the day by means of a deterministic pattern function. To formalize the notion of the daily maximum irradiation level, we introduce the

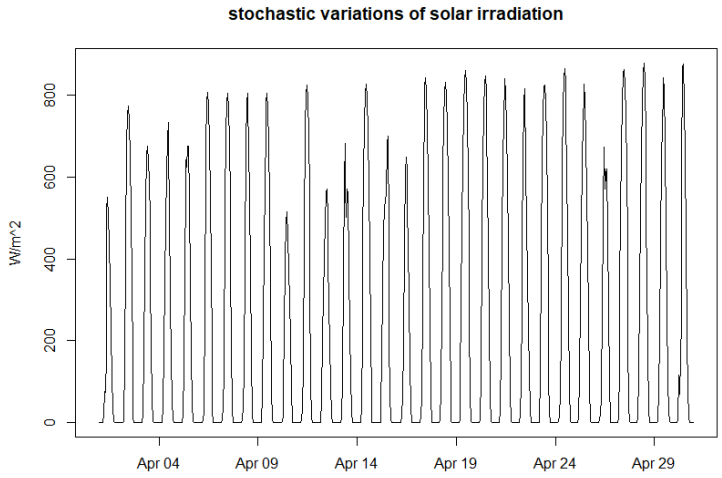

Figure 23: Intraday pattern and yearly cycle of solar irradiation

The figure shows a typical pattern of the variation of solar irradiation within a month (April 2018). The weather cell is located in southern Germany (Baden Wuerttemberg).

day count function that maps all hours in the data set of length T to its respective day of the year

$$d : [0, T] \mapsto \mathbb{N}_0$$

d thus takes values between 1 and 365[42] depending on which day a respective hour t belongs to. The dayily maximum process of local hourly solar irradiation levels then reads as follows

$$\tilde{y}_j^{s,k} = \max_{t:d_t=j}(y_t^{s,k}), \quad j = 1, ..., d_T \tag{9}$$

To account for seasonal variations over the year we de-trend the data $\overline{y}_j^{s,k} := y_j^{s,k} - \mu_j^{s,k}$. The resulting process shows serial and spatial correlation as well as non-normality. The latter issue can be dealt with the transformation-approach outlined in section 3.3.1 to arrive at an approximately normally distributed time series that preserves the original covariance structure as usual. The resulting multivariate time series of (daily) maxima of solar irradiation is then once again captured by means of a VAR model as usual. We furthermore assume solar irradiation and wind speed to be independent from each other since time series correlations

[42]For simplicity, we map the 29th of February in leap years to the 28th of February.

of the corresponding de-trended weather variables are smaller than 25 % (in absolute terms) and cluster near 0.

To capture (local) intraday-variations we make use of a deterministic time-dependent and site-specific function δ_j^k mapping the daily maximum to hourly values as follows:

$$y_t^{s,k} = \delta_{d_t}^k(t, y_{d_t}^{s,k})$$

with

$$\delta_j^k(t, x) = x \sum_{k=1}^{24} \alpha_k \mathbf{1}_{\text{(k-th hour of the day)}}$$

where δ_j^k is estimated for every day of the year and every weather cell separately.

We find that, similarly to the case of wind speed, our modeling approach seems to be capable of reproducing the salient features of solar irradiation (see Figure 24, Figure 25, Figure 26, Figure 27, and Figure 28 respectively).

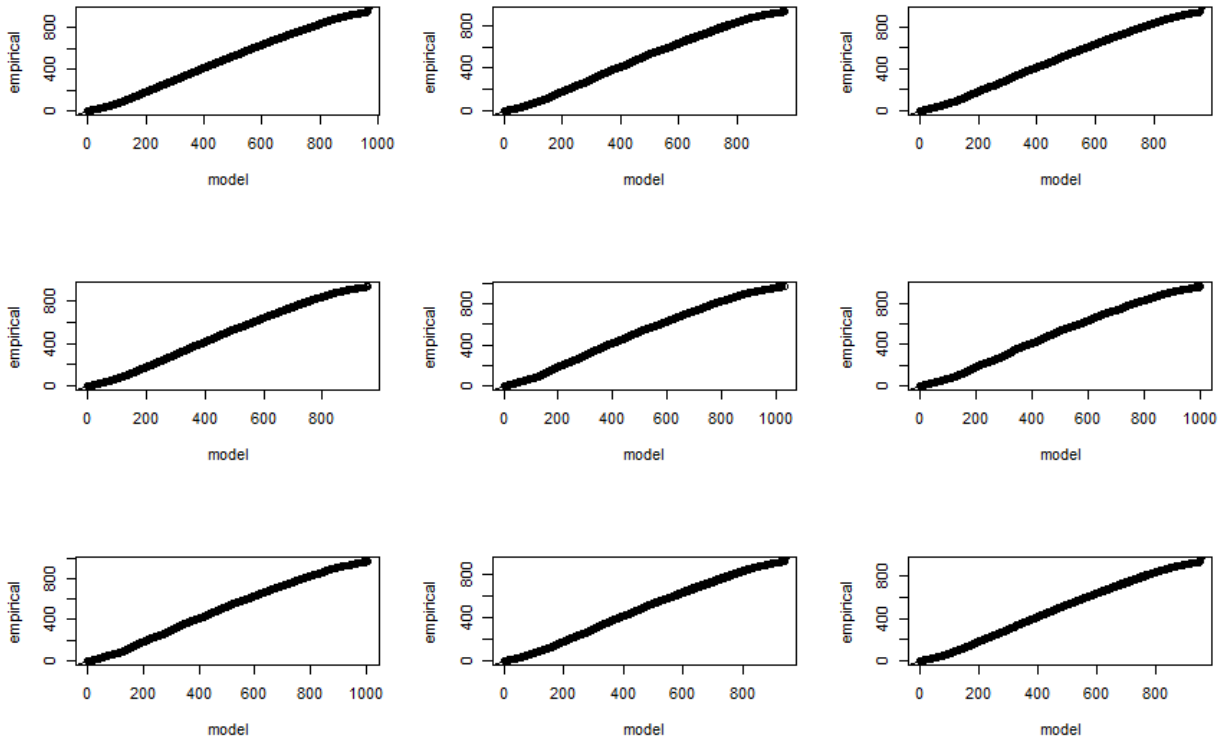

Figure 24: QQ-plots of simulated vs. observed solar irradiation

The figure shows QQ-plots of exemplary selected weather cells of solar irradiation levels.

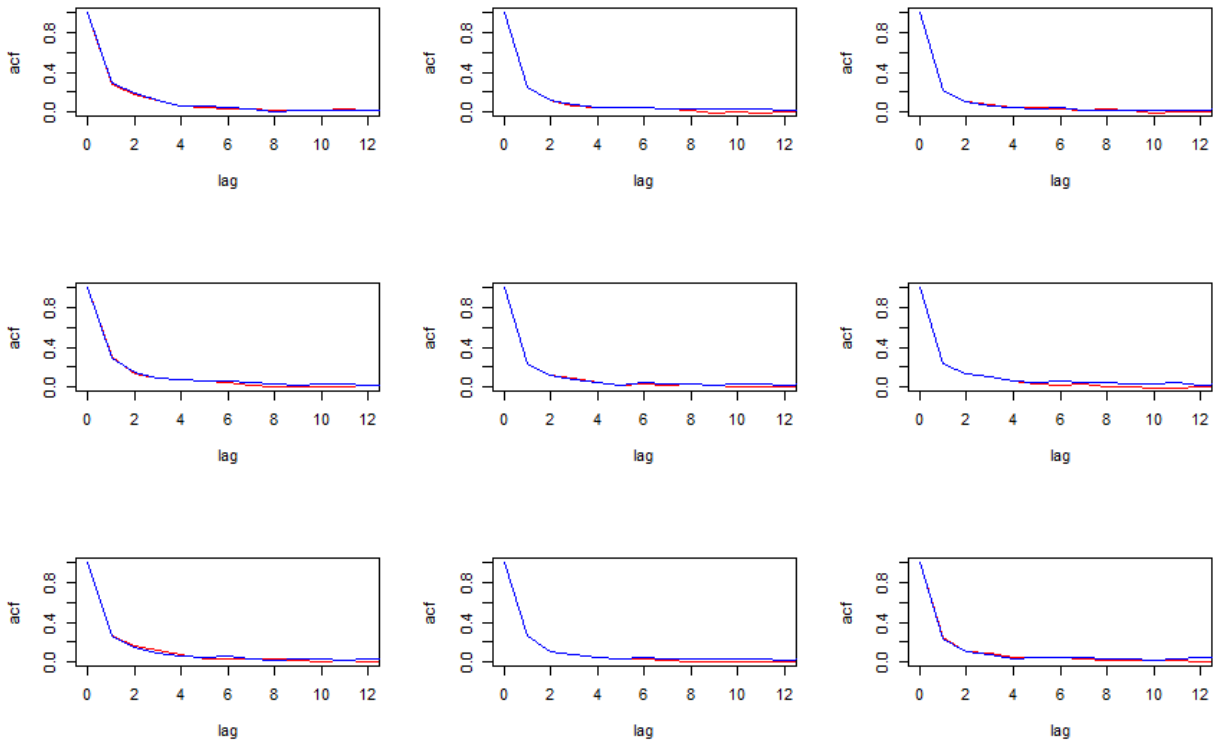

Figure 25: ACFs of observed and simulated solar irradiation

The figure visualizes the actual (blue) as well as the simulated (red) autocorrelation function of de-trended and transformed daily maximum of solar irradiation $\hat{y}_t^{s,k}$ for nine exemplary selected locations throughout Germany.

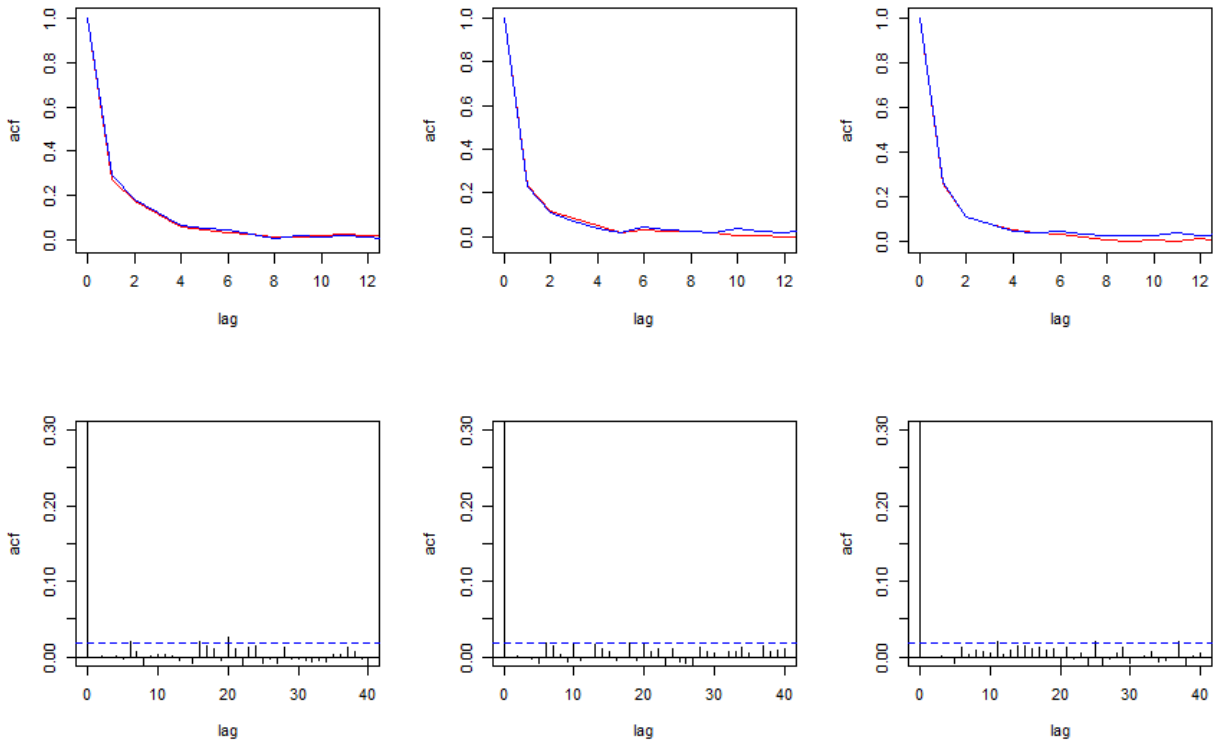

Figure 26: ACFs of solar irradiation and model residuals

The top three graphs visualize the actual (blue) as well as the simulated (red) autocorrelation function of de-trended and transformed daily maximum of solar irradiation $\hat{y}_t^{s,k}$ for three exemplary locations in Germany. The bottom three graphs show the autocorrelation function of the corresponding model residuals with 5 % confidence bounds.

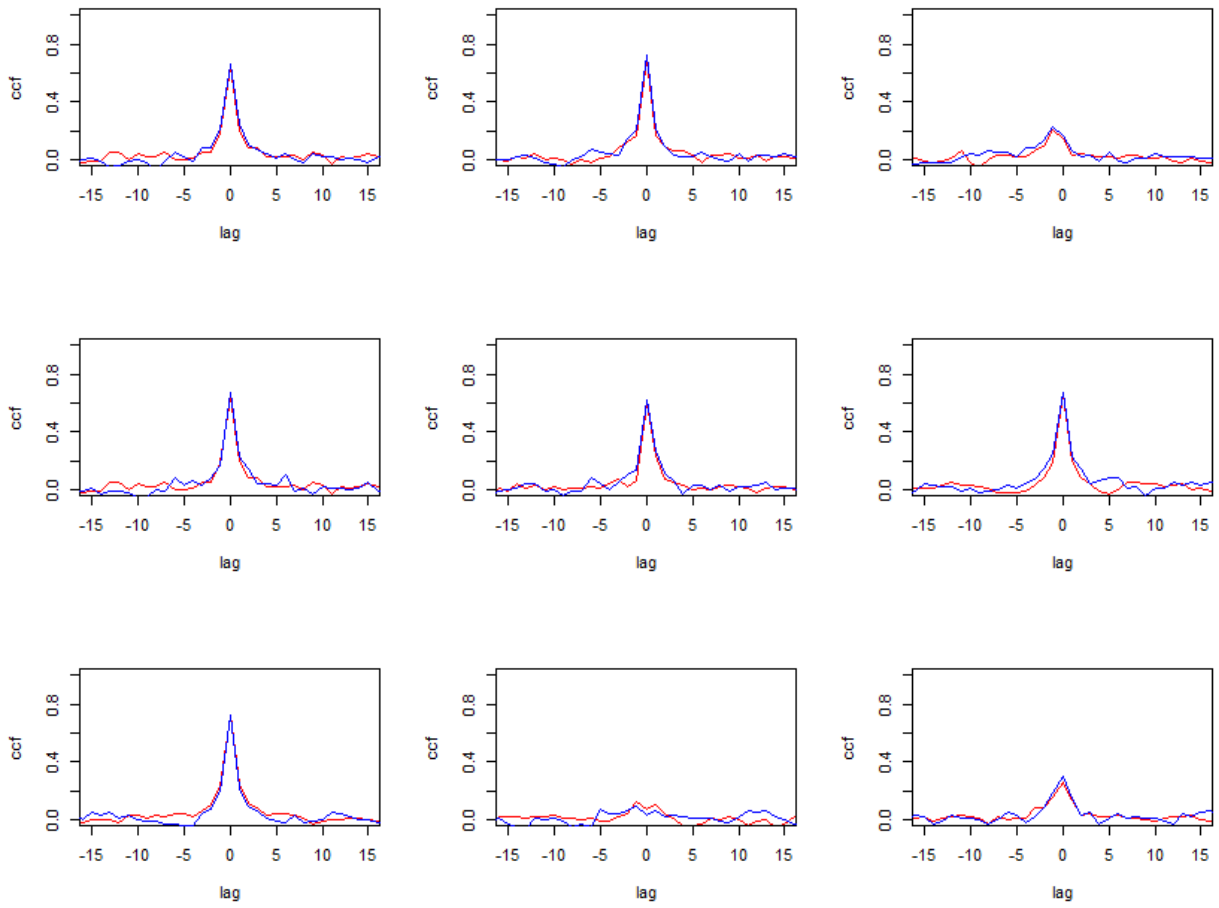

Figure 27: CCFs of observed and simulated solar irradiation

The figure visualizes the actual (blue) as well as the simulated (red) crosscorrelation function of de-trended and transformed daily maximum of solar irradiation $\hat{y}_t^{s,k}$ for nine exemplary selected pairs of locations through Germany.

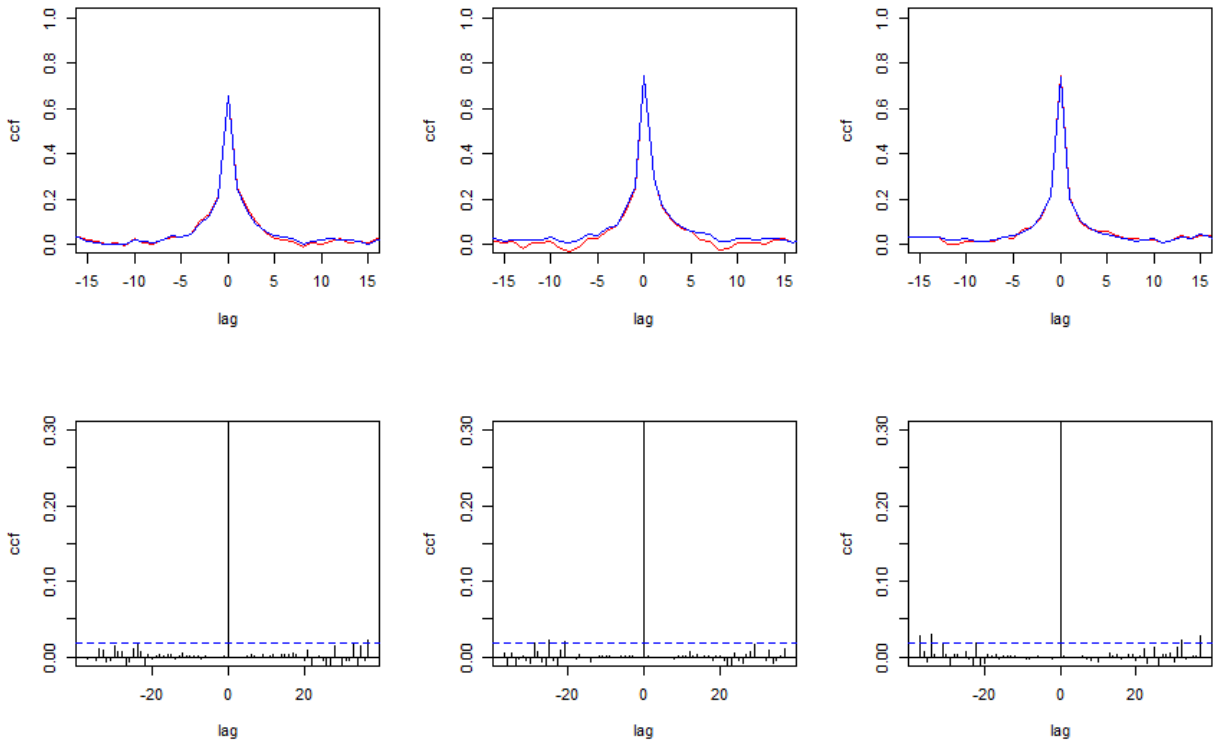

Figure 28: CCFs of solar irradiation and model residuals

The top three graphs visualize the actual (blue) as well as the simulated (red) crosscorrelation function of de-trended and transformed daily maximum of solar irradiation $\hat{y}_t^{s,k}$ between three pairs of exemplary locations in Germany. The bottom three graphs show the crosscorrelation function of the corresponding model residuals with 5 % confidence bounds.

D Supply Curve Estimation

The following Figure 29 displays monthly supply functions estimated with smoothing splines for peak hours:

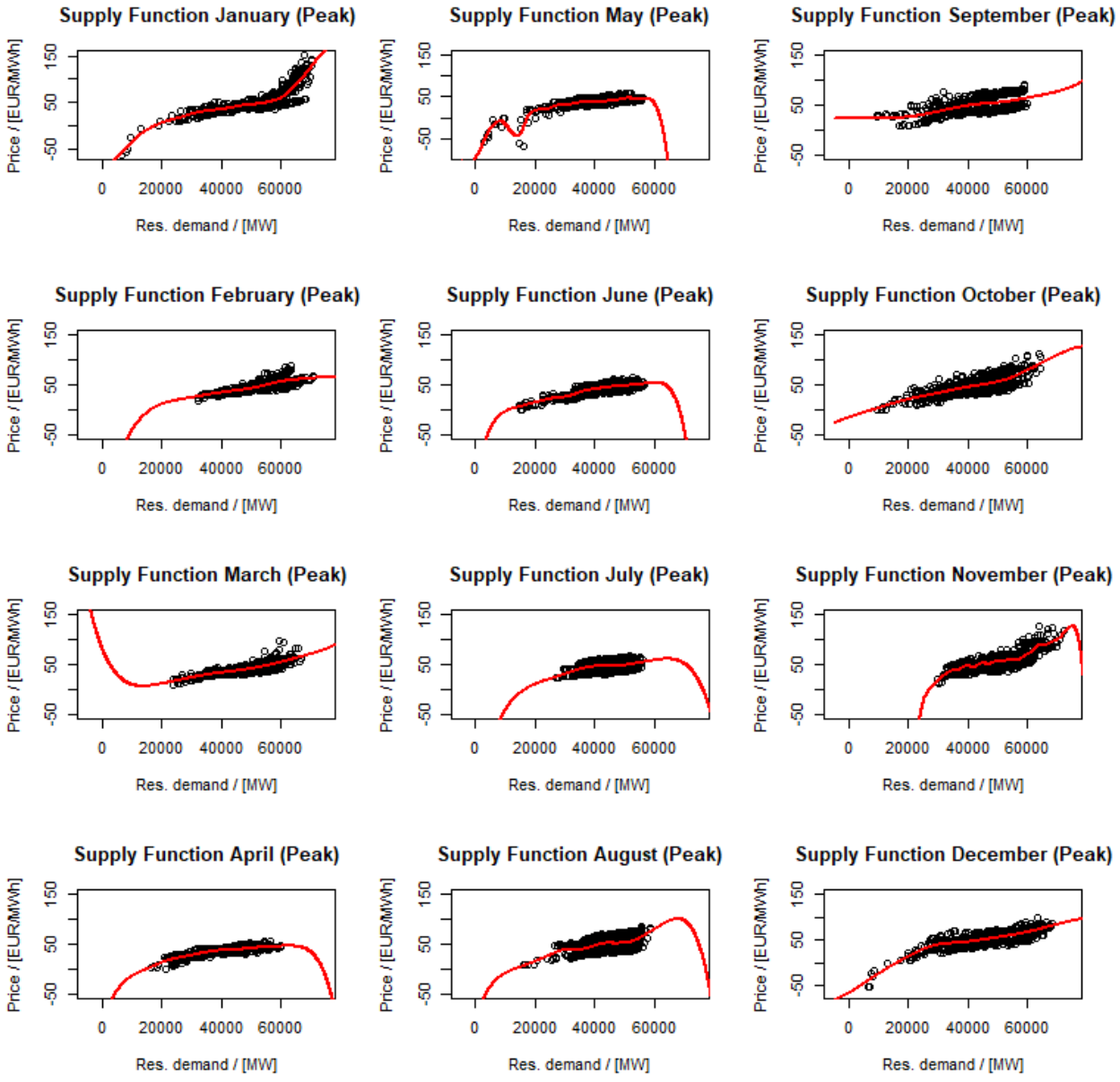

Figure 29: Monthly supply functions with smoothing spline estimation for peak hours

The following Figure 30 displays monthly supply functions estimated with smoothing splines for offpeak hours:

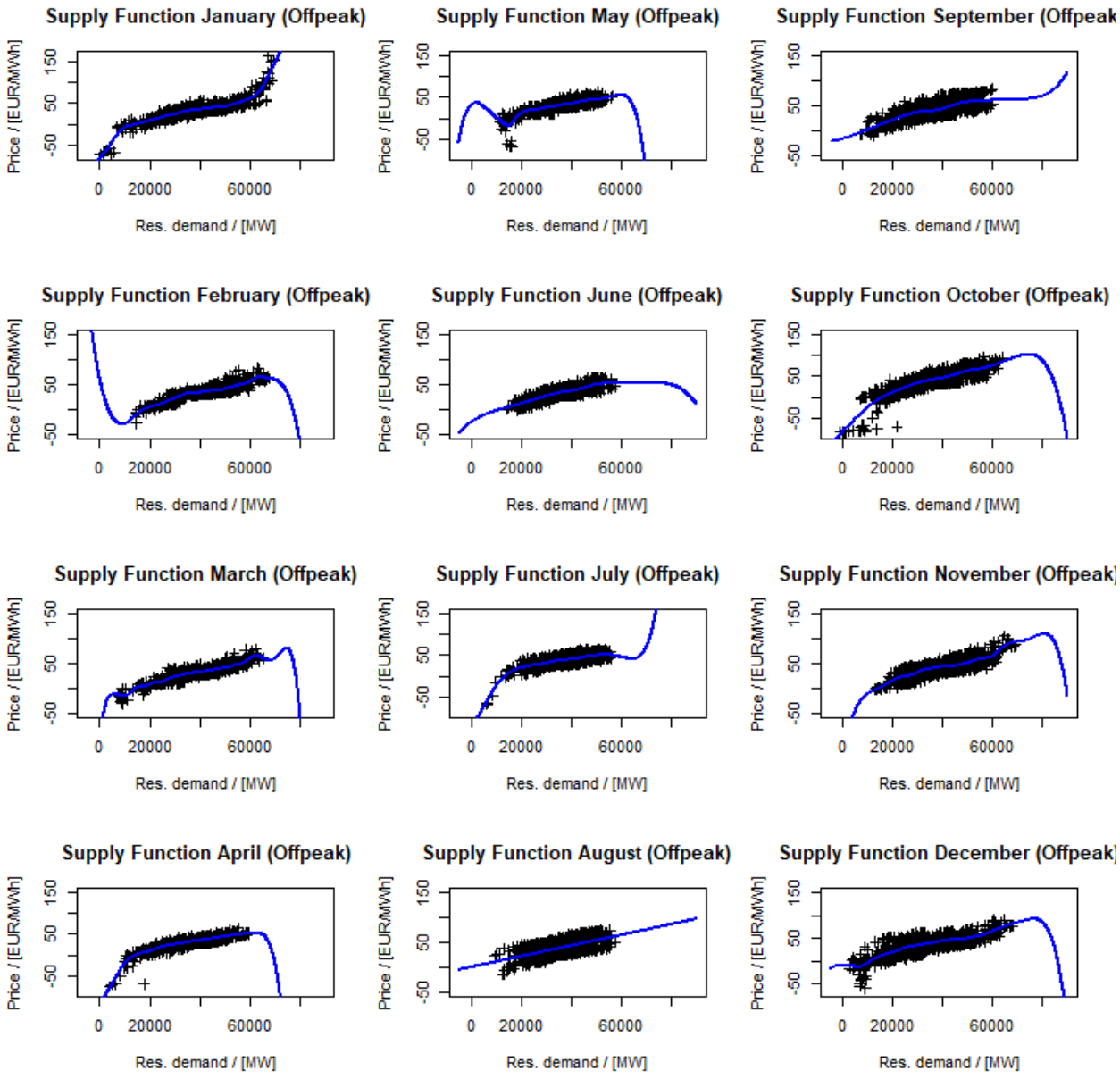

Figure 30: Monthly supply functions with smoothing spline estimation for offpeak hours

Figure 31 displays on the top the time series of simulated day-ahead prices using the monthly supply curve estimations and observed residual demand (red) as well as the observed day-ahead prices (black) from 01.01.2017 to 31.12.2018. The bottom of Figure 33 depicts the corresponding QQ-plots.

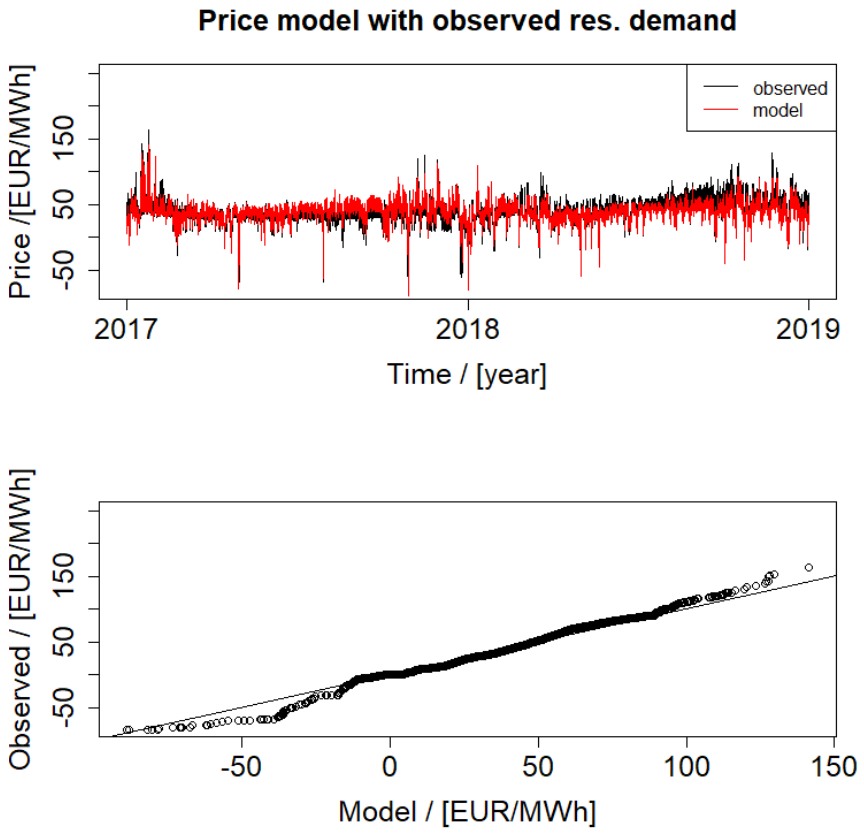

Figure 31: Time series and QQ-plot of supply curve estimation with observed residual demand and without residual volatility process

The figure depicts the time series (top) and QQ-plots (bottom) of simulated and observed day-ahead prices from 01 January 2017 to 31 December 2018 using for the simulation the described supply curve estimation methodology with the observed residual demand and without a residual volatiltiy process in the respective time period.

Table 6 depcits the parameters of the cubic smoothing spline estimations:

	df	λ	spar	ratio	CV crit.	PV crit.	Adj. R^2
Jan. (peak)	11.56	$7.95 \cdot 10^{-7}$	0.16	0.16	156.46	0.16	0.81
Feb. (peak)	5.69	$2.63 \cdot 10^{-5}$	0.37	0.37	49.32	0.37	0.63
Mar. (peak)	5.01	$4.66 \cdot 10^{-5}$	0.40	0.40	42.80	0.40	0.67
Apr. (peak)	8.95	$3.22 \cdot 10^{-6}$	0.24	0.24	21.61	0.24	0.72
May (peak)	24.41	$2.61 \cdot 10^{-8}$	-0.05	-0.05	44.95	-0.05	0.82
Jun. (peak)	9.29	$2.68 \cdot 10^{-6}$	0.23	0.23	40.29	0.23	0.72
Jul. (peak)	5.20	$3.78 \cdot 10^{-5}$	0.39	0.39	85.27	0.39	0.16
Aug. (peak)	8.35	$3.35 \cdot 10^{-6}$	0.24	0.24	176.98	0.24	0.26
Sep. (peak)	5.45	$2.92 \cdot 10^{-6}$	0.37	0.37	196.04	0.37	0.30
Oct. (peak)	6.22	$1.51 \cdot 10^{-5}$	0.33	0.33	151.38	0.33	0.56
Nov. (peak)	19.33	$1.15 \cdot 10^{-7}$	0.04	0.04	117.88	0.04	0.67
Dec. (peak)	7.66	$5.73 \cdot 10^{-6}$	0.27	0.27	117.68	0.27	0.69
Jan. (offpeak)	18.65	$1.15 \cdot 10^{-7}$	0.04	0.04	62.65	0.04	0.88
Feb. (offpeak)	19.17	$1.15 \cdot 10^{-7}$	0.04	0.04	34.21	0.04	0.83
Mar. (offpeak)	25.56	$2.86 \cdot 10^{-8}$	-0.04	-0.04	29.21	-0.04	0.85
Apr. (offpeak)	14.79	$3.12 \cdot 10^{-7}$	0.10	0.10	35.17	0.10	0.83
May (offpeak)	17.74	$1.49 \cdot 10^{-7}$	0.06	0.06	54.33	0.06	0.71
Jun. (offpeak)	7.95	$4.98 \cdot 10^{-6}$	0.27	0.27	38.48	0.27	0.73
Jul. (offpeak)	13.89	$3.66 \cdot 10^{-7}$	0.11	0.11	73.13	0.11	0.57
Aug. (offpeak)	2.00	$4.10 \cdot 10^{3}$	1.50	1.50	169.75	1.50	0.34
Sep. (offpeak)	7.39	$6.69 \cdot 10^{-6}$	0.28	0.28	145.00	0.28	0.46
Oct. (offpeak)	9.12	$2.55 \cdot 10^{-6}$	0.23	0.23	168.50	0.23	0.77
Nov. (offpeak)	15.25	$2.90 \cdot 10^{-7}$	0.10	0.10	104.63	0.10	0.63
Dec. (offpeak)	14.54	$3.77 \cdot 10^{-7}$	0.11	0.11	160.39	0.11	0.66

Table 6: Hyper parameters of monthly supply curve estimated with smooting splines

The table shows the hyper parameters and valuation criteria of the monthly supply curve estimated with smoothing splines in the years 2017 and 2018.

E Modeling Residual Volatility

We consider various model specifications to model the residual volatility process $\sigma_t = s_t - f_t(\hat{d}_t)$. After experimenting with different lags and specifications we found a SARIMA(2,0,2) x $(1,0,1)_{24}$ model to be the superior one (in terms of AIC). Likewise Gaussian GARCH noise was incapable of capturing the tails of the distribution appropriately causing us to instead opt for student-t distributed error terms.

The formal discription of the model for the residual volatility process σ_t thus reads as follows:

$$\phi(B)\Phi(B^{24})\sigma_t = \theta(B)\Theta(B^{24})\epsilon_t \tag{10}$$

with

$$\begin{aligned}\epsilon_t &= \eta_t \xi_t \\ \xi_t^2 &= \omega_0 + \omega_1 \epsilon_{t-1}^2 + \omega_2 \xi_{t-1}^2, \quad \omega_i \in \mathbb{R}^+, \quad i = 0, 1, 2\end{aligned}$$

where $\omega_0 \geq 0, \omega_1 > 0, \omega_2 > 0, \omega_1 + \omega_2 < 1, \phi(z) = 1 - \phi_1 z - \phi_2 z^2, \Phi(z) = 1 - \Phi_1(z), \theta(z) = 1 - \theta_1(z) - \theta_2 z^2, \Theta(z) = 1 - \Theta_1(z), \phi_1, \phi_2, \Phi_1, \theta_1, \theta_2, \Theta_1 \in \mathbb{R}$, B corresponds to the backshift operator ($B^j x_t = x_{t-j}$), and η_t is a student-t distributed random variable with v degrees of freedom. We can write (10) explicitly

$$\begin{aligned}\sigma_t = \phi_1 \sigma_{t-1} + \phi_2 \sigma_{t-2} + \Phi_1 \sigma_{t-24} - \Phi_1(\phi_1 \sigma_{t-25} + \phi_2 \sigma_{t-26}) \\ + \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-1} - \Theta_1 \epsilon_{t-24} + \Theta_1(\theta_1 \epsilon_{t-25} + \theta_2 \epsilon_{t-26})\end{aligned}$$

making it more clear how lagged and contemporaneous values are related to each other. The model is estimated by maximum likelihood and results are shown in Table 7.

To evaluate the corresponding methodology of smoothing splines estimation (monthly, peak and offpeak) with a SARIMA residual volaitlity process using GARCH noise, we inspect the performance using the observed residual demand values as input and compare day-ahead prices modeled with 1000 trajectories with the observed day-ahead prices. Please note that the simulated price distribution (1000 trajectories) is winsorized at the 99.9 % / 0.1 % quantile. Figure 32 inspects the time series of day-ahead prices, Figure 33 reveals the QQ-plots, Figure 34 compares the density functions, and Table 8 displays the statistical moments.

A. SARIMA	ϕ_1	ϕ_2	Φ_1	θ_1	θ_2	Θ_1
	1.6384	-0.6454	0.9745	-0.8920	-0.0287	-0.8953
	$(0.0162)^{***}$	$(0.0032)^{***}$	$(0.0014)^{***}$	$(0.0175)^{***}$	$(0.0112)^{*}$	$(0.0063)^{***}$
B. GARCH	ω_0	ω_1	ω_2	v		
	1.6973	0.3057	0.5905	4.0829		
	(0.6333)	$(0.0713)^{***}$	$(0.5905)^{***}$	$(0.1737)^{***}$		

Table 7: Parameter estimates of the residual volatility model

The table shows parameter estimates and standard errors (in parantheses) for SARIMA(2,0,2) x $(1,0,1)_{24}$-GARCH(1,1) model. Estimates are obtained by maximum likelihood. ***, **, and * denotes statistical significane at the 0.1%, 1%, and 5% level.

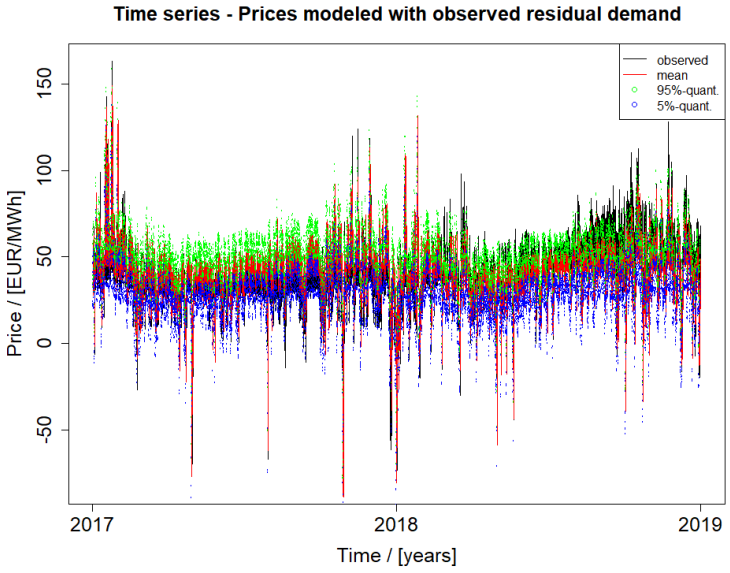

Figure 32: Time series comparison of observed spot prices and modeled spot prices using supply curve methodology with observed residual demand

The figure exhibits the times series of observed day-ahead prices (black) with the corresponding modeled day-ahead prices (mean and quantiles) using the described supply curve estimation methodology with the observed residual demand from 01 January 2017 to 31 December 2018.

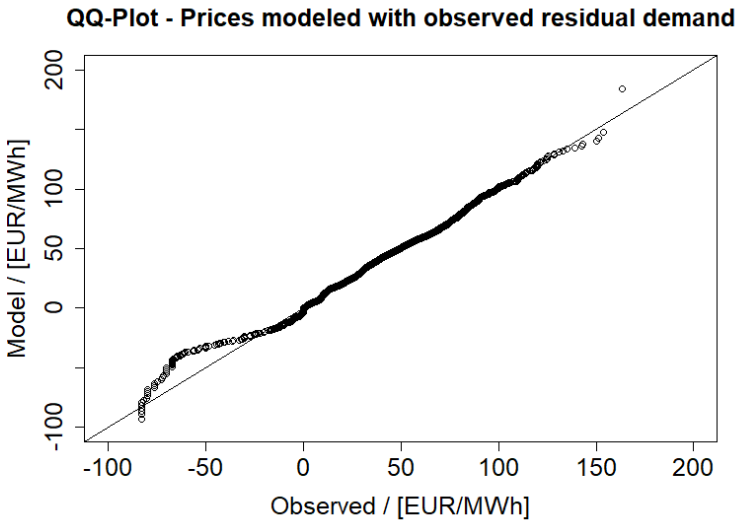

Figure 33: QQ-plot of observed spot prices and modeled spot prices using supply curve methodology with observed residual demand

The figure depicts the QQ-plot of observed (x-axis) vs. simulated (y-axis) day-ahead prices from 01 January 2017 to 31 December 2018 using the described supply curve estimation methodology with the observed residual demand in the respective time period.

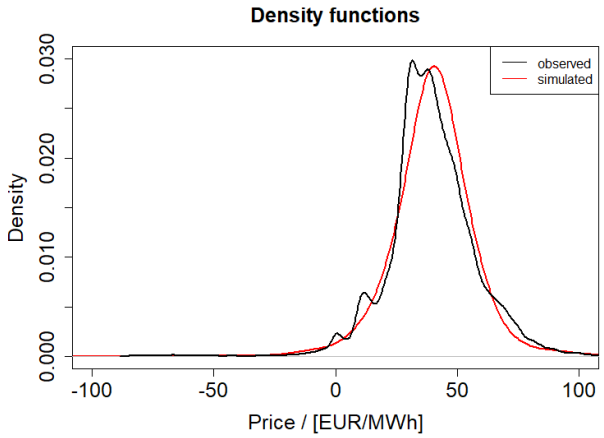

Figure 34: Density function of of observed spot prices and modeled spot prices using supply curve methodology with observed residual demand

The figure depicts the density functions of the observed and the simulated day-ahead prices from 01 January 2017 to 31 December 2018 using the described supply curve estimation methodology with the observed residual demand in the respective time period.

	Observed	Model
Mean	39.36	39.91
Variance	339.90	314.90
Skewness	-0.23	-0.27
Kurtosis	5.04	4.45

Table 8: Statistical moments of of observed spot prices and modeled spot prices using supply curve methodology with observed residual demand

The table shows the first four statistical moments of the observed and the simulated day-ahead prices from 01 January 2017 to 31 December 2018 using the described supply curve estimation methodology with the observed residual demand in the respective time period.

F Market-Wide Demand in Germany

As expected, demand in Germany exhibits pronounced seasonal patterns. Figure 35 shows the hourly time series of demand for Germany from beginning 2015 until the end of 2018 revealing several salient patterns and features. First, demand tends to be considerably higher during winter seasons, which can mainly be attributed to an increased demand for electrical light. Additionally, sharp drops in consumption levels for winter holidays (24th of December until about the 4-6th of January) are noticeable caused by the seasonal shutdown of major production facilities in the car manufacturing business as well as of steel mills and aluminum smelters.

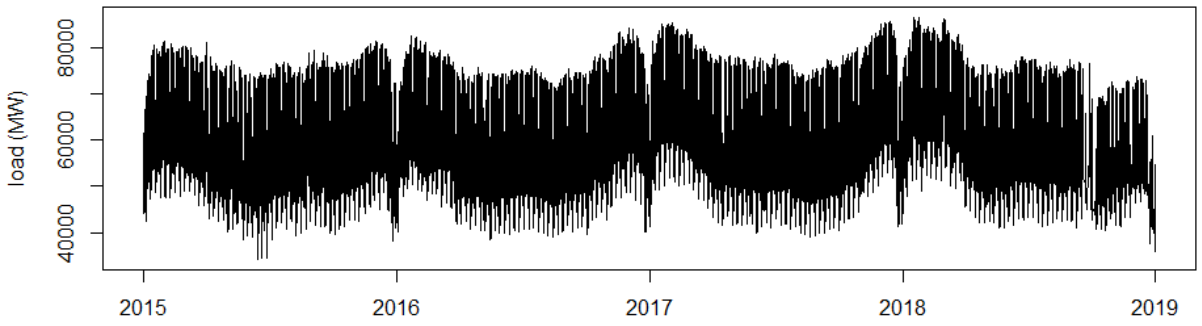

Figure 35: Yearly electricity demand variation

The figure depicts the hourly time series of (expected) load in the day-ahead market from 01 January 2015 to 31 December 2018.

The left graph of Figure 35 offers a more detailed look at the weekly seasonality. Demand

usually peaks in the first half of the week and drops considerably on the weekend with Sunday generally having the lowest demand levels overall. The intraday variation of electricity consumption is very large with changes of almost 50 % during the day (see right graph of Figure 36). Consumption usually rises significantly in the morning hours when people get up and prepare for work and peaks during lunch time and later in the evening (rush hour).

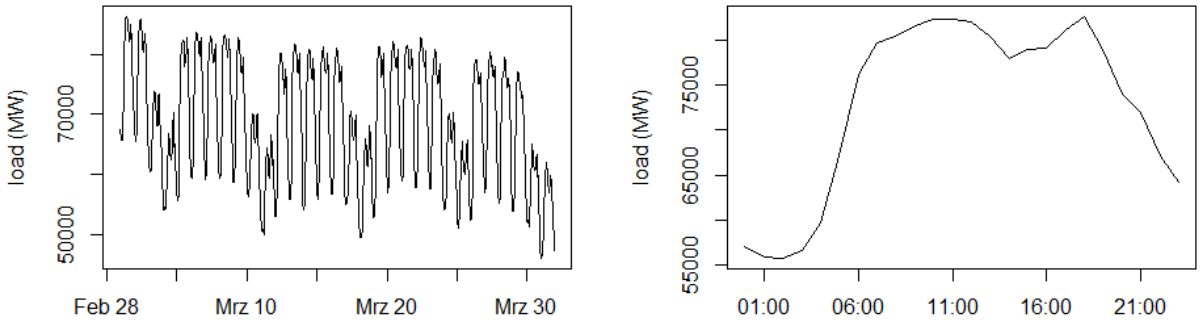

Figure 36: Weekly and intraday pattern of electricity demand

The left graph shows a detailed view of (expected) load in March 2018 whereas the right graph depict the timeseries of load for the 5th March, 2018.

Following Burger et al. (2003) we model demand with a flexible trend function paired with a SARIMA(2,0,2) x $(1,0,1)_{24}$ component. This results in the following model specification:

$$d_t = \mu_t^d + \psi_t$$
$$\hat{\phi}(B)\hat{\Phi}(B^{24})\psi_t = \hat{\theta}(B)\hat{\Theta}(B^{24})\epsilon_t$$

where μ_t^d corresponds to the trend-function with indicator variables for hourly-, weekly-, monthly-, and holiday-patterns, $\hat{\phi}(z) = 1 - \hat{\phi}_1 z - \hat{\phi}_2 z^2$, $\hat{\Phi}(z) = 1 - \hat{\Phi}_1 z$, $\hat{\theta}(z) = 1 - \hat{\theta}_1(z) - \hat{\theta}_2(z^2)$, $\hat{\Theta}(z) = 1 - \hat{\Theta}_1(z)$, $\hat{\phi}_1, \hat{\phi}_2, \hat{\Phi}_1, \hat{\theta}_1, \hat{\theta}_2, \hat{\Theta}_1 \in \mathbb{R}$, B corresponds to the backshift operator ($B^j x_t = x_{t-j}$), and ϵ_t is a normally distributed random variable. The model estimated by maximum likelihood and results can be found in Table 11.

$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\Phi}_1$	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\Theta}_1$
1.6360	-0.6461	0.9697	-0.5959	-0.2166	-0.8655
$(0.0251)^{***}$	$(0.0236)^{***}$	$(0.0029)^{***}$	$(0.0247)^{***}$	$(0.0077)^{***}$	$(0.0053)^{***}$

Table 9: Parameter estimates of the demand model
The table shows parameter estimates and standard errors (in parantheses) for SARIMA(2,0,2) x $(1,0,1)_{24}$ model. Estimates are obtained by maximum likelihood. ***, **, and * denotes statistical significance at the 0.1%, 1%, and 5% level.

G Reduced-Form Model Specification and Calibration

Given the occurrence of negative prices for the case of electricity, arithmetic spot price models have gained wide acceptance during recent years. We follow Benth et al. (2007), Meyer-Brandis and Tankov (2008), and Benth et al. (2013) and model hourly day-ahead spot prices by means of a multi-factor spike model specified as follows:

$$s_t = \mu_t^s + x_t + y_t \tag{11}$$

with

$$\begin{aligned} dx_t &= -\kappa_x x_t d_t + \sigma_x dw_t, \quad \kappa_x, \sigma_x \in \mathbb{R}^+ \\ dy_t &= -\kappa_y y_t d_t + dn_t, \quad \kappa_y \in \mathbb{R}^+ \end{aligned}$$

where μ_t^s corresponds to the deterministic seasonality function which accounts for trends over the year, weekday- and the pronounced intraday-pattern. x_t is a standard Vasicek mean-reversion process capturing "normal" price variations whereas y_t is a spike process to account for rare price spikes. $d\omega_t$ is a standard brownian motion and dn_t a compound Poisson process with constant jump intensity λ and double-exponentially distributed jump sizes z_t, i.e. exponentially distributed negative as well as positive jumps with density:

$$f_z(x) = p\beta_1 e^{-\beta_1 x} \mathbb{1}_{x<0} + (1-p)\beta_2 e^{-\beta_2 x} \mathbb{1}_{x>0}, \quad p, \beta_1, \beta_2 \in \mathbb{R}^+$$

where $\beta_1, \beta_2 > 0$ and $p \in [0,1]$. $\beta_1(\beta_2)$ controls the shape of the positive (negative) jump size distribution whereas p determines the fraction of positive and negative spikes respectively.

The calibration of the above model specification necessitates the identification of latent state variables x_t and y_t as well as model parameters governing the Gaussian and non-Gaussian

mean-reversion process. Following Cartea and Figueroa (2005), Benth et al. (2007), and Bieger-König (2013) we identify spikes by a recursive filtering algorithm which simultaneously estimates the trend function μ_t^s as well as the time series of y_t. Given the values of y_t and μ_t^s we can infer the time series of the Gaussian mean-reversion process x_t. Model parameters can then be estimated by means of standard econometric techniques.

Figure 37 shows the time series of filtered state variables whereas Table 10 summarizes the paramter estimates. The distribution of x_t exhibits low skewness (-0.0931) and excess kurtosis (4.1476). The filtering algorithm thus works quite well in identifying heavy tailed price movements allowing the disentaglement of both processes and estimation of parameter values afterwards. We find an hourly spike intensity of 0.6 % which corresponds to roughly 52 spikes annually on average. Furthermore, positive price spikes are less likely to occur with $p \approx 0.37$ and have smaller magnitude (+ 52 EUR/MWh vs. -77 EUR/MWh on average). This stands in contrast to studies using data before 2010. For instance, Bieger-König (2013) finds that positive spikes are slightly larger and almost three times as frequent. An effect which we attribute to the growing presence of renewable generation.

κ_x	σ_x	κ_y	λ	p	β_1	β_2
0.1140	11.4915	0.2778	0.0060	0.3714	0.0217	0.0145

Table 10: Parameter estimates of the reduced-form power price model

The table shows parameter estimates for the reduced-form power price model given by (11).

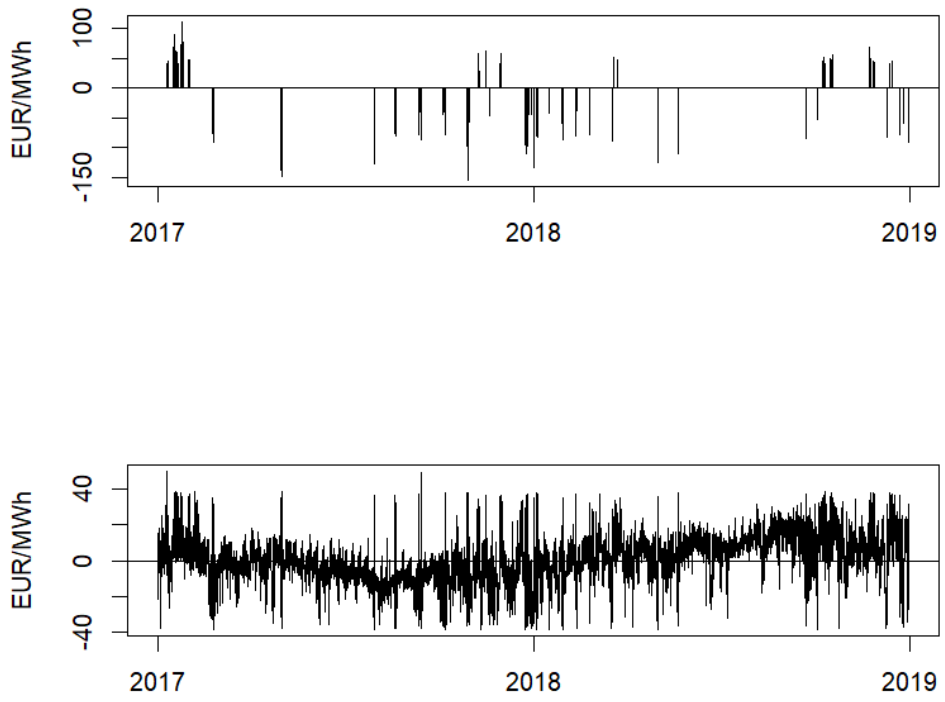

Figure 37: Filtered state variables

The top graph shows the filtered Non-Gaussian process whereas the bottom graph depicts the Gaussian counterpart.

H Evaluation of Model Performance

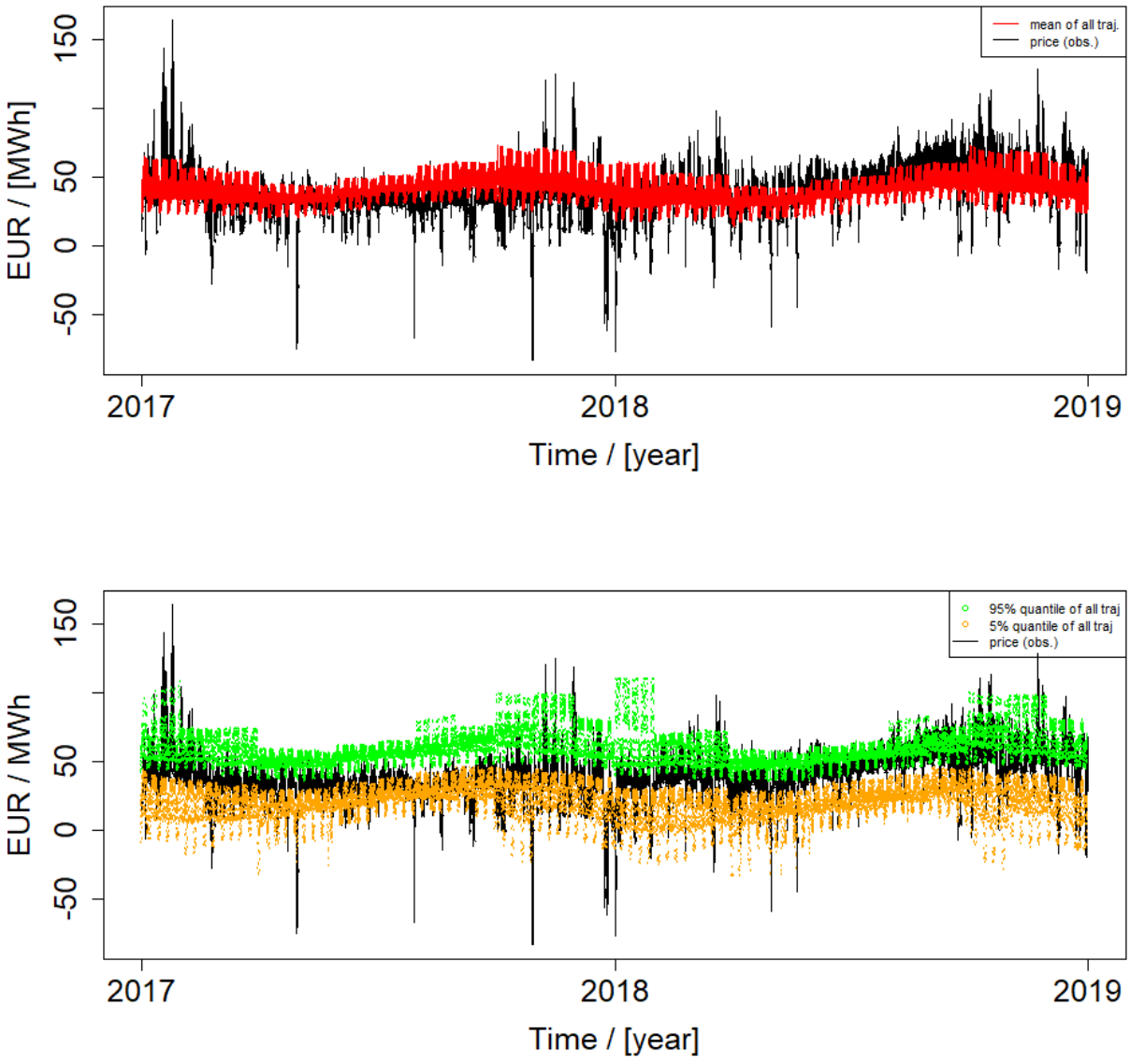

Figure 38: Time series of modeled and observed day-ahead prices

The top graph depicts the time series of the observed (black) and the simulated means of all trajectories in the same hour (red) from 01 January 2017 to 31 December 2018. The lower graph shows the time series of the observed (black) and simulated 95 % quantile (green) and 5 % quantile of all trajectories in the same hour (orange) from 01 January 2017 to 31 December 2018.

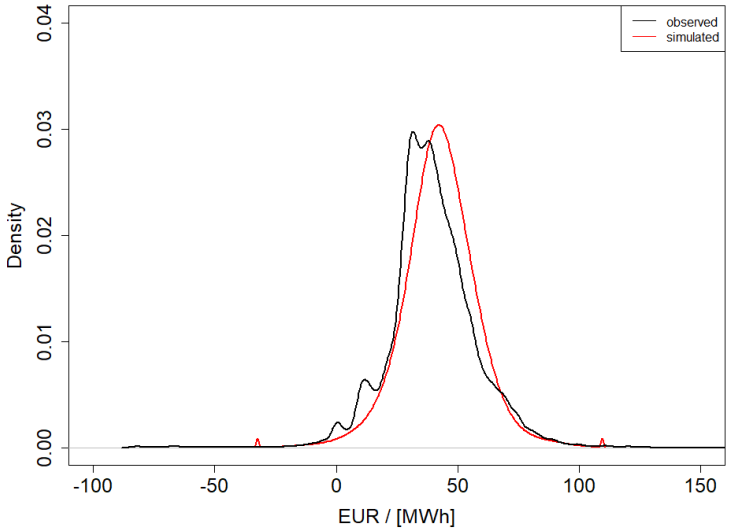

Figure 39: Density functions of modeled and observed day-ahead prices

The top graph depicts the density functions of the observed (black) and simulated (red) from 01 January 2017 to 31 December 2018.

	Observed	Model
Mean	42.67	39.33
Variance	340.27	244.38
Skewness	-0.23	-0.06
Kurtosis	8.02	5.04

Table 11: Statistical moments of modeled and observed day-ahead price distributions

The table shows the first four statistical moments of the observed and the simulated day-ahead prices from 01 January 2017 to 31 December 2018.

	df	λ	*spar*	*ratio*
observed	4	$1.27 \cdot 10^{-4}$	0	0
simulated	4	$1.27 \cdot 10^{-4}$	0	0

Table 12: Spline Interpolation Parameters

The table shows the parameter specification in the spline interpolation of Figure 11 using smoothing splines.

Working Paper Series in Production and Energy

recent issues

No. 65 Martin Klarmann, Robin Pade, Wolf Fichtner, Nico Lehmann: Energy Behavior in Karlsruhe and Germany

No. 64 Florian Zimmermann, Dogan Keles: State or Market: Investments in New Nuclear Power Plants in France and Their Domestic and Cross-border Effects

No. 63 Paul Heinzmann, Simon Glöser-Chahoud, Nicolaus Dahmen, Uwe Langenmayr, Frank Schultmann: Techno-ökonomische Bewertung der Produktion regenerativer synthetischer Kraftstoffe

No. 62 Christoph Fraunholz, Kim K. Miskiw, Emil Kraft, Wolf Fichtner, Christoph Weber: On the Role of Risk Aversion and Market Design in Capacity Expansion Planning

No. 61 Zoe Mayer, Rebekka Volk, Frank Schultmann: Evaluation of Building Analysis Approaches as a Basis for the Energy Improvement of City Districts

No. 60 Marco Gehring, Franziska Winkler, Rebekka Volk, Frank Schultmann: Projektmanagementsoftware und Scheduling: Aktuelle Bestandsaufnahme von Funktionalitäten und Identifikation von Potenzialen

No. 59 Nico Lehmann, Jonathan Müller, Armin Ardone, Katharina Karner, Wolf Fichtner: Regionalität aus Sicht von Energieversorgungs- und Direktvermarktungsunternehmen – Eine qualitative Inhaltsanalyse zu Regionalstrom in Deutschland

No. 58 Emil Kraft, Marianna Russo, Dogan Keles, Valentin Bertsch: Stochastic Optimization of Trading Strategies in Sequential Electricity Markets

No. 57 Marianna Russo, Emil Kraft, Valentin Bertsch, Dogan Keles: Short-term Risk Management for Electricity Retailers Under Rising Shares of Decentralized Solar Generation

No. 56 Anthony Britto, Joris Dehler-Holland, Wolf Fichtner: Optimal Investment in Energy Efficiency as a Problem of Growth-Rate Maximisation

No. 55 Daniel Fett, Christoph Fraunholz, Dogan Keles: Diffusion and System Impact of Residential Battery Storage under Different Regulatory Settings

No. 54 Joris Dehler-Holland, Marvin Okoh, Dogan Keles: The Legitimacy of Wind Power in Germany

No. 53 Florian Diehlmann, Markus Lüttenberg, Lotte Verdonck, Marcus Wiens, Alexander Zienau, Frank Schultmann: Public-Private Collaborations in Emergency Logistics: A Framework based on Logistical and Game-Theoretical Concepts

Impressum

Karlsruher Institut für Technologie

Institut für Industriebetriebslehre und Industrielle Produktion (IIP)
Deutsch-Französisches Institut für Umweltforschung (DFIU)

Hertzstr. 16
D-76187 Karlsruhe

KIT – Universität des Landes Baden-Württemberg und
nationales Forschungszentrum in der Helmholtz-Gemeinschaft

Working Paper Series in Production and Energy
No. 66, October 2022

ISSN 2196-7296

www.iip.kit.edu