Gianfreda, A., Ravazzolo, F., & Rossini, L. (2023). Large Time-Varying Volatility Models for Hourly Electricity Prices. Oxford Bulletin of Economics and Statistics, 85(3), 545-573. https://doi.org/10.1111/obes.12532

Large Time-Varying Volatility Models for Hourly Electricity Prices*

Angelica Gianfreda[a,d] Francesco Ravazzolo[b,e] Luca Rossini[c]

[a]University of Modena and Reggio Emilia, Modena, Italy [b]BI Norwegian Business School, Norway

[c]University of Milan, Italy [d]EMG, London Business School, UK

[e] Free University of Bozen-Bolzano, Italy

October 27, 2022

Abstract

We study the importance of time-varying volatility in modeling hourly electricity prices when fundamental drivers are included in the estimation. This allows us to contribute to the literature of large Bayesian VARs by using well-known time series models in a large dimension for the matrix of coefficients. Based on novel Bayesian techniques, we exploit the importance of both Gaussian and non-Gaussian error terms in stochastic volatility. We find that using regressors as fuel prices, forecasted demand and forecasted renewable energy is essential to properly capture the volatility of these prices. Moreover, we show that the time-varying volatility models outperform the constant volatility models in both the in-sample model-fit and the out-of-sample forecasting performance.

Keywords: Modelling; Forecasting BVARs; Gaussian Stochastic Volatility; Non-Gaussian Stochastic Volatility; Fossil Fuels; Renewable Energy Sources; Electricity Demand; Professional Forecasts.

*The authors thank three anonymous referee, the associate editor and the editor for the useful comments. The authors gratefully acknowledge Matteo Iacopini and Jamie Cross for their useful feedback. This paper is part of the research activities at the Free University of Bozen-Bolzano, funded by Europe Energy S.p.A., and at the Centre for Applied Macroeconomics and Commodity Prices (CAMP) at the BI Norwegian Business School. Angelica Gianfreda acknowledges the financial support for the project *Energy Risk Modelling under UNcertainties*, funded by the Free University of Bozen-Bolzano.

1 Introduction

Electricity is a *non-* or *partially* storable commodity, it must be produced when and where demanded. The impossibility to store economically electricity and the variability introduced into the system by new regulations and the imperfect predictability of fundamental drivers are sources of uncertainties reflected in electricity prices and their volatility. Knittel and Roberts (2005) and Escribano et al. (2011) provide a detailed overview of all stylized facts of electricity prices, like multiple seasonality, mean reversion, short-lived *spikes* (or sudden jumps in prices), and high volatility. These characteristics do originate from the convex supply curve, the price inelastic demand in the short-run, and the limited storability of electricity.

In recent years, worldwide energy policies have supported, and they are still fostering, green generation to reduce carbon emissions and mitigate climate change. The increasing share of electricity generated from renewable energy sources (RES) has a twofold effect. On one hand, the equilibrium price is lowered because RES, entering the supply curve before other more costly technologies, shift the supply curve towards the right. And, prices can become negative when extremely high RES generation is coupled with low levels of demand, as in Germany. On the other hand, RES (as wind, solar, and hydro to less extent) add complexity to the market because they are variable, intermittent, and not easily predictable since they are strictly related to weather conditions. If the wind blows and/or the sun shines, green and economic generation satisfies the demand, and electricity prices are low (or negative); otherwise, demanded electricity is covered with more expensive thermal conventional plants running with fossil fuels.

Obviously, uncertainties affect also the demand and its forecasts. Indeed, it is known that demand reacts non linearly to temperature, as emphasized by Henley and Peirson (1997) and more recently by Damm et al. (2017). Consequently, there are recent and increasing concerns about demand and supply varying over time with greater uncertainty and, then, amplifying the electricity price features and volatility patterns.

The vast majority of the literature on electricity prices has considered univariate models for individual hours of the day with possible time–varying volatility patterns, asymmetries, and shocks induced by fundamental drivers. Koopman et al. (2007), Huurman et al. (2012), Gianfreda and Grossi (2012), Paraschiv et al. (2014), Ketterer (2014), Frömmel et al. (2014), Chan and Grant (2016), Erdogdu (2016), Ciarreta and Zarraga (2016), Jeon and Taylor (2016) and Laporta et al. (2018), among many others, implement univariate GARCH–type specifications for individual hourly or daily averaged price series. Only a few papers have considered cross-time dependencies. For instance, Maciejowska and Nowotarski (2016) and Ziel (2016) observe that prices for early morning hours depend more on the latest information than on information contained at the same hour of the previous day. This is an indication of cross-hourly dependence among prices which raises doubts about cross-relations in the volatility dynamics: we aim at exploring this issue employing Bayesian VAR models. Other proper multivariate specifications have been explored by Bauwens et al. (2013), Efimova and Serletis (2014) and Raviv et al. (2015). However, closer to our approach is the analysis undertaken by Kostrzewski and Kostrzewska (2019) on the PJM market. They compare univariate Bayesian stochastic volatility models with a double exponential distribution of jumps and exogenous variables with non-Bayesian ones, and specifications with and without exogenous variables as indicators for jumps, special days (that is Mondays, Saturdays, and Sundays), the minimum hourly price observed over the previous day, and finally temperatures - the latter ones being a proper exogenous variable. Instead, we adopt a multivariate setting in which true exogenous variables are related to both demand and supply curves. Specifically, we test the effects of uncertain forecasted demand levels (produced accounting also for temperatures, as it will be described later), uncertain forecasted renewable power generated from both wind and solar photovoltaic units, and additionally fossil fuel prices. To the best of our knowledge, the analysis of multivariate models with time-varying volatility for electricity prices, demand, RES, and fossil fuels has not been deeply investigated and we aim at filling this gap by adopting Bayesian approaches. This contribution is an extension of Gianfreda et al. (2020) by adding stochastic volatility and then contributing to the limited results on multivariate volatility models. In that paper, the authors compare several univariate and multivariate frequentist and Bayesian models augmented with fundamental variables (demand, RES, fossil fuels, and seasonality) to predict hourly day-ahead electricity prices. Therefore, following Gianfreda et al. (2020), we consider Bayesian Vector Autoregressive (BVAR) representations since better forecasting performances emerge from multivariate models given the larger information contained, as suggested by Stock and Watson (2002).

The availability of large electricity data sets referring to market system information allows scholars and researchers to implement models previously used for macroeconometrics, see Mumtaz and Zanetti (2015); Huber and Feldkircher (2019); Koop et al. (2019); Huber et al. (2020); Chan (2020a) among others.

The most challenging aspect for the use of multivariate time series models is the amount of information available and this feature is actually needed to include all relevant information in modelling and forecasting electricity prices. It is well-known that they show strong dependencies from their past values (given the highly repetitive nature of electricity auctions underlying the determination of these prices) and strong seasonality that from demand is transferred into prices. Indeed, demand is heavily influenced by industrial activities with different dynamics over days of the week and months of the year. Hence, modelling these hourly prices, their past values (observed on one, two, and seven days before), seasonality, exogenous variables for supply (that is RES and fossil fuels) and demand requires the handling of a matrix of coefficients of size (162×24), given the 24 hourly prices with 3 lags, 12 monthly dummies, two dummies for Saturdays and Sundays and one dummy for holidays, plus hourly (forecasted) demand and RES (wind and solar), and 3 fossil fuels (coal, gas, and CO_2).

One can state that traditional factor models are successfully used to handle large datasets, however, the recent literature in large Bayesian Vector Autoregressive models (VAR) has provided valid and important alternatives. In the macroeconomic literature (Mumtaz and Zanetti, 2013; Clark and Ravazzolo, 2015; Carriero et al., 2016; Chiu et al., 2017; Carriero et al., 2019; Chan, 2020a), the use of time-varying volatility is known to improve the full sample analysis by capturing the peak and booms and also to improve the forecasting accuracy.

The large dimensionality of time-varying volatility models applied to electricity hourly data[1] can result in sizeable estimation errors. To mitigate this, we follow recent macroeconomic literature and apply Bayesian estimation techniques. Contrarily to what happens in the macroeconomic analysis, we do not encounter any ordering problem in the lower triangular matrix of the structural VAR since we have a strong time dependence due to the natural ordering of consecutive hours.

From the computational and operational point of view, Carriero et al. (2019) and Chan (2020a) have recently developed Bayesian estimation methods for VARs with stochastic volatility that allow reducing the computational timing when the number of coefficients increases. Therefore, following the specification in Chan and Eisenstat (2018) and Cross et al. (2020), we use a Bayesian VAR with stochastic volatility for capturing the movements of the volatility in the electricity prices. Differently from Chan (2020a), we do not assume a stochastic volatility component constant across the variables, but we assume that the time-varying volatility changes across hours. Moreover, as

[1] We remember that our models require to estimate a (24×24) time-varying covariance matrix.

anticipated and differently from Carriero et al. (2019) and Cross et al. (2020), we do not have any ordering problems in the estimation, because our variables follow the consecutive time dependence structure of the 24 hours in a day.

The main finding in our paper refers to the importance of modeling multivariate time-varying volatility in the variation of the electricity prices and, in particular, the fact that the assumption of constant volatility on average overestimates the time-varying volatility over time, thus leading to imprecise estimation. Moreover, the inclusion of the fat tail error term in the stochastic volatility produces further improvements across central hours, where the estimation of different degrees of freedom is important. These findings are based on three explicative countries (Germany, Italy and Denmark) characterized by high RES penetration in the considered years during which all relevant forecasts are available to us. When forecasting German, Italian and Danish electricity prices from 2018 to 2019, we find evidence of strong improvements once forecasted demand and renewable energy sources are included in the analysis with respect to the baseline model with only a lag representation of the response variable. Furthermore, if we include time-varying volatility in the form of Gaussian or Student-t distributions, we observe strong improvements in Germany, Denmark and moderate improvements in Italy. These results are consistent in both point and density forecasting, and also when focusing on quantile density forecasting for tail comparisons.

Indeed, we additionally contribute to the literature of density forecasting in Panagiotelis and Smith (2008), Huurman et al. (2012), Jónsson et al. (2014), Gianfreda and Bunn (2018) and Gianfreda et al. (2020) by comparing several density metrics. Differently from Kostrzewski and Kostrzewska (2019) who use unconditional and conditional coverage tests together with the Diebold-Mariano tests, we assess the density forecasts through the continuous ranked probability scores (CRPS), their quantile-weighted variants as the averaged center quantiles (CQ-CRPS), and the averaged right and left tails quantiles (RQ-CRPS and LQ-CRPS). As argued by Gianfreda and Bunn (2018), price asymmetries induced by wind generation are significant and attention should be paid to their modeling. The presence of a fat tail in the stochastic volatility is particularly emphasized in Germany during the central hours and across the different considered metrics. This phenomenon is particularly evident in this market since prices are free to fluctuate from a floor price of -500 €/MWh to a cap price of 3000€/MWh, whereas in other countries the floor is generally set to zero.

The paper proceeds as follows. Section 2 describes the electricity market and data used. The methodology is described in Section 3 together with details on the estimation process. Results are presented in Section 4, which also contains the recursive out-of-sample forecasting performance of the Bayesian VAR models with stochastic volatility. Finally, Section 5 concludes and briefly discusses some future research directions.

2 Electricity Market: Prices and Drivers

Electricity is a commodity traded in wholesale markets, generally organized in different sequential sessions: from the day-ahead and intra-day to the balancing ones.

On a voluntary basis, bids to buy and offers to sell electricity are submitted in the day-ahead session for each hour of the following day, in pairs of prices and quantities by both consumption units and generators. This session opens several days in advance and closes at noon on the day before physical delivery. Then, individual bids are aggregated and supply offers are ordered giving the priority of dispatch to more efficient and less polluting units, characterized by lower marginal costs. Hence, wind, solar, and in general all renewable energy plants enter the supply curve before more costly thermal conventional or nuclear units; this is the so-called 'merit order criterion'. Subsequently, the equilibrium price is computed hourly under a cost minimising objective function and it is identified by the intersection of the aggregated curves of supply and demand. These 24 prices are often called *day-ahead*, *forward*, or *auction* prices because they are determined one day before delivery on the day-ahead market; and we refer to day-ahead and spot interchangeably. Electricity prices are prone to the microstructural features of the market, and so influenced by installed capacity, by the international dynamics of fossil fuel prices (in varying proportions related to the shares of thermal technologies in the generation mix), plant maintenance and outages, interconnections with foreign markets, and weather conditions influencing both demand and RES forecasts. Indeed, the intra-day sessions, taking place after the day-ahead market, allow all units to modify their day-ahead plans, since for instance sudden outages and/or changing weather conditions (among many other factors) may affect generators' planned schedules of production and consumers' forecasted consumption.

Finally, the balancing markets represent the last sessions where only generators with the required degree of flexibility are allowed to act, and where system security, grid stability, and the instantaneous match between demand and supply are granted by the transmission system operator. The prices determined on these sessions are proper real-time prices (usually called 'balancing' prices) and are regulated by different pricing mechanisms[2]; we would like to emphasize that we are not considering these prices but only the day-ahead ones.

Uncertainties affecting the supply curve are mainly concerning the intermittent production from renewable energy sources, the international movements of fossil fuel prices, and technical operations of plants and of the whole electricity system. Among these, the penetration of renewables represents the major source of uncertainty. Given their increasing shares through the years, electricity markets have become more and more weather dependent and prices are exposed to atmospheric conditions. Then also climate change is going to influence future price dynamics and exacerbate price variability. Weather conditions are fundamental factors influencing wind and solar generation, but also electricity demand, then focussing on the understanding of price volatility is becoming crucial in these markets for the management of operational risks and hedging strategies. Accurate price forecasts accounting for time-dependent volatility are then essentials in determining the amount of risk to cover.

To this aim, we consider the following fundamental data: day-ahead prices, forecasted demand, forecasted wind and solar PV generation, and fossil fuel prices. Specifically, we refer to three[3] European countries characterized by high penetration of renewables and different generation mixes, these are Germany, Italy and Denmark.

2.1 Data Description and Preliminary Analysis

Hourly day-ahead auction prices are directly collected from the corresponding power exchanges: the *European Energy Exchange* EEX for Germany[4], the *Gestore dei Mercati Energetici* GME for Italy[5], considering the single national Italian price (*prezzo unico nazionale*, PUN); and, the *NordPool* for Denmark[6]. These hourly electricity prices are quoted in €/MWh and have a daily frequency. They have been pre-processed for time-clock changes to exclude the 25th hour in October and to interpolate the missing 24th hour in March; hence, there are no missing observations. Their forward nature is important in understanding the right timing and the need for forecasted demand, wind, and solar quantities. Indeed, market operators run their forecasting models on day t to obtain a set of 24 prices (and quantities for delivery on the following day $t+1$) to be submitted before the

[2]Additional details and further insights on balancing prices can be found in Gianfreda et al. (2018) and Gianfreda et al. (2019).

[3]Please note that availability of forecasts may be an issue, hence this influenced our selection of studied countries.

[4]Precisely, we had access to the ftp from www.eex.com thanks to the *Europe Energy*

[5]http://www.mercatoelettrico.org

[6]http://www.nordpoolgroup.com

closure of the market (around noon of the same day t). Then, the information available refers to the same day t as far as the forecasted quantities are concerned (for demand, wind and solar), but also to the previous day $t-1$ when the settlement prices for fossil fuels are determined (and plant capacities are declared).

Therefore, hourly forecasted quantities are collected from Refinitiv. In detail, we use values forecasted by the *operational* weather model provided by the *European Centre for Medium-Range Weather Forecast*, EC. This model runs at midnight and updates from 05.40 a.m. to 06.55 a.m., hence providing the latest information available to market operators to prepare their bidding strategy to be submitted to the day-ahead market by noon. In case of missing or unavailable forecasts for a specific model (as for the Italian market), we adopted this strategy to reconstruct the full required series: when the forecasts from the EC model running at midnight (its acronym is ECop00) were not available, we replaced missing observations by those produced by other forecasting models according to their time of publication. In details, at the second step of replacement, we considered the forecasts provided by the *Global Forecast System* (GFS) running an *ensemble* model at midnight (that is GFSen00), alternatively in case of further missing observations we used the results of the *operational* model that was still running at midnight (that is GFSop00), and, if necessary, we use the same replacement scheme using respectively GFSen18, GFSop18, ECen12, according to the time of publication of their results. At the end of the process, minimal residual missing observations (at the beginning of the Italian sample) were replaced with interpolated values. The two weather models differ in terms of randomness and resolution. The *operational* model is a deterministic model with high resolution and no involved randomness, whereas the *ensemble* model is a probabilistic model with lower resolution but with random variations of the initial weather conditions. Therefore, the latter one simulates more weather instability by considering different weather scenarios.

The forecasted series of solar power production show an additional problem: they exhibit a block structure of null values in hours early in the mornings and late in the evenings, creating collinearity issues. Therefore, following Gianfreda et al. (2020), we have pre-processed these series using a linear transformation[7]. As fossil fuel prices are concerned, we use the settlement prices for coal (as for the Intercontinental Exchange API2 cost, insurance, and freight Amsterdam, Rotterdam and

[7]Draws from a Uniform distribution are generated and then added to the original zero values to obtain small numbers. This results in having (column) blocks of very small values close to but different from zero, instead of having (column) blocks of zeros.

Antwerp, with ticker LMCYSPT), natural gas (as for the one month forward ICE UK[8], with ticker NATBGAS) and CO_2 (as for the EEX-EU CO_2 Emissions E/EUA in €, with ticker EEXEUAS), downloaded from Datastream and all converted in €/MWh using the WMR&DS exchange rates from US$ to Euros (with ticker USEURSP) and from GBP to Euro (with ticker UKEURSP). They have a daily frequency and so a constant structure over the 24 hours; missing weekends and holidays have been interpolated. Finally, we use monthly and dummy variables for calendar and weekend seasonality, disentangling between Saturdays, Sundays and holidays.

Given that extreme prices affect price variability and that these spikes represent peculiar characteristics of the electricity markets incorporating important information (as arbitrage opportunities from a day–ahead trading perspective), we maintain them in our series[9], differently from Weron (2006) and Afanasyev and Fedorova (2019). Moreover, following Karakatsani and Bunn (2010) and Paraschiv et al. (2014), to avoid masking statistical price properties and volatility dynamics that we want to capture and model, we use price levels and not logarithmic prices as in Conejo et al. (2005), Garcia et al. (2005), Weron and Misiorek (2008), Bordignon et al. (2013) among others.

To summarize, we use hourly data for prices, forecasted demand, wind, and solar PV generation, together with repeated daily data (across the 24 hours) for fossil fuel prices from 01 January 2016 to 31 December 2019. Figure 1 shows the sample distributions of electricity prices, with high variability (outside the upper and lower quartiles) with different degrees across the three considered markets. We can observe outliers varying between ±200 €/MWh in Germany; between zero and 150€/MWh in Italy and between -50 €/MWh and 150€/MWh in Denmark. The sample distributions of forecasted demand, wind, and solar PV are reported in Figure 2. The historical evolution of prices and drivers is presented in the supplementary material, since the intra-daily dynamics help more in understanding the degree of uncertainties in demand and RES levels and show how their connected variability can influence hourly prices, with peculiar dynamics across the 24 hours and with stronger combined effects (of high variability in demand, wind, and solar) in peak hours. All together suggests the importance of considering fat-tailed and asymmetric distributions.

According to both point and density metrics, Gianfreda et al. (2020) recommend multivariate

[8]As suggested by Gianfreda et al., 2016, it represents a pure hub benchmark and can be used for all EU markets.

[9] We performed a robustness check setting outliers to upper and lower thresholds, determined as in BUNN2010 by using price sample means and three times their standard deviations. Results were qualitatively similar, hence omitted but available on request.

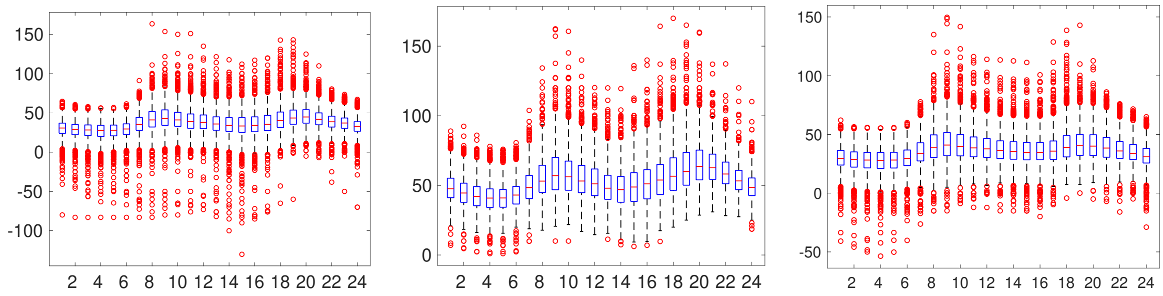

Figure 1: Sample Price Distributions for Germany (on the left), Italy (on the center) and Denmark (on the right).

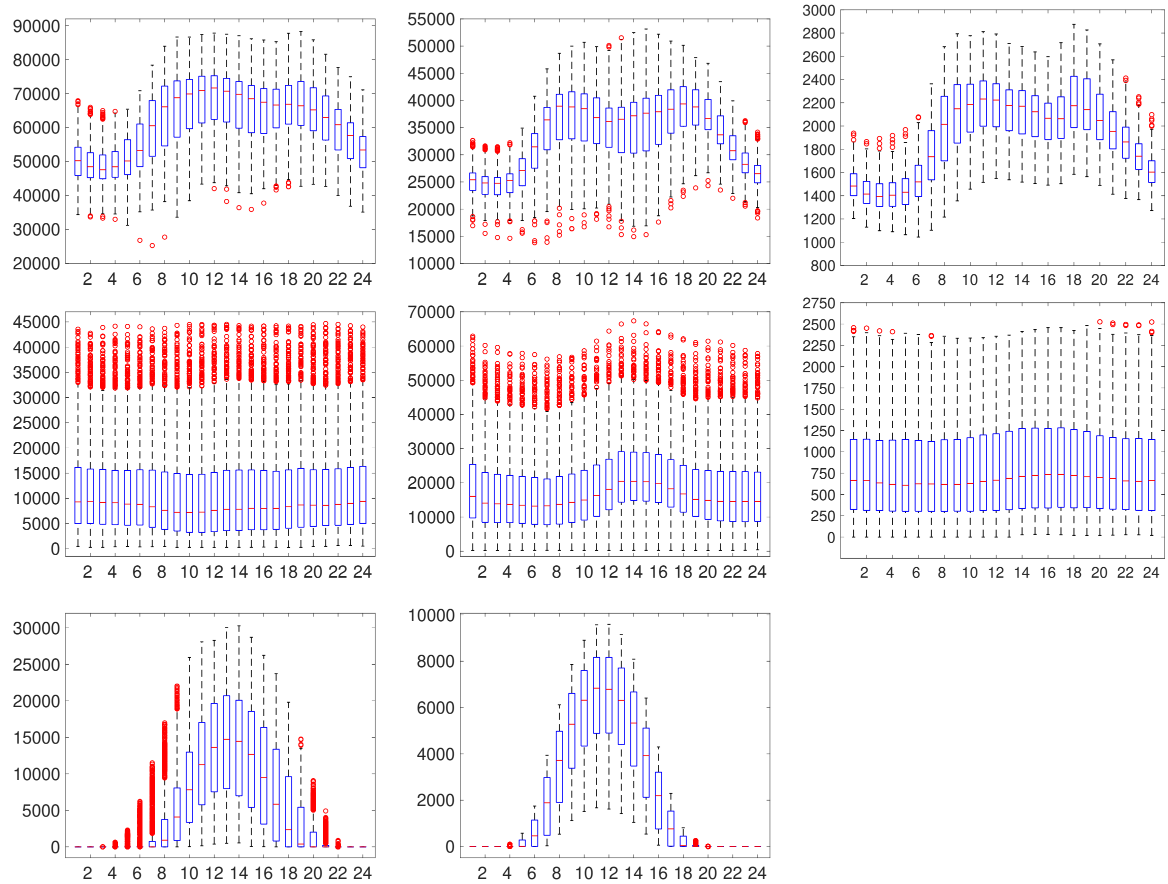

Figure 2: Sample Distributions for Forecasted Demand (top row), Forecasted Wind (middle row) and Forecasted Solar PV (bottom row) for Germany (on the left), Italy (on the center) and Denmark (on the right).

models for forecasting electricity prices. Therefore, we inspect the dynamics of hourly electricity prices in a univariate context to understand if simple and less parameterized models are more suitable than multivariate ones. To support the multivariate formulation, we run 24 univariate models for electricity prices $y_{h,t}$ for hour h observed on day t and prices observed one, two and 7 days before[10] ($y_{h,t-l}$ with $l = 1, 2, 7$). Moreover, we consider monthly dummies ($d_{k,t}$, where

[10] We follow Knittel and Roberts (2005), Weron and Misiorek (2008) and Raviv et al. (2015), who show that these specifications provide accurate forecasts because they capture seasonal patterns in electricity prices. Then, we restrict

$d_{1t}, \ldots, d_{12t}$ represent the twelve months of the year) and dummies for Saturdays, Sundays and holidays for each country (these are indicated with d_{13t}, d_{14t} and d_{15t}); thus $K = 15$. Regarding the exogenous variables, we include forecasted demand ($x_{h,t}$), wind ($w_{h,t}$) and solar ($z_{h,t}$) generation for hour h and fundamentals prices for coal (c_{t-1}), natural gas (g_{t-1}) and CO_2 (m_{t-1}). Finally, we add also the first lag of all the other remaining hours $y_{j,t-1}$, formally

$$y_{h,t} = \sum_{k=1}^{K} \psi_k d_{k,t} + \sum_{l \in \{1,2,7\}} \phi_l y_{h,t-l} + \alpha_1 x_{h,t} + \alpha_2 w_{h,t} + \alpha_3 z_{h,t} + \beta_1 c_{t-1} + \beta_2 g_{t-1} + \beta_3 m_{t-1} + \sum_{j \neq h} \gamma_j y_{j,t-1} + \varepsilon_{h,t}. \tag{1}$$

Then, from the resulting 24 residual series of each model, $\hat{\varepsilon}_{h,t}$, the variance-covariance matrix has been computed. Uncorrelated residuals make the multivariate VAR specification not necessary, however, we find evidence of large correlations, across studied markets as emphasized by the heat maps in Figure 3. Therefore, VAR models seem more appropriate to estimate this covariance structure, and it could result in improved efficiency of mean equation estimates and consequently in point and density forecast accuracy.

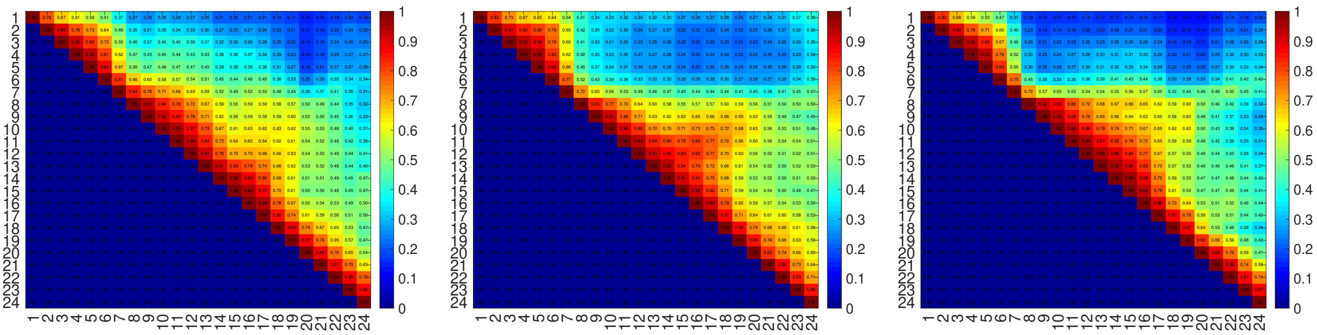

Figure 3: Heat maps of the Upper Triangular Correlation Matrix for Residuals of the Univariate Model as formulated in Eq. 1 for Germany (left), Italy (center) and Denmark (right).

3 Methodology

This section introduces first the multivariate models specification. Then, details on the estimation procedure are presented, and finally the metrics used to assess both point and density forecasting performances are described.

lags to $t-1$, $t-2$ and $t-7$, corresponding to the previous day, two days before, and one week before the delivery time. They recall first similar conditions that may have characterized the market over the same hours and similar days (like congestions and blackouts) and, second, the demand level during the days of the week. Besides, this formulation reduces the risk of overparameterization.

3.1 Model Specifications: VARs with Stochastic Volatility

In this section, we outline the class of models we wish to compare. Firstly, we consider the most general model with time-varying volatility (Chan and Eisenstat, 2018; Cross et al., 2020) and then other models are specified as restricted versions of the general one. In particular, we highlight the differences between a Gaussian stochastic volatility model and a fat-tail stochastic volatility model, thus with the Student-t error term. Therefore, our multivariate specifications allow us to compare whether features such as intermittent and unpredictable supply and variable demand (and consequent negative prices) only increase volatility or actually change the (tail) distribution, as suggested by Gianfreda and Bunn (2018) in their analysis of univariate time series.

Let $\mathbf{y}_t = (y_{1t}, \ldots, y_{Ht})'$ denote the $(H \times 1)$ vector of day-ahead hourly electricity prices, with $H = 24$. Consider the following vector autoregressive (VAR) model with stochastic-volatility (SV):

$$A_0\mathbf{y}_t = B_1\mathbf{y}_{t-1} + \ldots + B_p\mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t, \qquad \boldsymbol{\varepsilon}_t \sim \mathcal{N}(0, \Sigma_t), \tag{2}$$

where $B_1, \ldots, B_p$ are the $(H \times H)$ VAR matrix of coefficients; A_0 is a $(H \times H)$ lower triangular matrix with ones on the diagonal and Σ_t is a time-varying diagonal matrix of the form $\Sigma_t = \text{diag}(\exp(h_{1t}), \ldots, \exp(h_{Ht}))$. Following Chan and Eisenstat (2018), we reformulate the model as follows:

$$\mathbf{y}_t = \tilde{X}_t\boldsymbol{\beta} + W_t\boldsymbol{\gamma} + \boldsymbol{\varepsilon}_t, \qquad \boldsymbol{\varepsilon}_t \sim \mathcal{N}(0, \Sigma_t), \tag{3}$$

where $\tilde{X}_t = I_H \otimes (\mathbf{y}'_{t-1}, \ldots, \mathbf{y}'_{t-p})$ and W_t contains the appropriate elements of $\mathbf{y}_t$. Regarding the coefficients in $B_1, \ldots, B_p$ and A_0, we can split them in two different groups. The first group consists of $\boldsymbol{\beta}$, which is a $(k_\beta \times 1)$ vector containing the coefficients associated with the lagged observations, the dummy variables and the exogenous variables. The second group contains a $(k_\gamma \times 1)$ vector, $\boldsymbol{\gamma}$, of coefficients that characterizes the contemporaneous relations between the variables and it consists of the free elements of A_0 stacked by rows.

In particular, $k_\gamma = H(H-1)/2$, while the size of the vector of coefficients varies along the model specifications: if we include the lagged observations, then $k_\beta = H^2p$, whereas if we add a vector of dummies denoted by $\mathbf{d}_t = (d_{1t}, \ldots, d_{Kt})'$, (with $K = 15$ since it includes $d_{1t}, \ldots, d_{12t}$ for the twelve months, d_{13t}, d_{14t} and d_{15t} for Saturdays, Sundays and holidays), then $k_\beta = (Hp+K)H$. The model in Eq. (3) can be written in a stacked form:

$$\mathbf{y}_t = X_t\boldsymbol{\theta} + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}, \Sigma_t),$$

where $X_t = (\tilde{X}_t, W_t)$ and $\boldsymbol{\theta} = (\boldsymbol{\beta}', \boldsymbol{\gamma}')$ is of dimension $k_\theta = k_\beta + k_\gamma$. If we assume $p = 3$, then there are $k_\theta = 2364$ parameters to be estimated.

In order to complete the model specification of the VAR(p) with stochastic volatility, we need to include the time-varying volatility in the model. Thus, we include the log-volatilities $\boldsymbol{h}_t = (h_{1t}, \ldots, h_{Ht})$ for $t = 1, \ldots, T$. Following Cogley and Sargent (2005), we assume that the latent log-volatilities $\boldsymbol{h}_t$ evolve according to a random walk process

$$\boldsymbol{h}_t = \boldsymbol{h}_{t-1} + \boldsymbol{u}_t, \quad \boldsymbol{u}_t \sim \mathcal{N}(\mathbf{0}, \Omega),$$

where $\boldsymbol{u}_t$ is a vector of i.i.d. residuals, $\Omega = \mathrm{diag}(\sigma^2_{h_1}, \ldots, \sigma^2_{h_H})$ and $\boldsymbol{h}_0$ is treated as a parameter to be estimated.

As stated in Cross et al. (2020), the prior specifications for the state variance, Ω, and for the initial state, $\boldsymbol{h}_0$, follow an independent prior distribution such as

$$\boldsymbol{h}_0 \sim \mathcal{N}(\boldsymbol{a}_{\boldsymbol{h}}, V_h), \quad \sigma^2_{h_i} \sim \mathcal{IG}(\nu_{h_i}, S_{h_i}) \quad \text{for } i = 1, \ldots, H,$$

where $\mathcal{IG}(\cdot, \cdot)$ denotes an inverse Gamma distribution. Regarding the hyperparameters, we set $\boldsymbol{a}_{\boldsymbol{h}} = \mathbf{0}$, $V_h = 10 \times I_H$; $\nu_{h_i} = 10$ and $S_{h_i} = 0.1^2(\nu_{h_i} - 1)$.

Regarding the prior distribution of the vectorized matrix of coefficients $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, we assume an independent Normal prior[11] specification of the form:

$$\boldsymbol{\beta} \sim \mathcal{N}(\underline{\boldsymbol{\mu}}_\beta, \underline{V}_\beta); \quad \boldsymbol{\gamma} \sim \mathcal{N}(\underline{\boldsymbol{\mu}}_\gamma, \underline{V}_\gamma),$$

where $\underline{\boldsymbol{\mu}}_\beta, \underline{\boldsymbol{\mu}}_\gamma$ are the prior means set equal to $(\mathbf{0}, \mathbf{0})$ and $\underline{V}_\beta, \underline{V}_\gamma$ are the prior covariance matrixes set equal to $(10 \times I_{k_\beta}, I_{k_\gamma})$. We can specify Eq. 3 in a different form by using a different representation of the $\tilde{X}_t$ matrix. In particular, we include some exogenous variables in the VAR, which leads to a VARX specification with stochastic volatility. The exogenous variables included in the analysis refer to both the demand and supply curves. As far as the former one is concerned, we include the forecasted hourly demand $\boldsymbol{x}_t = (x_{1t}, \ldots, x_{Ht})'$ which contains variability around the expected levels of demand. As far as the supply is concerned, we consider fossil fuel prices, which however do not change over the 24 hours and are determined over the previous day, and the variability induced by the forecasted values for RES. Then, to summarize, we have included $\boldsymbol{w}_t = (w_{1t}, \ldots, w_{Ht})'$ for

[11] Please note that we have also used a shrinkage prior of the Dirichlet-Laplace form for all the components of the $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ in both VAR and VARX specifications. However, the results were similar to those with the Normal prior, hence they have been omitted.

forecasted wind generation; $\boldsymbol{z}_t = (z_{1t}, \ldots, z_{Ht})'$ for forecasted solar power generation; and m_{t-1}, g_{t-1} and c_{t-1} for CO_2, gas and coal prices determined on the previous day, respectively.

From Eq 3, we redefine the matrix $\tilde{X}_t = I_H \otimes (\mathbf{y}'_{t-1}, \ldots, \mathbf{y}'_{t-p}, \mathbf{d}'_t, \mathbf{x}'_t, \mathbf{w}'_t, \mathbf{z}'_t, m_{t-1}, g_{t-1}, c_{t-1})$, thus transforming Eq. 3 into a VARX(p)-SV, where $k_\beta = (Hp + K + 3H + 3)H$ is the dimension of the vector of coefficients $\boldsymbol{\beta}$.

In this paper, we consider a different specification of the time-varying covariance matrix that augments the (random walk) stochastic volatility specification to include fat tails. Thus, we introduce a VAR(p) with stochastic volatility with a Student-t error term and we introduce the degrees of freedom of the Student-t distribution. Differently from a strand of the literature (see Chan, 2020a,b), we include the degrees of freedom that change with the variables, such as electricity demand and solar power generation which show higher values during daytime. Thus, $\nu_j > 0$ depends on $j = 1, \ldots, H$ to account for the varying dynamics across the hours.

The model specification follows Eq. (3) but the variance matrix Σ_t has a novel component. The VAR(p)-tSV is given by

$$\begin{aligned} \boldsymbol{y}_t &= \tilde{X}_t \boldsymbol{\beta} + W_t \boldsymbol{\gamma} + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(\mathbf{0}, \text{diag}(\exp(h_{1t})/\lambda_{1t}, \ldots, \exp(h_{Ht})/\lambda_{Ht})), \\ \boldsymbol{h}_t &= \boldsymbol{h}_{t-1} + \boldsymbol{u}_t, \quad \boldsymbol{u}_t \sim \mathcal{N}(\mathbf{0}, \text{diag}(\sigma^2_{h_1}, \ldots, \sigma^2_{h_H})), \end{aligned} \tag{4}$$

where $\lambda_{jt} \sim \mathcal{IG}(\nu_j/2, \nu_j/2)$ for every t and $j = 1, \ldots, H$. We consider two different options on the degree of freedom of the fat tail component. In the first one, the degrees of freedom ν_j are parameters to be estimated for each variable, thus we assume a Uniform prior for ν_j, i.e. $\nu_j \sim \mathcal{U}(0, \underline{\nu})$, where we fix $\underline{\nu}$ equal to 50; while in the second case, we fix the degrees of freedom $\nu_j = 5$ to ensure fat tails[12].

As for the Gaussian stochastic volatility model, we consider a specification with exogenous variables also for the case with fat tails. Thus, we have a different specification of $\tilde{X}_t = I_H \otimes (\mathbf{y}'_{t-1}, \ldots, \mathbf{y}'_{t-p}, \mathbf{d}'_t, \mathbf{x}'_t, \mathbf{w}'_t, \mathbf{z}'_t, m_{t-1}, g_{t-1}, c_{t-1})$ and we can define a VARX(p)-tSV model to be used in the in-sample estimation and in the forecasting exercise.

As anticipated, we do not encounter any ordering problems in the lower triangular matrix A_0 as instead in Carriero et al. (2019); Cross et al. (2020). In fact, in our analysis, we have a strong time dependence due to the 24-hour consecutive specification, and the variable ordering in the

[12]The choice of ν_j equal to 5 is due to the posterior mean over the 24 hours of the degrees of freedom estimated in the VARX-tSV model.

VAR-SV and VAR-tSV model will not affect the in-sample analysis and the relative forecasting performance of the models. We remember that the 24 hourly prices are set jointly the day before delivery, this implies that the set of 24 hourly prices has to be submitted jointly to the system operator. This is a technical market requirement for the functioning of the system.

For comparison, we consider also a constant variance model, thus a VAR(p) and a VARX(p), where the error term is distributed with a constant covariance matrix $\Sigma_t = \Sigma$. As far as the prior assumption is concerned, we use the usual Bayesian prior specification for the covariance matrix and we consider an inverse Wishart distribution, $\mathcal{IW}(\underline{\nu}_0, \underline{\Psi}_0)$, where $\underline{\nu}_0$ are the prior degrees of freedom and $\underline{\Psi}_0$ is the prior scale matrix. A summary of the models used is presented in Table 1.

Models	Description
VAR	with Gaussian constant volatility
VARX	with Gaussian constant volatility
VAR-SV	with Gaussian stochastic volatility
VARX-SV	with Gaussian stochastic volatility
VAR-tSV	with stochastic volatility and t innovations (estimated ν_j)
VARX-tSV	with stochastic volatility and t innovations (estimated ν_j)
VAR-tSV$_{\boldsymbol{\nu}}$	with stochastic volatility and t innovations (fixed $\boldsymbol{\nu} = \nu_j = 5$)
VARX-tSV$_{\boldsymbol{\nu}}$	with stochastic volatility and t innovations (fixed $\boldsymbol{\nu} = \nu_j = 5$)

Table 1: List of Competing Bayesian VAR Models.

3.2 Estimation Procedure

Recalling that hourly prices with a reduced 7-lag structure are considered, with abuse of notation we use p = 3 to denote the maximum number of lags in the remainder of the article. Thus, to further inspect the effect of fat tails, we have investigated the differences between estimating or fixing the degrees of freedom at the same value for all the 24 hours.

Regarding the estimation method, we refer the reader to the usual Bayesian literature. In particular, the joint posterior distributions are not available in closed form, thus we need to apply a Markov Chain Monte Carlo (MCMC) method based on a Gibbs sampling and draw each parameter from its posterior full conditional distribution. For the VAR model with stochastic volatility (VAR-SV), the Gibbs sampler could be summarized through the following steps:

1. sample $\boldsymbol{\beta}$ from its posterior full conditional distribution;
2. sample $\boldsymbol{\gamma}$ from it posterior full conditional distribution given the other parameters;
3. sample the path $h_{i,0}, \ldots, h_{i,T}$ for each $i = 1, \ldots H$ by sampling each $h_{j,t}$ from its posterior full conditional distribution;

4. sample $\sigma^2_{h_i}$ for each $i = 1 \dots H$ from its posterior full conditional distribution.

Whereas for the VAR model with Student-t distribution, we need to add the following two steps for the degrees of freedom:

5. sample $\lambda_{i,t}$ for each $i = 1, \dots, H$ and $t = 1, \dots, T$ from its posterior full conditional distribution;

6. sample ν_i for each $i = 1, \dots, H$ using an independent Metropolis-Hastings step as in Chan and Hsiao (2014). The proposal distribution is a Gaussian $\mathcal{N}(\hat{\nu}, K^{-1})$, where $\hat{\nu}$ is the mode of the log-posterior density and K is the negative Hessian evaluated at the mode.

In order to check which is the best model across the variety of models estimated, we adopt two criteria for the comparison of Bayesian models, which will be used to compare the estimation performances of the models. In particular, we apply the deviance information criterion (DIC) introduced in the seminal paper by Spiegelhalter et al. (2002), which is defined as the sum of the posterior mean deviance and the effective number of parameters. Hence, it indicates a tradeoff between the model fit and its complexity; see Chan and Eisenstat (2018) for a more detailed discussion. In addition, we implement the Bayes Factor and the relative ratio with respect to a benchmark model, that is the VAR model. The Bayes factor selects the model that it is more likely to have occurred given the observed data.

As far as computational times are concerned, they generally range from 60 seconds for the simplest model to less than 8 minutes for more complex ones; further details are provided in the supplementary material.

3.3 Assessment of the Forecasting Performance

In order to assess the goodness of our forecasts, we adopt point and density metrics which are described in this section. As a point forecast measure, we apply the root-mean-square errors for each of the hourly prices ($RMSE_h$), as well as the RMSEs for the daily average ($RMSE_{\text{Avg}}$) and for the average restricted only to the central peak hours from 8 a.m. to 8 p.m. ($RMSE^P_{\text{Avg}}$). Specifically, they are computed as follows

$$\text{RMSE}_h = \sqrt{\frac{1}{T-R}\sum_{t=R}^{T-1}\left(\hat{y}_{h,t+1|t} - y_{h,t+1}\right)^2},$$

where T is the number of observations, R is the length of the rolling window and $\hat{y}_{h,t+1|t}$ are the individual hourly price forecasts, and

$$\mathrm{RMSE}_{\mathrm{Avg}} = \frac{1}{24}\sum_{h=1}^{24}\mathrm{RMSE}_h \quad \text{and finally} \quad \mathrm{RMSE}^P_{\mathrm{Avg}} = \frac{1}{13}\sum_{h=8}^{20}\mathrm{RMSE}_h. \tag{5}$$

On the other hand, we evaluate density forecasts by using the average continuous ranked probability score (CRPS) and the quantile CRPS (Gneiting and Raftery, 2007; Gneiting and Ranjan, 2011).[13] These measures have advantages over the log score, in particular, they reward better the values from the predictive density that are close to - but not equal to - the outcome, and they are less sensitive to outliers. The CRPS is defined as the lower the number the better the score, and it is given by

$$\mathrm{CRPS}_{h,t}(y_{h,t+1}) = \int_{-\infty}^{\infty}\left(F(z) - \mathbb{I}\{y_{h,t+1} \leq z\}\right)^2 dz = E_f|Y_{h,t+1} - y_{h,t+1}| - 0.5E_f|Y_{h,t+1} - Y'_{h,t+1}|, \tag{6}$$

where F denotes the cumulative distribution function associated with the predictive density f, $\mathbb{I}\{y_{h,t+1} \leq z\}$ denotes an indicator function taking the value 1 if $y_{h,t+1} \leq z$ and 0 otherwise, and $Y_{h,t+1}$ and $Y'_{h,t+1}$ are independent random draws from the posterior predictive density.

Regarding the quantile CRPS, the quantile-weighted versions of the continuous ranked probability score is defined as:

$$S_{h,t}(y_{h,t+1}) = \int_0^1 \mathrm{QS}_\alpha\left(F^{-1}(\alpha), y_{h,t+1}\right)\omega(\alpha)d\alpha, \tag{7}$$

where $\mathrm{QS}_\alpha\left(F^{-1}(\alpha), y_{h,t+1}\right)$ is the quantile score defined as

$$\mathrm{QS}_\alpha\left(F^{-1}(\alpha), y_{h,t+1}\right) = 2\left(\mathbb{I}\{y_{h,t+1} \leq F^{-1}(\alpha)\} - \alpha\right)\left(F^{-1}(\alpha) - y_{h,t+1}\right)$$

with $F^{-1}(\alpha)$ the quantile forecast and $\alpha \in (0,1)$. When $\omega(\alpha) = 1$, we have an uniform weight, thus an unweighted continuous ranked probability score.

In order to put the emphasis on the centre or on the tails of the distribution, the nonnegative weight function can assume different specifications on the unit interval $\omega(\alpha)$. In particular, $\omega(\alpha) = \alpha(1-\alpha)$ focusses on the centre, whereas $\omega(\alpha) = (2\alpha - 1)^2$ puts more emphasis on

[13] Our analysis focuses mainly on relative model performance. We have also investigated absolute performance via probability integral transforms (PITS). Models with stochastic volatility and exogenous variables are well calibrated. Results are available upon request.

the tails. Moreover, the emphasis on the right and left tails can be denoted by $\omega(\alpha) = \alpha^2$ and $\omega(\alpha) = (1-\alpha)^2$, respectively. And in what follows, we indicate the average center quantile CRPS with *CQ-CRPS*, the average right tail quantile CRPS with *RQ-CRPS*, and the average left tail quantile CRPS with *LQ-CRPS*, respectively. As for the RMSE, we can compute the averages for the CRPS, CQ-CRPS, RQ-CRPS and LQ-CRPS over the 24 hours and over the peak hours on day $t+1$. We report the RMSEs, average CRPS, and average quantile CRPSs for the baseline VAR model with constant volatility, whereas for the other VAR models we report the ratios between the computed metric of the current model over the computed metric of the baseline VAR model. Then, entries of less than 1 indicate that the given current model yields forecasts more accurate than those provided by the baseline.

Moreover, to provide a rough gauge of whether the differences in forecast accuracy are significant, we apply the Diebold and Mariano (1995) t-tests for equality of the average loss to compare predictions produced by alternative models with those produced by the benchmark model for a given horizon h[14]. The differences in accuracy that are statistically different from zero are denoted with one, two, or three asterisks, corresponding to significance levels of 10%, 5%, and 1%, respectively. The underlying p-values are based on t-statistics computed with a serial correlation-robust variance, using the pre-whitened quadratic spectral estimator of Andrews and Monahan (1992). Our use of the Diebold-Mariano test, with forecasts from models that are in many cases nested, is a deliberate choice, as in Clark and Ravazzolo (2015). And, as noted by Clark and West (2007) and Clark and McCracken (2012), this test is conservative and might result in under-rejection of the null hypothesis of equal predictability. We report p-values based on one-sided tests, taking the benchmark model (e.g. VAR model with constant volatility) as the null and the other current models (e.g. VARX model with and without time-varying volatility) as the alternative.

Finally, the Model Confidence Set procedure proposed by Hansen et al. (2011) across models for a fixed horizon has been employed to jointly compare their predictive power without disentangling between constant or time-varying volatility. The R package MCS detailed in Bernardi and Catania (2016) has been used, and the differences have been tested separately for each hour and model, repeating the full process for all three countries. Results are presented and discussed in the following section.

[14] When testing density forecasts, we use equal weights without adopting a weighting scheme, as in Amisano and Giacomini (2007).

4 Empirical Results

In this section, we illustrate the performance of the proposed Bayesian VAR models with different volatility structures. We first present results based on comparisons across estimated models assessing their performances in terms of DIC and Bayes Factor: these results based on a full sample estimation show that the introduction of stochastic volatility fits the data substantially better. Finally, the forecasting performances of the Bayesian VARs models are compared using different point and density measures.

4.1 Model Comparisons

In this section, we present results for the Bayes Factor and DIC for all considered models and studied countries. For the Bayes Factor, results are provided with respect to the baseline model (that is the VAR model), where entries less than 1 indicate that the current model yields more accurate estimations than those provided by the baseline model. Table 2 shows that including time-varying volatility provides better results with respect to the benchmark model. This finding is consistent for all countries, hence supporting the evidence of fat tails. As expected, the superiority of the VARX-tSV model is confirmed for all countries, although with a difference: while estimating the degrees of freedom seems the best solution for Germany and Denmark, fixing them is more appropriate for Italy. This may be due to the zero lower bound in action in the Italian market, where prices cannot become negative even if there is a combination of low demand and high RES generation.

Bayes Factor	VARX	VARX-SV	VARX-tSV	VARX-tSV$_{\nu}$
Germany	0.758	0.849	**0.722**	0.736
Italy	0.925	0.872	0.847	**0.836**
Denmark	0.846	0.949	**0.819**	0.862

Table 2: Bayes Factor under each model specification on the full sample size ($T = 1454$) for different countries. Values less than 1 indicate that the given current model yields more accurate estimations than those provided from the VAR model.

Moving to the second criterium of models comparison, we consider the DIC as computed in Chan and Eisenstat (2018). In particular, we provide the posterior mean and numerical standard errors in parentheses of the DIC computed over 20 different chains. Each chain consists of 20.000 posterior draws after a burn-in period of 5.000 iterations, while the integrated likelihood is evaluated every 20 post burn-in draws. Results from the in-sample analysis are reported in

Table 3. In details, and recalling that higher DIC values correspond to better fit, we observe that this metric confirms that models with stochastic volatility are the best fitted models, whereas models not including exogenous variables (VAR) or with constant volatility (VARX) are the worst ones, and this is consistent for all countries. As the Bayes Factor, the VARX-tSV with fixed and estimated degrees of freedom are found to be the best fitting models for Germany. For Italy and Denmark, the best models are the VARX-SV immediately followed by the VARX-tSV. Overall, this confirms the relevance of including time-varying volatility in the modelling strategy.

	VAR	VARX	VARX-SV	VARX-tSV	VARX-tSV$_\nu$
Germany					
DIC	192693	183949	165651	164472	**163362**
	(148.44)	(53.12)	(798.39)	(4.01)	(4.30)
Italy					
DIC	176047	174590	**162306**	163178	162947
	(90.31)	(92.80)	(2.92)	(3.43)	(2.86)
Denmark					
DIC	185466	182512	**155119**	159940	159726
	(137.02)	(94.56)	(280.65)	(325.63)	(57.32)

Table 3: The DIC estimates (numerical standard errors in parentheses) under each model specification on the full sample size ($T = 1454$) for different countries.

4.2 Dynamics of the Estimated Volatility

In this section, we present the dynamics of estimated volatility over the full sample, from January 2016 to December 2019.

The comparison of estimated stochastic volatility models with the one estimated with constant volatility for Germany is presented in Figure 4, where it is possible to observe volatility representations changing across the 24 hours (in simple words, they are time-varying estimates over the hours as described in Section 3). More specifically, we compare the VARX model with constant volatility computed over 1 year centered rolling window[15] (on the left), with the the posterior means of the VARX model with stochastic volatility (middle) and those of the VARX with fat tail stochastic volatility (on the right). Note that we left the same scale to emphasize the difference across models.

It is interesting to observe that there are substantial changes across the hours of the days for the latter models and, in particular, looking at the fat tail representation of the error, the

[15] We follow Clark and Ravazzolo (2015) who proposed this approach and provide all technical details for implementations on macroeconomic data.

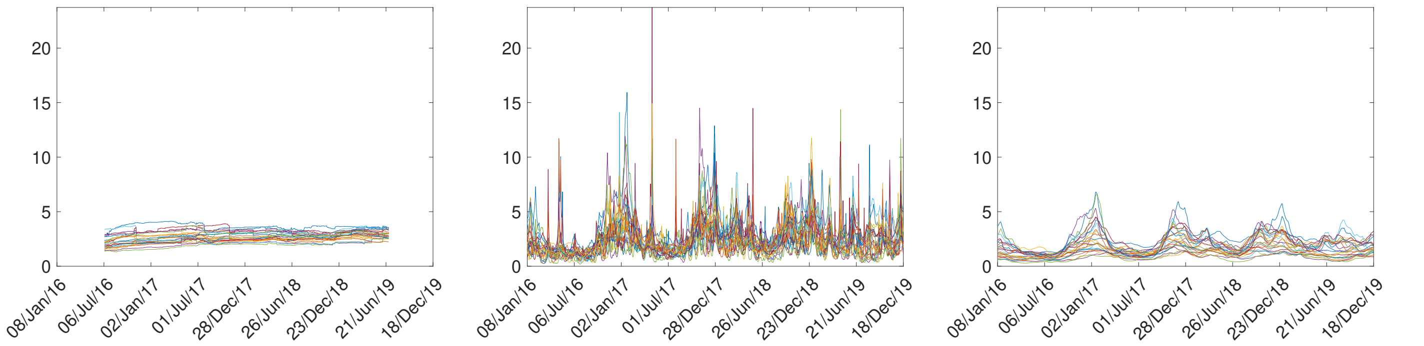

Figure 4: Model comparisons over all the 24 hours between the VARX computed over 1 year centred rolling window (on the left), the VARX-SV (middle) and the VARX-tSV (on the right) estimated in Germany. Posterior Means of the Stochastic Volatility Models are expressed in standard deviations $\exp(h_t/2)$.

volatility shows high values during winters and lower values during summers recalling the calendar seasonality for all the four studied years. The VARX model with stochastic volatility seems to be more prone to spikes at different hours and different periods of the year. In general, the assumption of constant volatility could therefore imprecisely estimate the time-varying pattern of volatility. Similar results are observed for Italy and Denmark, and are available in the supplementary material[16], which also contains the comparisons of the three models across all 24 hours.

The volatility estimates from the VARX model with constant volatility are on average higher than those estimated by the time-varying volatility models. And, to investigate this fact we provide the number of times that the posterior means of the VARX model are higher than those of the VARX model with stochastic volatility (first row) and (higher than) the VARX with fat tail stochastic volatility (on the second row), respectively; see Figure 5.[17] Looking at these results, the posterior means of the VARX model are on average higher than those of VARX with fat tailed stochastic volatility for all hours and for most of the hours excluding early hours at night when compared to the VARX with stochastic volatility. The average of these statistics over the 24 hours for German price is 64% in the VARX vs VARX-SV comparison, and 89% in the VARX vs VARX-tSV comparison. For Italy, results are qualitatively similar, with a 24-hour average of 60% for the VARX vs VARX-SV comparison, and 87% for the VARX vs VARX-tSV comparison. And for Denmark, results show an average of 73% in the VARX versus VARX-SV comparison, and of 98% for the VARX vs VARX-tSV comparison.

[16] Furthermore, we repeated our investigations on subperiods of years (2016-2017 and 2018-2019) and on individual years to provide an additional robustness check. These results are omitted for lack of space, but are available on request.

[17] When computing these statistics, we discard the initial 6-month data and the final 6-month data in order to compute rolling volatility from the VARX model.

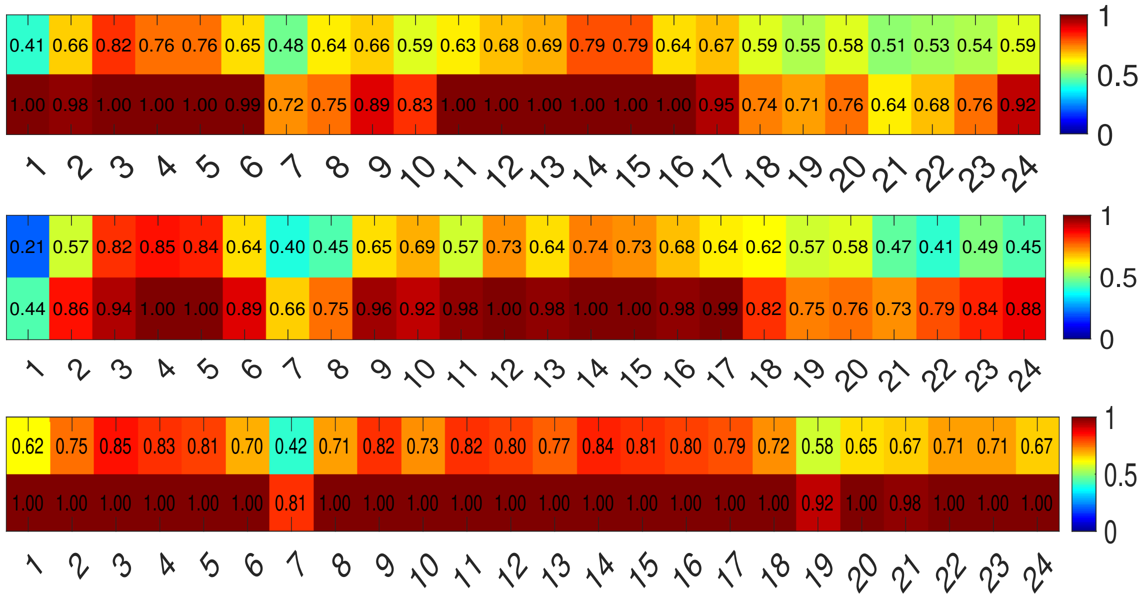

Figure 5: Number of times over the adjusted full sample that the posterior means of the VARX model are higher than those of the VARX model with stochastic volatility (first row) and than those of the VARX with fat tail stochastic volatility (second row) for Germany (top panel), Italy (middle panel) and Denmark (bottom panel).

In addition, Figure 6 shows the posterior means of the three specifications - VARX,[18] VARX-SV and VARX-tSV at selected hours (10, 14 and 18, that is when RES show their higher production in connection with high levels of demand); whereas results for the other hours are provided in the supplementary material. Model comparisons across the same hour emphasize indeed that the posterior means of the volatility estimated with a constant volatility structure (VARX) are higher than those with a time-varying structure (that is VARX-SV and VARX-tSV), and also that introducing fat tails in the error term (VARX-tSV in grey lines) systematically reduces the volatility over the considered sample and reductions are substantial especially during winters with this phenomenon being more clearly visible at hours 10 and 18. Surprisingly and more interestingly, the highest reductions are found at 14 when instead the economics of the energy systems suggests the major uncertainty due to the combination of forecasts errors for demand with those for wind and, especially, solar generation.

Thus, to further inspect the effect of fat tails, we have investigated the differences between estimating or fixing the degrees of freedom at the same value over all the 24 hours. For the same three hours indicated before, Figure 7 shows the values of posterior means when $\boldsymbol{\nu}$ is estimated (black lines) and when instead $\boldsymbol{\nu}$ is fixed (to 3 and 5). While black and red lines for estimated

[18] It has been computed over 1 year centered rolling window, but we have also tested the usage of a 2 years centered rolling window with similar results; thus, they have been omitted but are available upon request.

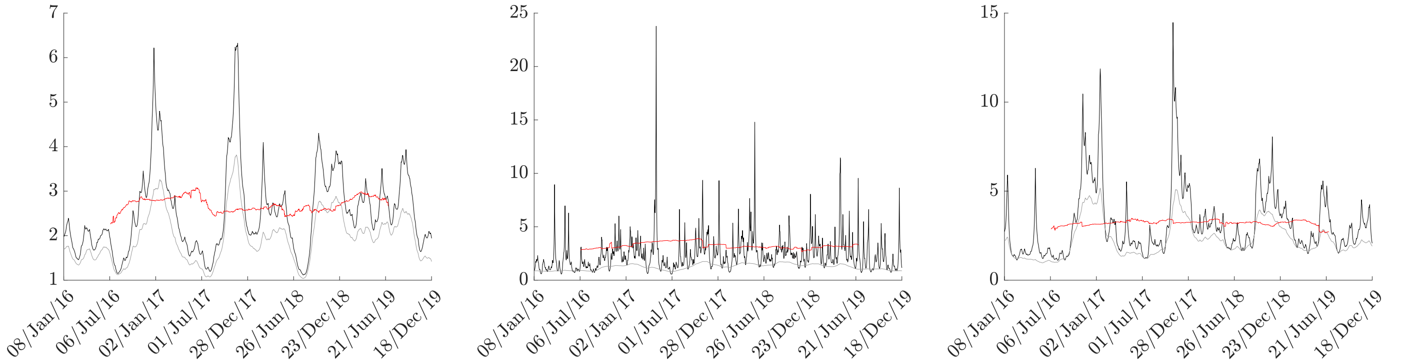

Figure 6: Posterior Means of the Stochastic Volatility Models, in standard deviations $\exp(h_t/2)$, observed in Germany. Model comparisons across hours 10 (left), 14 (middle) and 18 (right) for VARX (red line), VARX-SV (black line) and VARX-tSV (grey line).

and fixed (to 5) $\boldsymbol{\nu}$ are almost perfectly overlapping at hours 10 and 18 (as well as at hours 7-10 and 19-23 but less so for $\boldsymbol{\nu}$ fixed to 3), the three dynamics decouple substantially at hour 14 (and generally across hours 11-17) when high volatility is observed for the degrees of freedom fixed to 5 and low values occur for estimated degrees of freedom. Then, this suggests how important is accounting for time-varying changes estimated accounting for the characteristics of the 'current' sample, especially for hours with greater variability. These results confirm what was already suggested by Gianfreda and Bunn (2018) concerning the time-varying shape and tail dynamics exhibited by hourly electricity prices, even if they were considering each hour individually. Similar comments apply to Italy, whereas results for Denmark are a bit different given the decoupled dynamics observed across all 24 hours (in the supplementary material). This may be due to the different Danish generation mix.

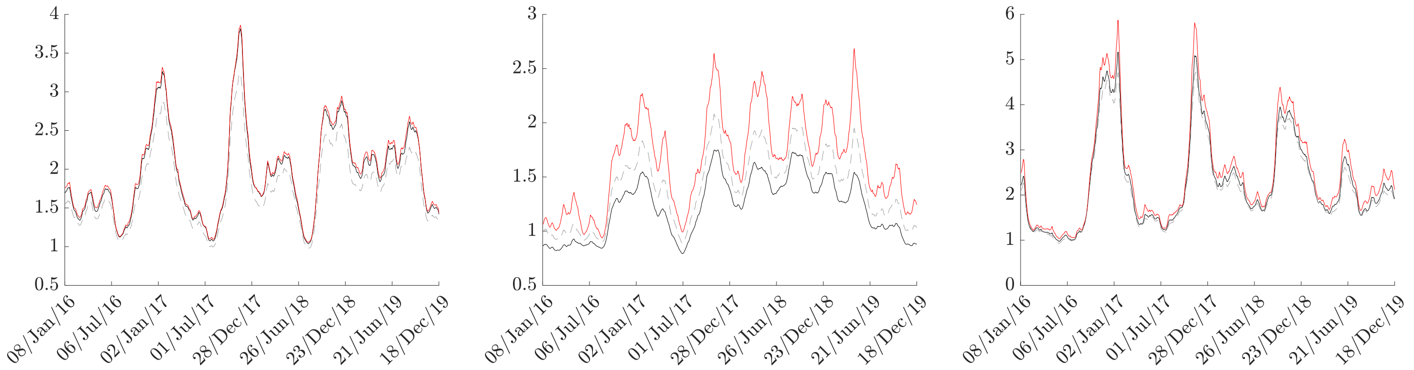

Figure 7: Posterior Means of the Stochastic Volatility Models VARX-tSV - in standard deviations $\exp(h_t/2)$ - with estimated $\boldsymbol{\nu}$ (black line), fixed $\boldsymbol{\nu} = 5$ (red line) and also $\boldsymbol{\nu} = 3$ (grey dashed) at hours 10 (left), 14 (centre) and 18 (right) in Germany.

When looking at the distribution of degrees of freedom $\boldsymbol{\nu}$, Figure 8 shows the posterior means and distribution of ν over the 24 hours. The credibility intervals provide the uncertainty about estimated values. Results shows that uncertainty is high around hours 7-10 and 18-23, instead it is low over hours 12-17, following the intra-daily demand profiles. In addition, in Italy uncertainty is more pronounced over the first three hours, probably because of the strategic behaviour of hydro units bidding on both the day-ahead and real time markets (Gianfreda et al., 2018). Reading these results together with those reflecting the differences in using estimated or fixed degrees of freedom, we can state that estimating or fixing the degrees of freedom does not imply dramatic changes at certain hours; however, attention must be paid to central hours since the low uncertainty is coupled with substantial differences between estimated or ex-ante fixed values, with the latter ones producing substantial overestimation. Hence, we recommend a dynamic estimation better adapting to all hours.

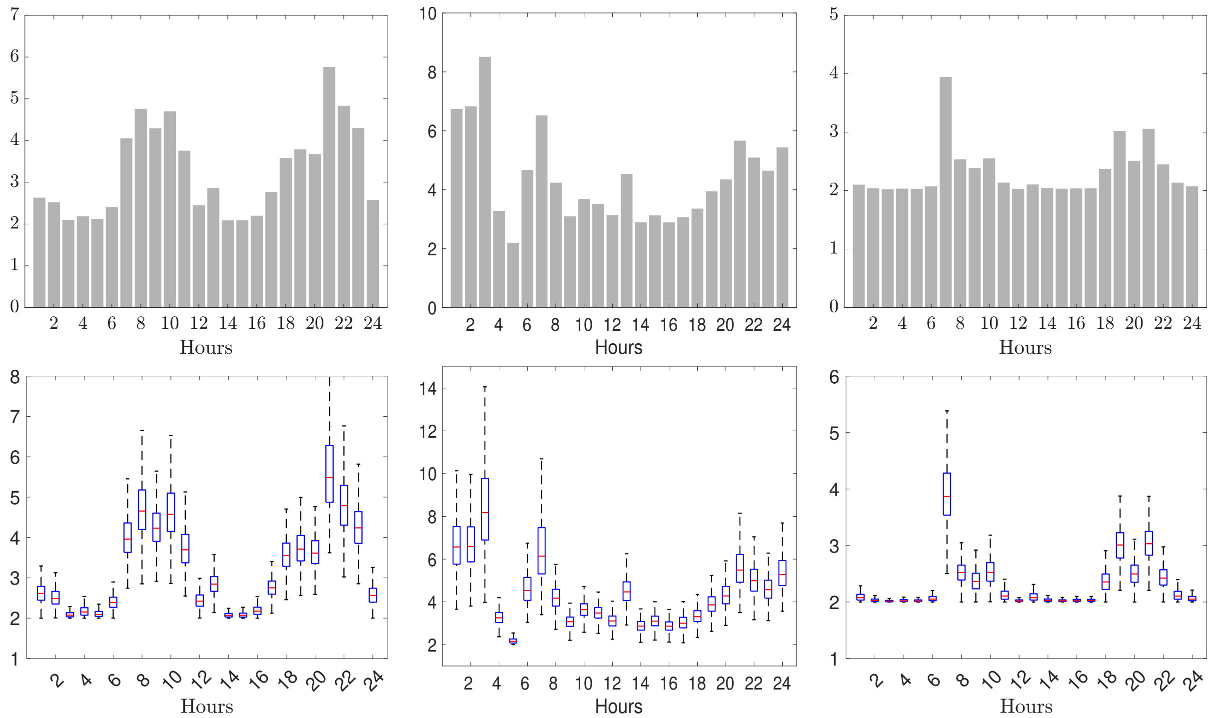

Figure 8: Posterior Means (top row) and Posterior Distribution (bottom row) of the Degrees of Freedom ν across the 24 hours for the VARX-tSV model in Germany (on the left column), Italy (in the middle) and Denmark (on the right). The red horizontal line represents the posterior mean, whereas the blue box indicates the 75% credibility interval.

4.3 Forecasting Results

Our results are based on a one-step ahead rolling forecasting process with a window size of two years. Please note that the initial estimation sample goes from 1 January 2016 to 31 December 2017 and then the forecasting evaluation period starts on 1 January 2018 and ends on 31 December 2019; for a total of 731 in-sample forecasts.

Results refer to the performances of our different multivariate models[19] from the simplest one (with lags and dummy variables, the benchmark VAR) to more complex ones including constant or different time-varying volatility specifications as well as fundamental drivers (fuel prices and forecasts for demand and renewable energy sources). Results for Germany, Italy and Denmark are reported in Tables 4, 5 and 6, for a selection of hours; whereas, results for all 24 hours and comprehensive model comparisons are provided in the supplementary material.

As expected, the forecasting performance decreases across the peak hours (that is between hour 8 and hour 20) and this is consistent with the high uncertainty affecting demand and supply levels during daytime. Hence, the benchmark VAR models show the highest RMSEs at hours 14-16 in Germany (around 13€/MWh); at peak hours 9-10, and 15-20 in Italy (around 8€/MWh) and at 8-11 in Denmark (around 11€/MWh). These differences can be explained by the intra-daily dynamics of forecasted demand and RES generation, which differ substantially across countries because of the diverse geographical conditions affecting, for instance, solar radiation and wind speed. Further details and comparisons of the intra-daily dynamics of demand and RES generation are presented in Gianfreda et al. (2020) and Gianfreda et al. (2016).

The most interesting aspect addressed here is the expected forecasting improvement resulting from accounting for volatility together with fundamental drivers (demand, RES, and fuels).

First of all, our results show that the Bayesian multivariate models with stochastic volatility (VARX-SV) exhibit substantial improvements with respect to the the VARX and the benchmark VARs with constant volatility. Including fundamental drivers in the modeling of price variability increases forecast accuracy and even larger gains are obtained by extending the models to account for time-varying volatility. When looking at the RMSE and CRPS metrics, the Model Confidence Set shows that the VAR and VARX models are never included in the model set when predicting German electricity prices; whereas, they may be retained at early-morning hours for Italian prices and from late-afternoon to night hours for Danish prices.

In Germany, further improvements can be observed when we include the Student-t stochastic volatility. Again, this occurs for both point and density metrics. Moreover, we observe no difference between estimating the degrees of freedom or fixing them equal to an ex-ante selected value (for instance 5) across hours. This confirms our previous findings and proves that both ways produce

[19]We also performed the forecasting exercise against the model proposed by Kostrzewski and Kostrzewska (2019), but the results from our model beats the SVDEJX. These results have been omitted for lack of space, but they are available on request.

similar forecasting results. Therefore, practitioners can make an indifferent decision regarding their modeling strategy, but we suggest estimating the degrees of freedom to address different degrees of fat tails over the 24 hours. On the contrary, in Italy there is no empirical evidence of differences between Gaussian and fat tail stochastic volatility models across the considered forecasting measures. On one hand, this is in line with the Italian lower bound for market prices. On the other hand, this may be due to the different levels of RES penetration observed in this country, especially for wind. Moreover, these improvements are less evident in Denmark, where accounting for fat tails does not provide further gains with respect to Gaussian time-varying error. For this country, it seems that estimating the degrees of freedom provides low improvements with respect to fixing them, due to the fact that Denmark has little variation in the estimation of $\boldsymbol{\nu}$ across hours with respect to the other countries.

Looking specifically at the metrics, the average reductions in loss function are similar for both metrics of about 20% in Germany and Denmark, and 5% in Italy. Forecasting gains in terms of the RMSE increase from hour 8 to the end of the day, as shown by full tables in the supplementary material, and more clearly by the last column of Tables 4, 5 and 6. In general, the RMSEs decrease from the VARX to the VARX-tSV across all hours, more strongly in Germany and Denmark than in Italy. Similarly for the CRPSs, for which we observe reductions of almost 10% in Germany and Denmark, and 5% in Italy. In Germany, the use of fat tails (Student t distribution) leads to improvements of 20% in both metrics, while we do not observe similar reductions in Italy and Denmark.

Considering the density forecasting and looking at the centre of distributions, results for the CQ-CRPS show small differences between the benchmark model and the model with time-varying volatility, of around 0.6 for Germany, 0.8 for Italy and 0.79 for Denmark.

When looking at the tails, results in terms of the average LQ-CRPS and RQ-CRPS confirm the expected lower gains in Italy and Denmark, of about 0.8 and 0.79 with respect to the benchmark levels for both left and right tails. As expected, we observe substantial forecasting improvements in Germany, however and surprisingly, of the equal magnitude for both left and right tails, of about 0.6 with respect to the benchmark levels.

As argued by Gianfreda and Bunn (2018) in their analysis on German individual hours selected according to intra-daily profiles, wind and solar generation reduce the skewness of hourly electricity prices, with this phenomenon being more evident at hours 12-13 because solar is at its maximum level. Moreover, they add that both increase the kurtosis of electricity prices at peak hour 19. On the contrary, hour 3 shows higher volatility because more negative price spikes are observed compared to hours 12-13 and 19.

Therefore, considering Germany and its off-peak1 prices (that is in the early morning from hour 1 to 7), lower quantiles are of most practical interest for the occurrence of down spikes. Then, looking for instance at hour 3 we observe that the LQ-CRPS shows limited improvements when considering constant or time-varying with Student-t volatility. This may be due to lower values observed during off-peak1 hours than those observed in peak (8-20) and off-peak2 (21-24) hours: there is indeed a jump in this metric from 0.413 at hour 1 to 1.158 at hour 15, which, however, still confirms the convenience in implementing more complex models with regressors and time-varying volatility with estimated or fixed degrees of freedom. Instead, at hours 12-13, both the high and low quantiles are of interest and here we observe substantial and significant improvements in both the LQ-CRPS and RQ-CRPS from the VARX-tSV when indifferently estimating or fixing ν. Similar comments apply at hours 19-21 when the high quantiles are the most interesting ones for the risk of high prices because RES decreases but demand is still at high levels. These comments, however, do not apply to the Italian prices since they are affected by RES with a lower magnitude compared to Germany and, more importantly, they are not allowed to become negative. However, it is still interesting to observe that during peak hours the inclusion of time-varying volatility and regressors improves both point and density metrics. See Tables 4, 5 and 6 for a sample of hours in all three markets, whereas results for all hours are reported in the extended Tables in the supplementary material.

5 Conclusions

Modeling day-ahead electricity prices has become extremely important for understanding the energy system and providing empirical support to policymakers in a context where uncertainty is progressively increasing as a consequence of changing weather conditions due to climate change. In this regard, appropriate models may also produce useful forecasts, which help market operators plan their generation schedules while accounting for (and then adapting to) the imperfect predictability of both demand and RES generation. This is also extremely relevant for the transmission system operators who must guarantee the continuous balance between demand and supply. In this framework, we address the less explored issue of multivariate models in which

Table 4: Forecasting Metrics for Germany. Note that real values are used for the Bayesian VAR model, and ratios for all the other models

Hour	1	4	7	10	13	16	19	22	Avg	Avg_{8-20}
RMSE										
VAR	5.017	6.842	9.519	11.432	11.946	13.409	11.556	10.097	8.162	10.827
VARX	1.099	0.933	0.831**	0.773***	0.760***	0.753***	0.754***	0.676***	0.730	0.710
VARX-SV	1.048	0.885***	0.760***	0.704***	0.700***	0.707***	0.698***	0.679***	0.668	0.652
VARX-tSV	1.035	0.881***	0.748***	0.690***	0.682***	0.707***	0.693***	0.675***	0.664	0.645
VARX-tSV ($\boldsymbol{\nu} = 5$)	1.039	0.878***	0.747***	0.693***	0.678***	0.702***	0.695***	0.679***	0.662	0.644
CRPS										
VAR	2.548	3.438	4.747	5.822	6.123	6.696	6.015	5.271	4.141	5.553
VARX	1.103	0.953*	0.855***	0.772***	0.751***	0.742***	0.749***	0.682***	0.732	0.707
VARX-SV	1.034	0.867***	0.751***	0.681***	0.686***	0.669***	0.667***	0.646***	0.645	0.626
VARX-tSV	0.996	0.846***	0.724***	0.660***	0.642***	0.640***	0.660***	0.644***	0.627	0.605
VARX-tSV ($\boldsymbol{\nu} = 5$)	1.002	0.847***	0.724***	0.664***	0.642***	0.640***	0.658***	0.642***	0.627	0.605
CQ-CRPS										
VAR	0.248	0.335	0.463	0.570	0.601	0.654	0.590	0.518	0.405	0.544
VARX	1.104	0.961**	0.863***	0.775***	0.754***	0.743***	0.751***	0.683***	0.733	0.706
VARX-SV	1.040	0.874***	0.755***	0.685***	0.687***	0.666***	0.669***	0.647***	0.647	0.626
VARX-tSV	1.001	0.850***	0.725***	0.662***	0.642***	0.638***	0.661***	0.646***	0.628	0.605
VARX-tSV ($\boldsymbol{\nu} = 5$)	1.007	0.852***	0.726***	0.667***	0.644***	0.639***	0.661***	0.644***	0.629	0.606
RQ-CRPS										
VAR	0.388	0.503	0.732	0.907	0.930	0.997	0.937	0.782	0.629	0.852
VARX	1.163	0.997	0.864***	0.779***	0.751***	0.760***	0.757***	0.705***	0.747	0.716
VARX-SV	1.038	0.883***	0.746***	0.682***	0.687***	0.668***	0.677***	0.661***	0.642	0.624
VARX-tSV	0.986	0.839***	0.710***	0.656***	0.633***	0.633***	0.671***	0.660***	0.617	0.599
VARX-tSV ($\boldsymbol{\nu} = 5$)	0.998	0.845***	0.710***	0.659***	0.633***	0.633***	0.671***	0.658***	0.617	0.599
LQ-CRPS										
VAR	0.415	0.580	0.763	0.923	0.993	1.112	0.951	0.872	0.673	0.893
VARX	1.046	0.905***	0.837***	0.760***	0.745***	0.722***	0.737***	0.658***	0.714	0.697
VARX-SV	1.026	0.845***	0.754***	0.677***	0.684***	0.672***	0.654***	0.628***	0.634	0.611
VARX-tSV	0.999	0.848***	0.735***	0.662***	0.647***	0.647***	0.646***	0.628***	0.634	0.611
VARX-tSV ($\boldsymbol{\nu} = 5$)	0.999	0.843***	0.735***	0.664***	0.648***	0.646***	0.644***	0.626***	0.633	0.610

Notes:

[1] Please refer to Section 3 for the details on models. The 'X' indicates models with exogenous variables. All forecasts are produced with a one-step-ahead rolling window process.

[2] ***, ** and * indicate that ratios are significantly different from 1 at 1%, 5% and 10%, according to the Diebold-Mariano test.

[3] Grey cells indicate models that belong to the Superior Set of Models delivered by MCS procedure at confidence level 10%.

[4] Results with a fixed $\boldsymbol{\nu} = 3$ are similar with those obtained with $\boldsymbol{\nu} = 5$, thus they have been omitted but are available in the supplementary material.

volatility dynamics are included. We have questioned whether the inclusion of a constant or a time-varying volatility structure can better detect the movements of electricity prices in three European countries, namely Germany, Italy and Denmark. Thus, we propose high dimensional VAR models with different stochastic volatility representations in which fundamental drivers are included as exogenous variables, these are forecasted demand, forecasted renewable energy sources, and fuels. In particular, we assume that the time-varying volatility changes across hours and drives

Table 5: Forecasting Metrics for Italy. Note that real values are used for the Bayesian VAR model, and ratios for all the other models

Hour	1	4	7	10	13	16	19	22	Avg	Avg_{8-20}
RMSE										
VAR	4.421	5.270	5.868	8.545	6.905	8.613	8.551	6.266	5.095	6.983
VARX	1.031	1.021	0.974	0.894***	0.911***	0.876***	0.982	0.964	0.945	0.890
VARX-SV	1.016	1.019	0.954	0.844***	0.886***	0.841***	0.896***	0.901***	0.898	0.841
VARX-tSV	1.002	1.019	0.952	0.856***	0.891***	0.850***	0.903***	0.896***	0.902	0.850
VARX-tSV ($\boldsymbol{\nu} = 5$)	1.001	1.021	0.950	0.851***	0.886***	0.847***	0.898***	0.893***	0.898	0.845
CRPS										
VAR	2.441	2.928	3.094	4.553	3.665	4.572	4.593	3.409	2.754	3.760
VARX	1.037	1.019	0.991	0.901***	0.897***	0.878***	0.979	0.963*	0.953	0.892
VARX-SV	1.018	0.999	0.957**	0.837***	0.867***	0.827***	0.875***	0.874***	0.892	0.829
VARX-tSV	0.999	0.999	0.957**	0.843***	0.863***	0.831***	0.878***	0.870***	0.894	0.832
VARX-tSV ($\boldsymbol{\nu} = 5$)	0.999	0.998	0.955**	0.837***	0.858***	0.828***	0.872***	0.868***	0.892	0.828
CQ-CRPS										
VAR	0.241	0.290	0.304	0.448	0.360	0.450	0.453	0.337	0.272	0.370
VARX	1.038	1.017	0.992	0.909***	0.897***	0.882***	0.981	0.964**	0.955	0.896
VARX-SV	1.021	1.005	0.965***	0.844***	0.870***	0.830***	0.875***	0.878***	0.898	0.834
VARX-tSV	1.004	1.004	0.965***	0.848***	0.865***	0.832***	0.877***	0.873***	0.900	0.836
VARX-tSV ($\boldsymbol{\nu} = 5$)	1.003	1.003	0.962***	0.844***	0.860***	0.829***	0.873***	0.871***	0.897	0.832
RQ-CRPS										
VAR	0.392	0.444	0.482	0.723	0.585	0.730	0.717	0.546	0.426	0.592
VARX	1.015	1.036	1.010	0.914***	0.917***	0.889***	0.989	0.973	0.976	0.912
VARX-SV	1.006	1.013	0.976	0.855***	0.898***	0.847***	0.909***	0.886***	0.928	0.859
VARX-tSV	0.987	1.010	0.975	0.862***	0.896***	0.855***	0.908***	0.885***	0.931	0.862
VARX-tSV ($\boldsymbol{\nu} = 5$)	0.986	1.010	0.974	0.856***	0.890***	0.852***	0.906***	0.884***	0.929	0.860
LQ-CRPS										
VAR	0.371	0.470	0.488	0.701	0.563	0.704	0.721	0.519	0.434	0.586
VARX	1.061	1.006	0.970	0.881***	0.874***	0.858***	0.966*	0.951**	0.926	0.868
VARX-SV	1.026	0.982	0.931***	0.813***	0.832***	0.802***	0.838***	0.859***	0.849	0.793
VARX-tSV	1.007	0.985	0.930***	0.820***	0.828***	0.804***	0.847***	0.852***	0.852	0.797
VARX-tSV ($\boldsymbol{\nu} = 5$)	1.008	0.984	0.927***	0.813***	0.823***	0.800***	0.837***	0.847***	0.848	0.791

Notes: Please see the notes to Table 4.

the dependence in a time-ordered structure, avoiding ordering problems as stated in the literature (Carriero et al., 2019; Cross et al., 2020).

Using only a lagged representation of the data or adding different exogenous variables, we find empirical evidence supporting VAR models with stochastic volatility against the conventional VAR. Indeed, most of the gains appear to come from allowing for a time-varying stochastic volatility rather than a constant volatility structure. In particular, the assumption of fat tails in the error term improves the detection of time-varying volatility, during some hours of the day. Furthermore, in a recursive forecasting exercise, we find that models with exogenous variables show improvements in both point and density forecasts. In addition, the combined inclusion of

Table 6: Forecasting Metrics for Denmark. Note that real values are used for the Bayesian VAR model, and ratios for all the other models

Hour	1	4	7	10	13	16	19	22	Avg	Avg_{8-20}
RMSE										
VAR	4.970	7.070	8.320	11.201	9.648	9.881	10.101	8.767	6.918	9.142
VARX	1.020	0.921	0.867***	0.846***	0.848***	0.809***	0.783***	0.803***	0.769	0.768
VARX-SV	0.980**	0.909***	0.840***	0.826***	0.835***	0.821***	0.793***	0.815***	0.785	0.773
VARX-tSV	0.984***	0.920***	0.841***	0.838***	0.842***	0.825***	0.811***	0.820***	0.794	0.786
VARX-tSV ($\boldsymbol{\nu} = 5$)	0.985**	0.917***	0.839***	0.831***	0.836***	0.820***	0.795***	0.817***	0.788	0.777
CRPS										
VAR	2.438	3.566	4.228	5.712	4.909	5.061	5.425	4.509	3.601	4.764
VARX	1.048	0.961*	0.889***	0.836***	0.865***	0.832***	0.764***	0.811***	0.783	0.772
VARX-SV	0.971*	0.894***	0.833***	0.795***	0.850***	0.835***	0.771***	0.810***	0.778	0.765
VARX-tSV	0.939***	0.880***	0.829***	0.797***	0.834***	0.815***	0.776***	0.815***	0.777	0.767
VARX-tSV ($\boldsymbol{\nu} = 5$)	0.947***	0.883***	0.827***	0.791***	0.826***	0.809***	0.759***	0.807***	0.769	0.756
CQ-CRPS										
VAR	0.237	0.345	0.411	0.558	0.478	0.492	0.534	0.440	0.351	0.465
VARX	1.055	0.970***	0.891***	0.839***	0.873***	0.842***	0.762***	0.815***	0.784	0.775
VARX-SV	0.964	0.891***	0.838***	0.796***	0.851***	0.838***	0.769***	0.816***	0.781	0.768
VARX-tSV	0.937***	0.878***	0.833***	0.797***	0.835***	0.820***	0.775***	0.818***	0.780	0.769
VARX-tSV ($\boldsymbol{\nu} = 5$)	0.945***	0.883***	0.832***	0.792***	0.830***	0.817***	0.759***	0.813***	0.774	0.760
RQ-CRPS										
VAR	0.351	0.503	0.654	0.945	0.784	0.782	0.889	0.698	0.559	0.761
VARX	1.082	1.012	0.906***	0.835***	0.865***	0.839***	0.767***	0.823***	0.795	0.779
VARX-SV	1.006	0.937***	0.817***	0.788***	0.842***	0.831***	0.771***	0.797***	0.772	0.758
VARX-tSV	0.949***	0.889***	0.818***	0.798***	0.830***	0.813***	0.775***	0.813***	0.772	0.765
VARX-tSV ($\boldsymbol{\nu} = 5$)	0.965**	0.896***	0.813***	0.792***	0.817***	0.801***	0.765***	0.800***	0.762	0.753
LQ-CRPS										
VAR	0.419	0.625	0.681	0.852	0.764	0.813	0.811	0.722	0.576	0.737
VARX	1.011	0.908***	0.868***	0.834***	0.854***	0.816***	0.761***	0.795***	0.767	0.761
VARX-SV	0.948***	0.862***	0.840***	0.802***	0.857***	0.839***	0.769***	0.814***	0.780	0.771
VARX-tSV	0.934***	0.878***	0.837***	0.798***	0.836***	0.814***	0.780***	0.813***	0.779	0.767
VARX-tSV ($\boldsymbol{\nu} = 5$)	0.933***	0.873***	0.833***	0.789***	0.831***	0.810***	0.753***	0.808***	0.770	0.754

Notes: Please see the notes to Table 4.

time-varying volatility plus exogenous variables leads to further improvements in both point and density metrics.

In the future, it would be interesting to extend these models by including a global shrinkage prior to both exogenous and lagged variables or, alternatively, to use different priors in all the parameters. Besides electricity markets, these models are expected to become of extreme interest in future applications. For instance, we have already assisted to the formation of negative interest rates in response to financial crises and, more recently, the question of time instability has been posed to account for sudden large drops in electricity demand and in the WTI prices, as observed in April 2020 during the COVID-19 pandemic. Then, it has become clear that electricity prices, as well as oil prices and interest rates, can become negative and attract the attention of all economists to understand their economic implications.

References

Afanasyev, D. O. and Fedorova, E. A. (2019). On the impact of outlier filtering on the electricity price forecasting accuracy. *Applied energy*, 236:196–210.

Amisano, G. and Giacomini, R. (2007). Comparing density forecasts via weighted likelihood ratio tests. *Journal of Business & Economic Statistics*, 25(2):177–190.

Andrews, D. and Monahan, J. (1992). An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. *Econometrica*, 60(4):953–966.

Bauwens, L., Hafner, C. M., and Pierret, D. (2013). Multivariate volatility modeling of electricity futures. *Journal of Applied Econometrics*, 28(5):743–761.

Bernardi, M. and Catania, L. (2016). Portfolio Optimisation Under Flexible Dynamic Dependence Modelling. *ArXiv e-prints*.

Bordignon, S., Bunn, D. W., Lisi, F., and Nan, F. (2013). Combining day-ahead forecasts for British electricity prices. *Energy Economics*, 35:88 – 103.

Carriero, A., Clark, T. E., and Marcellino, M. (2016). Common Drifting Volatility in Large Bayesian VARs. *Journal of Business & Economic Statistics*, 34(3):375–390.

Carriero, A., Clark, T. E., and Marcellino, M. (2019). Large Bayesian vector autoregressions with stochastic volatility and non-conjugate priors. *Journal of Econometrics*, 212(1):137–154.

Chan, J. C. and Hsiao, C. Y. (2014). *Estimation of Stochastic Volatility Models with Heavy Tails and Serial Dependence*, chapter 6, pages 155–176. John Wiley & Sons, Ltd.

Chan, J. C. C. (2020a). Large Bayesian VARs: A Flexible Kronecker Error Covariance Structure. *Journal of Business & Economic Statistics*, 38(1):68–79.

Chan, J. C. C. (2020b). *Large Bayesian Vector Autoregressions*, pages 95–125. Springer International Publishing, Cham.

Chan, J. C. C. and Eisenstat, E. (2018). Bayesian model comparison for time-varying parameter VARs with stochastic volatility. *Journal of Applied Econometrics*, 33(4):509–532.

Chan, J. C. C. and Grant, A. L. (2016). Modeling energy price dynamics: Garch versus stochastic volatility. *Energy Economics*, 54:182–189.

Chiu, C.-W. J., Mumtaz, H., and Pintér, G. (2017). Forecasting with var models: Fat tails and stochastic volatility. *International Journal of Forecasting*, 33(4):1124–1143.

Ciarreta, A. and Zarraga, A. (2016). Modeling realized volatility on the spanish intra-day electricity market. *Energy Economics*, 58:152 – 163.

Clark, T. E. and McCracken, M. W. (2012). Reality checks and comparisons of nested predictive models. *Journal of Business & Economic Statistics*, 30(1):53–66.

Clark, T. E. and Ravazzolo, F. (2015). Macroeconomic forecasting performance under alternative specifications of time-varying volatility. *Journal of Applied Econometrics*, 30(4):551–575.

Clark, T. E. and West, K. D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, 138(1):291–311.

Cogley, T. and Sargent, T. J. (2005). Drifts and volatilities: monetary policies and outcomes in the post WWII US. *Review of Economic Dynamics*, 8(2):262–302.

Conejo, Contreras, Espínola, and Plazas (2005). Forecasting electricity prices for a day-ahead poolbased electric energy market. *International Journal of Forecasting*, 21(3):435–462.

Cross, J. L., Hou, C., and Poon, A. (2020). Macroeconomic forecasting with large Bayesian VARs: Global-local priors and the illusion of sparsity. *International Journal of Forecasting*, 36(3):899–915.

Damm, A., Köberl, J., Prettenthaler, F., Rogler, N., and Töglhofer, C. (2017). Impacts of +2°C global warming on electricity demand in Europe. *Climate Services*, 7:12 – 30. IMPACT2C - Quantifying projected impacts under 2 °C warming.

Diebold, F. and Mariano, R. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13(3):253–263.

Efimova, O. and Serletis, A. (2014). Energy markets volatility modelling using garch. *Energy Economics*, 43:264 – 273.

Erdogdu, E. (2016). Asymmetric volatility in european day-ahead power markets: A comparative microeconomic analysis. *Energy Economics*, 56:398 – 409.

Escribano, A., Peña, I. J., and Villaplana, P. (2011). Modelling Electricity Prices: International Evidence. *Oxford Bulletin of Economics and Statistics*, 73(5):622–650.

Frömmel, M., Han, X., and Kratochvil, S. (2014). Modeling the daily electricity price volatility with realized measures. *Energy Economics*, 44:492–502.

Garcia, R. C., Contreras, J., van Akkeren, M., and Garcia, J. (2005). A GARCH forecasting model to predict day-ahead electricity prices. *IEEE Transactions on Power Systems*, 20(2):867–874.

Gianfreda, A. and Bunn, D. (2018). A stochastic latent moment model for electricity price formation. *Operations Research*, 66(5):1189–1203.

Gianfreda, A. and Grossi, L. (2012). Forecasting Italian electricity zonal prices with exogenous variables. *Energy Economics*, 34(6):2228–2239.

Gianfreda, A., Parisio, L., and Pelagatti, M. (2016). Revisiting long-run relations in power markets with high RES penetration. *Energy Policy*, 94:432 – 445.

Gianfreda, A., Parisio, L., and Pelagatti, M. (2018). A review of balancing costs in Italy before and after RES introduction. *Renewable and Sustainable Energy Reviews*, 91:549 – 563.

Gianfreda, A., Parisio, L., and Pelagatti, M. (2019). The RES-induced switching effect across fossil fuels: An analysis of day-ahead and balancing prices. *The Energy Journal*, 40.

Gianfreda, A., Ravazzolo, F., and Rossini, L. (2020). Comparing the forecasting performances of linear models for electricity prices with high RES penetration. *International Journal of Forecasting*, 36(3):974–986.

Gneiting, T. and Raftery, A. (2007). Strictly proper scoring rules, prediction and estimation. *Journal of American Statistical Association*, 102(477):359–378.

Gneiting, T. and Ranjan, R. (2011). Comparing density forecasts using threshold- and quantile-weighted proper scoring rules. *Journal of Business and Economic Statistics*, 29(3):411–422.

Hansen, P. R., Lunde, A., and Nason, J. M. (2011). The Model Confidence Set. *Econometrica*, 79:453–497.

Henley, A. and Peirson, J. (1997). Non-linearities in electricity demand and temperature: Parametric versus non-parametric methods. *Oxford Bulletin of Economics and Statistics*, 59(1):149–162.

Huber, F. and Feldkircher, M. (2019). Adaptive Shrinkage in Bayesian Vector Autoregressive Models. *Journal of Business & Economic Statistics*, 37(1):27–39.

Huber, F., Koop, G., and Onorante, L. (2020). Inducing Sparsity and Shrinkage in Time-Varying Parameter Models. *Journal of Business & Economic Statistics*, pages 1–15.

Huurman, C., Ravazzolo, F., and Zhou, C. (2012). The power of weather. *Computational Statistics and Data Analysis*, 56(11):3793–3807.

Jeon, J. and Taylor, J. W. (2016). Short-term density forecasting of wave energy using arma–garch models and kernel density estimation. *International Journal of Forecasting*, 32(3):991 – 1004.

Jónsson, T., Pinson, P., Madsen, H., and Nielsen, H. A. (2014). Predictive densities for day-ahead electricity prices using time-adaptive quantile regression. *Energies*, 7(9):5523–5547.

Karakatsani, N. and Bunn, D. W. (2010). Fundamental and behavioural drivers of electricity price volatility. *Studies in Nonlinear Dynamics & Econometrics*, 14(4):1–42.

Ketterer, J. C. (2014). The impact of wind power generation on the electricity price in germany. *Energy Economics*, 44:270 – 280.

Knittel, C. and Roberts, M. (2005). An empirical examination of restructured electricity prices. *Energy Economics*, 27(5):791–817.

Koop, G., Korobilis, D., and Pettenuzzo, D. (2019). Bayesian compressed vector autoregressions. *Journal of Econometrics*, 210(1):135–154.

Koopman, S. J., Ooms, M., and Carnero, M. A. (2007). Periodic seasonal reg-arfima-garch models for daily electricity spot prices. *Journal of the American Statistical Association*, 102(477):16–27.

Kostrzewski, M. and Kostrzewska, J. (2019). Probabilistic electricity price forecasting with bayesian stochastic volatility models. *Energy Economics*, 80:610 – 620.

Laporta, A. G., Merlo, L., and Petrella, L. (2018). Selection of value at risk models for energy commodities. *Energy Economics*, 74:628–643.

Maciejowska, K. and Nowotarski, J. (2016). A hybrid model for gefcom2014 probabilistic electricity price forecasting. *International Journal of Forecasting*, 32(3):1051 – 1056.

Mumtaz, H. and Zanetti, F. (2013). The impact of the volatility of monetary policy shocks. *Journal of Money, Credit and Banking*, 45(4):535–558.

Mumtaz, H. and Zanetti, F. (2015). Labor market dynamics: A time-varying analysis. *Oxford Bulletin of Economics and Statistics*, 77(3):319–338.

Panagiotelis, A. and Smith, M. (2008). Bayesian density forecasting of intraday electricity prices using multivariate skew t distributions. *International Journal of Forecasting*, 24(4):710 – 727.

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

Raviv, E., Bouwman, K. E., and van Dijk, D. (2015). Forecasting day-ahead electricity prices: Utilizing hourly prices. *Energy Economics*, 50:227 – 239.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(4):583–639.

Stock and Watson (2002). Forecasting using principal components from a large number of predictors. *Journal of American Statistical Association*, 97(460):1167–1179.

Weron, R. (2006). *Modeling and forecasting electricity loads and prices: a statistical approach*. Wiley.

Weron, R. and Misiorek, A. (2008). Forecasting spot electricity prices: A comparison of parametric and semiparametric time series models. *International Journal of Forecasting*, 24(4):744–763.

Ziel, F. (2016). Forecasting electricity spot prices using lasso: On capturing the autoregressive intraday structure. *IEEE Transactions on Power Systems*, 31(6):4977–4987.