International Journal of Forecasting 36 (2020) 974–986

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Comparing the forecasting performances of linear models for electricity prices with high RES penetration

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

[a] "Marco Biagi", Department of Economics, University of Modena and Reggio Emilia, viale Berengario, 51 41121 Modena, Italy
[b] Free University of Bozen-Bolzano, Italy
[c] Vrije Universiteit Amsterdam, Netherlands
[d] Energy Markets Group, London Business School, UK
[e] Centre for Applied Macroeconomics and Commodity Prices, BI Norwegian Business School, Norway

ARTICLE INFO

Keywords:
Point and density forecasting
Electricity markets
Hourly prices
Renewable energy sources
Demand
Fossil fuels

ABSTRACT

We compare alternative univariate versus multivariate models and frequentist versus Bayesian autoregressive and vector autoregressive specifications for hourly day-ahead electricity prices, both with and without renewable energy sources. The accuracy of point and density forecasts is inspected in four main European markets (Germany, Denmark, Italy, and Spain) characterized by different levels of renewable energy power generation. Our results show that the Bayesian vector autoregressive specifications with exogenous variables dominate other multivariate and univariate specifications in terms of both point forecasting and density forecasting.

1. Introduction

Despite the recent availability of high-frequency data for forecasted demand and renewable generation, the literature on forecasting electricity prices using these exogenous variables is still relatively scarce. Therefore, we aim to fill this gap by looking at linear models, in both univariate and multivariate frameworks, while comparing the frequentist approach with the Bayesian approach and evaluating both point forecasts and density forecasts.

We show that hourly prices can be predicted efficiently by taking advantage of intraday information available to market participants when fossil fuels are controlled for. We have explored linear autoregressive (AR) and vector AR (VAR) models, both with and without fundamental predicted drivers (forecasted demand, forecasted wind and solar power generation). These exogenous variables play an important role in formulating day-ahead conditional expectations, and their effects have motivated extensive research. Furthermore, in the last ten years, the amount of electricity generated from renewable energy sources (RES-E) has grown significantly because of political and financial support for these sources, which may play an essential role not only in reducing the energy dependence (on imported fossil fuels) of a country but also, and more importantly, in mitigating global warming (by reducing greenhouse gas emissions). The renewable energy sources' (RES) share of the total power capacity increased from 24% to 44% between 2000 and 2015 in Europe, reaching a total of more than 2,000 GW in 2016. The share of wind power increased from 2.4% to 15.6%, with total generation of approximately 300 TWh, covering more than 10% of EU demand. Denmark and Germany were among the leading countries for total wind power capacity per inhabitant. The global solar photovoltaic (PV) capacity totalled an estimated 106 GW in Europe at the end of 2016, which is more than 32 times the capacity observed in 2006. Germany, Italy, and Spain belong to

* Corresponding author at: "Marco Biagi", Department of Economics, University of Modena and Reggio Emilia, viale Berengario, 51 41121 Modena, Italy.

E-mail addresses: angelica.gianfreda@unimore.it (A. Gianfreda), francescoravazzolo@gmail.com (F. Ravazzolo).

https://doi.org/10.1016/j.ijforecast.2019.11.002

the group of top-ten countries worldwide for capacity and additions (REN21, 2017). These statistics support our choice of selected markets.

On the operational side, RES have added complexity to the management of the electricity system, and thus to electricity price modelling and forecasting. Consequently, a growing body of literature has investigated the effects of RES on electricity price dynamics in several markets around the world (Europe, United States, Canada, and Australia). Given the uncertainties in the forecasted levels of demand and RES-E, market operators are concerned about the forecasts of day-ahead prices.

Still, there is no empirical consensus about the superiority of multivariate models versus univariate models, and we aim to fill this gap when all fundamental drivers are considered, thus providing clear operational guidelines in forecasting hourly day-ahead electricity prices. Therefore, in this article we compare various univariate and multivariate linear models with and without RES-E forecasts and other fundamental drivers, estimated with use of frequentist and Bayesian approaches, for producing day-ahead forecasts of selected European electricity prices. Indeed, the advent of RES has raised numerous challenges for electricity markets in terms of managing, monitoring, modelling, and forecasting. Renewables (as wind and solar energy) have zero marginal production cost but are intermittent: if the wind blows and/or the sun shines, electricity prices are low; otherwise, when the sun stops shining or the wind stops blowing, traditional thermal plants running on fossil fuels must produce demanded electricity with higher generation costs. Consequently, some negative prices can arise when power from RES is sufficient to meet demand and some units must be paid to reduce production and/or increase demand.[1] Therefore, this emphasizes the importance of including RES-E and fossil fuels when one is looking for the best price forecasts. While there is consensus that including demand forecasts or RES-E forecasts (if the market penetration is not negligible) leads to more accurate forecasts, it is still an open question as to which RES-E forecast is more informative in which market, and also whether their inclusion can reduce the importance of fossil fuels. Hence, models with only a subset of exogenous variables have also been considered.

Our results show that demand and renewable energies increase the point and density accuracy of the predictive models, especially during peak hours. However, their inclusion does not reduce the importance of fossil fuels, which we suggest should be retained in the models. Moreover, we find evidence of better forecasting of the multivariate models, given that they allow for interrelationships among different hours of the day, and the Bayesian approach leads to further forecasting improvements. Finally, and for the first time since the increasing RES penetration, we show that the models with forecasted wind power only (besides forecasted demand and fuels) perform better than those with forecasted solar power only (besides forecasted demand and fuels). Furthermore, their simultaneous inclusion further improves the performance.

The rest of this article is structured as follows. Section 2 summarizes previous research on forecasting electricity prices and highlights our contributions. Section 3 contains the description of the market together with details of the data used. Section 4 presents our models, the estimation method, and the metrics used to assess our results. These are discussed in Section 5, together with the major findings. Finally, our conclusions are presented in Section 6.

2. Literature review

As emphasized in the reviews by Nowotarski and Weron (2018) and Weron (2014), there is increasing interest in electricity price forecasting. However, few studies have addressed the comparison of univariate and multivariate models within the frequentist and Bayesian approaches when considering both point forecasts and density forecasts and the forecasting ability of fundamental drivers. In performing extensive empirical comparisons, we aim to fill this gap while exploring several combinations of forecasted variables and fossil fuel prices.

Several studies have considered the univariate dimension for modelling purposes; for example, Chen and Bunn (2014), Gianfreda and Grossi (2012), Karakatsani and Bunn (2008), and Koopman, Ooms, and Carnero (2007). However, they did not include any forecasted renewable power generation. More recent articles have analysed the impact of RES on wholesale electricity price dynamics; see, for example, Gelabert, Labandeira, and Linares (2011), Jónsson, Pinson, and Madsen (2010), Ketterer (2014), Martinez-Anido, Brinkman, and Hodge (2016), Mauritzen (2013), Paraschiv, Erni, and Pietsch (2014), Pircalabu, Hvolby, Jung, and Høg (2017), Rintamäki, Siddiqui, and Salo (2017), and Woo, Horowitz, Moore, and Pacheco (2011). It is worth emphasizing that most of the authors modelled each hourly time series individually (i.e. 24 hourly time series separately), as in García-Martos, Rodríguez, and Sánchez (2007) and Misiorek, Trueck, and Weron (2006), hence ignoring the relationships among different hours of the day.

To overcome this issue, Maciejowska and Nowotarski (2016) proposed 24 separate AR models, including, among the regressors of the models, the early morning hours (up to 4 a.m.), the last prices (at hours 23 and 24) from the previous day, historical prices (at lags 1 and 7), a weekend dummy to capture seasonality, and load selected again at lags 1 and 7; however, no RES were included. In addition to AR models, Maciejowska and Weron (2015) also proposed VAR models for hourly and averaged daily prices, with 480 estimated parameters for working/weekend days, daylight hours, and a constant 7-lag order structure, which still does not involve demand and renewable power.

Therefore, following Conejo, Contreras, Espínola, and Plazas (2005), Maciejowska and Weron (2015), and Misiorek et al. (2006), we select AR models as benchmarks

[1] Negative prices are considered market signals of inflexibility: the system is not able to increase demand on one hand and reduce generation on the other hand because turning conventional power plants on and off would be inefficient and uneconomical.

because of their widespread use in the literature and their relatively good performance in predicting electricity prices. Moreover, we consider VAR representations to detect improvements in the forecasting performances. Indeed, we expect better forecasts from multivariate models than from univariate models given the larger information contained in a panel of data, as suggested by Stock and Watson (2002).

Being aware of the explosion in dimensionality, we push these models forward by including also forecasted demand and RES-E in both our univariate AR models and our multivariate VAR models. Furthermore, we consider exploring natural gas, coal, and CO_2 if their inclusion improves the forecasting ability, as shown by Maciejowska and Weron (2016). Hence, we manage a total of 161 parameters for each hour.

As far as forecasting is concerned, and has emerged from the reviews, few studies have considered density forecasting (e.g. Gianfreda & Bunn, 2018; Huurman, Ravazzolo, & Zhou, 2012; Jónsson, Pinson, Madsen, & Nielsen, 2014; Panagiotelis & Smith, 2008).

More recently, but without accounting for fundamental drivers and looking only at point forecasts, Raviv, Bouwman, and van Dijk (2015) compared the performances of models for the full panel of 24 hourly prices by studying Nord Pool from 1992 to 2010. On the basis of univariate AR and multivariate VAR models, they computed forecast combinations and empirically demonstrated that the useful predictive information contained in disaggregated hourly prices improves the forecasts of multivariate models. They showed that shrinking VAR models leads to further better forecasts, with the Bayesian VAR model outperforming the unrestricted VAR model. However, no density forecasting was performed and no RES were included in their models, as in Ziel and Weron (2018). Ziel and Weron (2018) proposed 58-parameter regression univariate and multivariate models accounting for different forms of seasonality, but no evidence of the uniform superiority of multivariate specifications was provided across all 12 markets, seasons, or hours studied. More specifically, and closer to our analysis, they concluded that, in Spain, the multivariate specification often outperforms the univariate specification in the morning hours, whereas in Germany and in the two Danish zones, the univariate specification often outperforms the multivariate specification in the late evening/night hours. However, these results depend on the specifications of their models and may produce different results if forecasted demand and RES-E are included. Therefore, this further supports our investigation, and we aim to provide even clearer evidence of linear univariate and multivariate forecasting performances by comparing frequentist and Bayesian models when more complexity is induced by uncertain and intermittent renewable generation.

3. Market structure and data description

3.1. The electricity market and its sessions

Wholesale electricity markets are platforms where electricity is traded. These are organized in sequential sessions: the day-ahead, intraday, and balancing sessions. In the day-ahead session, bids to buy and offers to sell electricity for each hour of the following day are submitted in pairs of prices and quantities by consumption units and generators on a voluntary basis (there is no obligation to act). This session opens several days in advance and closes one day before physical delivery. For this reason, these markets are often called *forward, auction, or day-ahead markets*, in which individual supply offers and demand bids are ordered giving priority of dispatch to more efficient and less polluting units with lower marginal costs (then wind and solar energy – RES in general – enter the supply curve before nuclear, coal, and gas units, which have higher marginal costs; this is the so-called merit order criterion). Hence, the price is computed under a cost-minimizing objective on an hourly basis and it is identified by the intersection of the aggregated curves of supply and demand. This day-ahead price is determined according to generators' planned schedules of production and by forecasted consumption programmes, which can be affected by sudden outages and weather conditions among many other factors.

Subsequently, the intraday sessions take place, wherein units are allowed to modify (by buying or selling) their day-ahead schedules as new information (such as better weather forecasts) becomes available. These operations are undertaken generally by units of intermittent and variable generation (but recently also some thermal units have started to participate in day-ahead and intraday sessions to explore higher profit opportunities in balancing sessions where prices are higher and the price as bid is used). The participation in the day-ahead and intraday sessions occurs on a voluntary basis; they are both managed by the system operator and a marginal pricing rule applies.

The balancing sessions represent the last sessions used by the transmission system operator to grant system security and grid stability and to match instantaneously demand and supply in the case of any unexpected imbalance. These are usually organized in an 'ex ante' planning phase (when generation resources are committed) and in a 'real-time' session (when the balancing is granted to restore frequency and quantity deviations); hence, several types of products are actually remunerated. Given that only generators with the required degree of flexibility are allowed to provide these services, these sessions are generally more concentrated than are the former ones, the participation is mandatory, and the pay-as-bid pricing mechanism is applied; for additional details, see Hirth and Ziegenhagen (2015) and Poplavskaya and de Vries (2019).

Day-ahead forecasts are particularly important for the market itself and for operators, because if the day-ahead forecasts (of quantities) are wrong, then energy must be acquired in the real-time market at a (potentially and generally) higher price, as highlighted by Gianfreda, Parisio, and Pelagatti (2018), who investigated all these market sessions and the bidding behaviour of (hydro, water pumping, and thermal conventional) units responsible for balancing in the northern zone of Italy.

Given the uncertainties in the forecasted levels of demand and, more importantly, those in the forecasted levels of RES-E (affecting the supply curve according to the levels of RES penetration), substantial variability is introduced. This also explains why one-step-ahead forecasts are gaining increasing interest. Moreover, market operators and traders are concerned about these forecasts of day-ahead prices because they are used in the balancing pricing mechanisms and can provide an indication of the magnitude of price spreads across sessions; see Bunn, Gianfreda, and Kermer (2018) and Lisi and Edoli (2018) for further details about imbalances and strategic speculations.

3.2. Data

We use hourly day-ahead prices (in levels) to estimate models for electricity traded/sold in Germany, Denmark, Italy, and Spain. These markets are particularly interesting given their high levels of RES penetration. Following Uniejewski, Nowotarski, and Weron (2016) and Ziel and Weron (2018), we refer to day-ahead prices and spot prices interchangeably to identify prices determined in a market today for delivery in a certain hour tomorrow, according to the literature on European electricity markets. Formally, they are forward prices determined one day in advance and with maturity on the following day.[2] This time difference is important in understanding the use of forecasted variables (as demand, wind, and solar energy) available to operators when they run their forecasting models to obtain a set of 24 prices to be submitted to power exchanges before the closure of the market.[3] We obtained national electricity prices directly from the corresponding power exchanges: the German hourly auction prices of the power spot market from the European Energy Exchange (EEX)[4]; the two-hourly zonal prices for Denmark from Nord Pool[5] (these were averaged to obtain a single price series for the whole country); the Italian hourly single national prices (*prezzo unico nazionale*, PUN) from the Italian system operator Gestore dei Mercati Energetici (GME)[6]; and the *precios horario del mercado spot diario* for Spain from the Operador del Mercado Ibérico, Polo Español (OMIE).[7] These hourly electricity prices (quoted in euros per megawatt hour), with daily frequency, were preprocessed for clock-time changes to exclude the 25th hour in October and to interpolate the missing 24th hour in March; hence, there are no missing observations.

As main drivers, we considered both the supply side and the demand side. As far as the supply side is concerned, we downloaded daily settlement prices for coal from Datastream and interpolated them for missing weekends and holidays (as for the Intercontinental Exchange API2 cost, insurance and freight Amsterdam, Rotterdam, and Antwerp, with ticker LMCYSPT), for carbon emissions (as for the EEX-EU CO_2 emissions E/EUA in euros, with ticker EEXEUAS), and for natural gas prices (as for the ICE UK, as it represents a pure hub benchmark and can be used for all EU markets, as suggested by Gianfreda, Parisio, & Pelagatti, 2016) all converted into euros per megawatt hour with use of the USEURSP rates from US dollars to euros (WMR&DS). In addition, we considered the forecasted renewable generation (from wind and solar PVs). We downloaded forecasted values for RES-E and demand directly from the market transmission system operators, apart for the German and Italian forecasts, which were provided by Thomson Reuters at hourly frequency. In these last two cases, the results from two weather providers (the European Centre for Medium-Range Weather Forecast and the Global Forecast System of the American weather service of the National Centers for Environmental Prediction) were inspected.[8] We decided to use only forecasts obtained with the European Centre for Medium-Range Weather Forecast operational model running at midnight, because this model is updated daily from 5.40 a.m. to 6.55 a.m., thus representing the latest information available to market operators to formulate their day-ahead bidding strategy.

While demand forecast models use weather forecasts accounting for temperature, precipitation, pressure, wind speeds, and cloud cover or radiation, forecasted wind values are obtained with use of the information on wind speeds and installed capacity. Finally, forecasted solar power production considers only PV installations, solar radiation, and installed capacity given the predominance of PV plants over solar thermal ones. It is worth recalling that the time series for solar power exhibits a block structure of null values in hours early in the mornings and late in the evenings, creating collinearity issues. Hence, we preprocessed these series by a linear transformation: drawing from a uniform distribution and adding these small numbers to the original zero values in the series. This results in (column) blocks of very small values close to but different from zero, instead of (column) blocks of zeros.

To summarize, we used daily fossil fuel prices (CO_2, gas, and coal, denoted by m, g, and c, respectively, and kept constant over the 24 h) and hourly data (with daily frequency) for electricity prices, forecasted demand (denoted by x), wind power (denoted by w), and solar PV generation (denoted by z) from 1 January 2011 to 31 December 2016 for Germany and Denmark and from 13 June 2014 to 13 June 2017 for Italy and Spain. We used the first four years as an estimation sample for Germany and Denmark, and the first two years for Italy and Spain,

---

2 However, it must be emphasized that in the United States the spot market is used to indicate the real-time market, whereas the day-ahead market is usually and more properly called the *forward market*.

3 Hence, we are not considering real-time prices determined by balancing needs to match instantaneously demand and supply. These prices are usually called *balancing prices* and are determined in other market sessions regulated by different pricing mechanisms; for further insights see Gianfreda et al. (2018), Gianfreda, Parisio, and Pelagatti (2019), Hirth and Ziegenhagen (2015).

4 Precisely, we accessed data from http://www.eex.com thanks to Europe Energy.

5 https://www.nordpoolgroup.com.

6 http://www.mercatoelettrico.org.

7 http://www.esios.ree.es.

8 Both use two types of weather models: the *operational* one, which is deterministic, with no involved randomness and high resolution; and the *ensemble* one, which is a probabilistic model, with lower resolution and variations around the initial set of weather conditions, hence providing different weather scenarios and, consequently, an idea of the weather instability. Both providers use one single run for the operational model and different runs for the ensemble at specific hours.

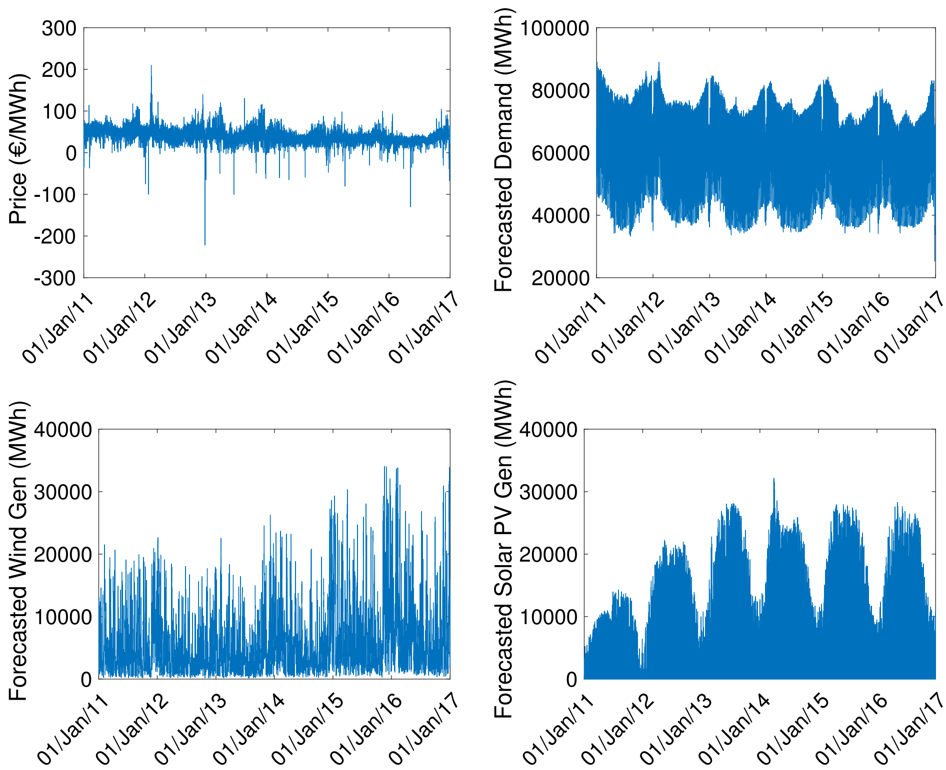

Fig. 1. Hourly series for electricity day-ahead prices (top left), forecasted demand (top right), forecasted wind power generation (bottom left), and forecasted solar PV generation (bottom right) observed in Germany from 1 January 2011 to 31 December 2016.

whereas we used the last two year/one year as the forecast evaluation period. The historical dynamics of these series observed in Germany are reported in Fig. 1. Prices show clearly the new stylized fact of 'downside' spikes together with mean reversion, whereas forecasted demand and solar generation exhibit clearer yearly seasonal patterns, with an increasing trend for solar power generation according to the new capacity additions through years. Similarly, forecasted wind power shows its dependence on weather conditions, albeit with an increasing trend,[9] corresponding again to investments in new capacity. To highlight calendar seasonality, monthly profiles for electricity prices, forecasted demand, wind and solar power generation are depicted in Fig. 2. Furthermore, to emphasize the weekly seasonality, Fig. 3 depicts the intraday dynamics across days of the weeks for demand and prices; obviously, wind and solar power generation are not presented, as they are weather dependent. Similar figures for the other countries are reported in Section S.1 in the supplementary material.

Finally, the intraday profiles for the yearly average values of forecasted demand and RES-E are presented in Fig. 4 to identify scenarios of high/low demand and/or RES-E expected to affect prices and, consequently, forecasts. We can observe that the ramp-up hours (during which the demand for electricity is expected to grow substantially) as well as the ramp-down hours (when demand is expected to decrease sharply) change across markets according to the time of the day or night and geographical locations. However, they confirm higher demand levels in the peak period (roughly between 8 a.m. and 8 p.m. for all markets). The intraday profiles for forecasted wind power generation show different dynamics: we can again identify scenarios for high wind power generation during peak hours in Denmark and Italy, whereas the opposite occurs in Germany and Spain. Obviously, the intraday profiles for forecasted solar PV generation are, instead, common for all markets, where available. Therefore, we can expect a stronger combined effect of high demand and wind power generation in Denmark, and high demand, wind generation, and solar power generation in Italy, but contrasting scenarios for demand, wind power generation, and solar power generation during the day in Germany and Spain: a low–high–low one (i.e. low demand and solar power generation versus high wind power generation) in the early and late hours versus a high–low–high one (i.e. high demand and solar power generation versus low wind power generation) for peak hours.

4. Forecasting models

We considered univariate and multivariate models for hourly prices with seasonality and with the introduction of exogenous variables relative to the forecasted demand and forecasted RES-E. Furthermore, we included fossil

[9] Given that trends can be observed in the series studied, we tested whether its inclusion does not improve substantially the forecasting performance.

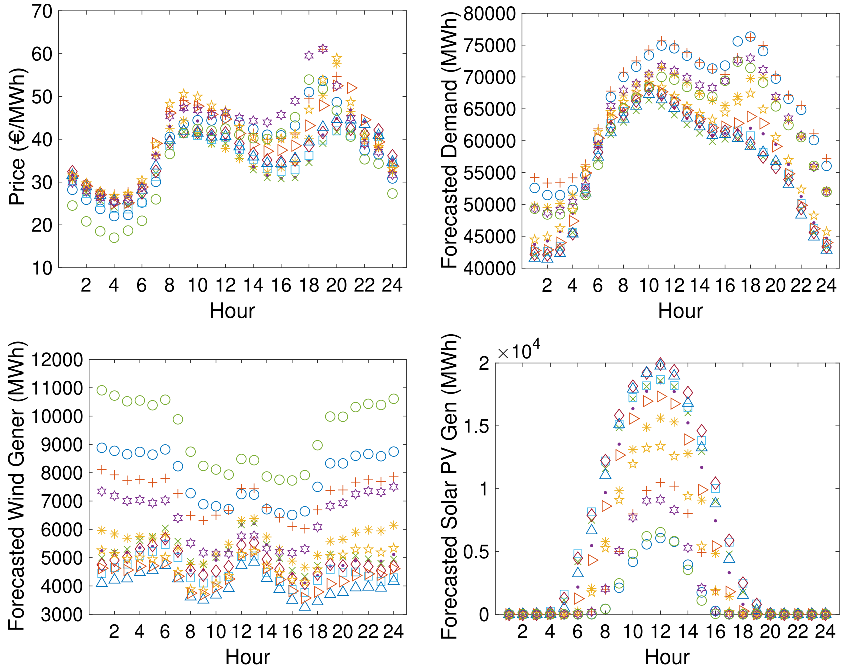

Fig. 2. Intraday profiles of monthly averages for electricity day-ahead prices (top left), forecasted demand (top right), forecasted wind power generation (bottom left), and forecasted solar PV generation (bottom right) observed in Germany. January (blue ∘), February (+), March (⋆), April (•), May (×), June (□), July (⋄), August (△), September (▹), October (⬠), November (○), December (green ∘).

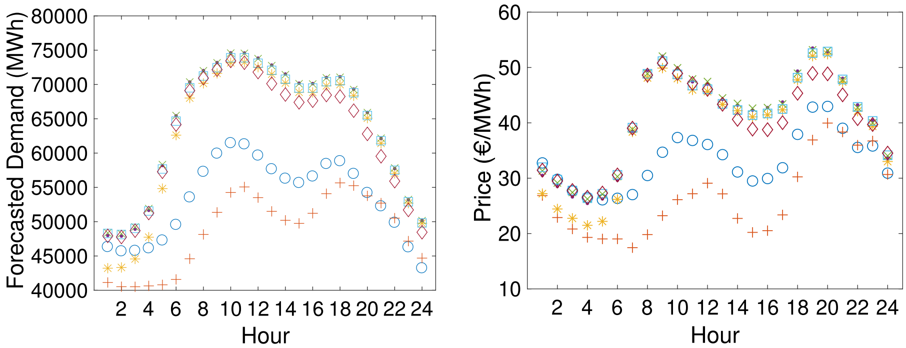

Fig. 3. Intraday profiles across days of the week for forecasted demand (left) and day-ahead electricity prices (right) in Germany. Saturday (∘), Sunday (+), Monday (⋆), Tuesday (•), Wednesday (×), Thursday (□), Friday (⋄).

fuels to account for marginal costs, hence reflecting the non-linearity of the supply curve. Specifically, coal, natural gas, and CO_2 settlement prices were included with a delay of one day, given that market operators know their values determined at the market closure on the day before (i.e. on day $t-1$) when they run their models early in the morning on day t to submit the 24-hour price forecasts by 11 a.m. (of the same day) for trades occurring on the following day, $t+1$.

Therefore, we have specified the following models to compare the forecasting performances when demand, forecasted RES-E, and fossil fuels are taken into account. There is consensus that including demand forecasts or RES-E forecasts (if the market penetration is not negligible) leads to more accurate forecasts. However, it is still an open question as to which RES-E forecast is more informative in which market and also whether their inclusion can reduce the importance of fossil fuels; hence, models with only a subset of exogenous variables were also considered. All together this resulted, first, in our inspecting simple models with only dummy variables for seasonality; second, in our adding regressors accounting for both demand and supply curves (i.e. dummies plus forecasted demand and lagged fossil fuels); third, in our considering if the forecasting ability for demand and RES-E reduces the need to include fuels; and finally, in our verifying if only forecasted wind power generation and/or forecasted solar power generation are/is efficient in providing good price forecasts. We follow common practice in the literature and restrict lags to $t-1$, $t-2$, and $t-7$, which correspond to the previous day, two days before, and one week before the delivery time, recalling,

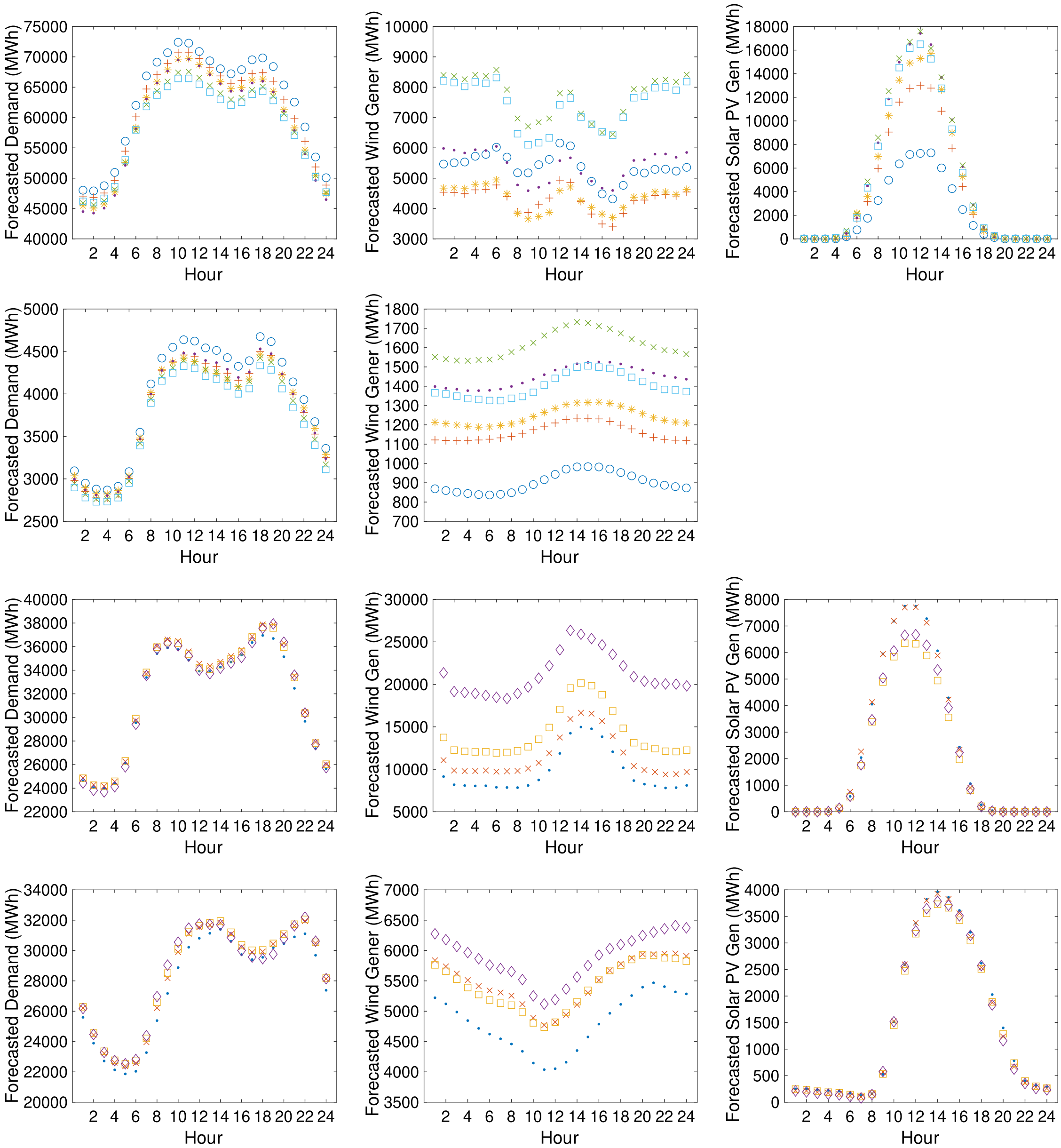

Fig. 4. Intraday profiles of yearly averages for forecasted demand (left), forecasted wind power generation (centre), and forecasted solar PV generation (right) observed in Germany (first row), Denmark (second row), Italy (third row), and Spain (last row). 2011 (blue ∘), 2012 (+), 2013 (⋆), 2014 (•), 2015 (×), 2016 (□).

first, similar conditions that may have characterized the market over the same hours and similar days (such as congestions and blackouts) and, second, the demand level during the days of the week. Knittel and Roberts (2005), Raviv et al. (2015), and Weron and Misiorek (2008) show that these specifications provide accurate forecasts because they capture seasonal patterns in electricity prices. In addition, this formulation reduces the risk of overparameterization. Hence, hourly prices with a reduced 7-lag structure are considered, and, with an abuse of notation in the remainder of the article, $p = 3$ is used in all our univariate and multivariate models to denote the number of lags included instead of the maximum lag.

4.1. Multivariate models

We consider and compare the performances of two different multivariate model specifications with and without exogenous variables used as benchmarks for the corresponding multivariate models. These are the VAR model, the VAR model with exogenous variables (VARX model) estimated by use os ordinary least squares, see Eqs. (2.3.2) and (2.3.4) in Kilian and Lutkepohl (2017), and their Bayesian formulations (BVAR and BVARX models, respectively) with a normal–Wishart prior; see Section 8 in the supplementary material.

4.1.1. VAR model

Let $\mathbf{y}_t = (y_{1t}, \ldots, y_{Ht})'$ denote the $(H \times 1)$ vector of hourly electricity prices, with $H = 24$. Moreover, we denote by $\mathbf{d}_t = (d_{1t}, \ldots, d_{Kt})'$ the $(K \times 1)$ dummy vector, with $(d_{1t}, \ldots, d_{12t})$ representing the 12 months of the year and (d_{13t}, d_{14t}) representing Saturdays and Sundays; hence $K = 14$. The VAR model of order p is formulated as follows:

$$\mathbf{y}_t = \Phi' X_t + \mathbf{e}_t, \quad t = 1, \ldots, T, \tag{1}$$

where Φ is the $((Hp + K) \times H)$ matrix containing the AR coefficients as well as the coefficients for all dummy variables, and $X_t = (\mathbf{y}_{t-1}, \ldots, \mathbf{y}_{t-p}, \mathbf{d}_t)$ is the matrix $((Hp + K) \times H)$ made by the lagged electricity prices and the dummy variables. The vector of errors $\mathbf{e}_t$ is assumed to be serially uncorrelated and normally distributed with zero mean and a full covariance matrix Σ.

4.1.2. VARX model

The VARX model includes the forecasted demand, as well as the forecasted wind and solar power generation, when available, and fossil fuel prices for coal, gas, and CO_2. The exogenous demand and RES variables are represented by the following vectors of dimensions $(H \times 1)$: $\mathbf{x}_t = (x_{1t}, \ldots, x_{Ht})'$, $\mathbf{z}_t = (z_{1t}, \ldots, z_{Ht})'$, and $\mathbf{w}_t = (w_{1t}, \ldots, w_{Ht})'$, respectively. On the other hand, fuel prices do not change over the 24 h and are determined on the previous day, $t - 1$. Thus, m_{t-1}, g_{t-1}, and c_{t-1} are the representations for CO_2, gas, and coal at previous time, respectively. From Eq. (1), we redefine the matrix X_t as $X_t = \left(\mathbf{y}_{t-1}, \ldots, \mathbf{y}_{t-p}, \mathbf{d}_t, \mathbf{x}_t, \mathbf{z}_t, \mathbf{w}_t, m_{t-1}, g_{t-1}, c_{t-1}\right)$ and, consequently, the matrix of coefficients Φ of size $((Hp + K + 3H + 3) \times H)$. The matrix X_t now comprises the vector of lagged hourly electricity prices, the vectors of dummy variables, and the exogenous variables. From Eq. (1), as the observations vary with time $t = 1, \ldots, T$, the VAR and VARX models of order p can be rewritten in a compact way:

$$\mathbf{Y} = \mathbf{X}\Phi + \mathbf{E}, \tag{2}$$

where $\mathbf{Y} = (\mathbf{y}'_1, \ldots, \mathbf{y}'_T)$ is a $(T \times H)$ matrix, and $\mathbf{X} = (X_1, \ldots, X_T)'$ is the $(T \times (Hp + K + 3H + 3))$ matrix of explanatory variables containing all the exogenous variables.[10] The $(T \times H)$ error matrix $\mathbf{E} = (\mathbf{e}'_1, \ldots, \mathbf{e}'_T)$ is normally distributed and serially uncorrelated with covariance matrix Σ.

4.1.3. BVAR models

Our multivariate models with or without exogenous variables were additionally estimated with the Bayesian method. From Eq. (2), a BVAR or BVARX model has the following stacked form:

$$\mathbf{y} = (I_H \otimes \mathbf{X})\boldsymbol{\alpha} + \boldsymbol{\varepsilon}, \tag{3}$$

where $\boldsymbol{\alpha} = \text{vec}(\Phi)$ and $\mathbf{y} = \text{vec}(\mathbf{Y})$ are vectorized matrices and $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \Sigma \otimes I_T)$, with I_T being a T-dimensional identity matrix. This stacked-form representation allows us to define and study the prior and posterior distribution of the matrix of coefficients and the covariance matrix, leading to a closed-form distribution. In particular, we define prior information on the matrix of coefficients and on the covariance matrix using a conjugate normal–Wishart prior.[11]

4.2. Univariate models

For all previous models, we formulated 24 (parsimonious) univariate AR specifications with the same assumptions on the lag order of the VAR specifications, whereas the errors were assumed to be normally distributed with zero mean and σ_h^2 variance for the hours $h = 1, \ldots, 24$. The AR model with only dummy variables was used as benchmark in the forecasting comparisons, and can be written as follows:

$$y_{h,t} = \sum_{l=1}^{p} \phi_l y_{h,t-l} + \sum_{k=1}^{K} \psi_k d_{kt} + \varepsilon_{h,t}.$$

On the other hand, the univariate frequentist and Bayesian AR models extended with exogenous variables, ARX and BARX models respectively, can be written as:

$$y_{h,t} = \sum_{l=1}^{p} \phi_l y_{h,t-l} + \sum_{k=1}^{K} \psi_k d_{kt} + \alpha_1 x_{ht} + \alpha_2 z_{ht} + \alpha_3 w_{ht} + \beta_1 m_{t-1} + \beta_2 g_{t-1} + \beta_3 c_{t-1} + \varepsilon_{h,t},$$

where x_{ht}, z_{ht}, and w_{ht} represent (forecasted) demand and renewable energy variables, whereas m_{t-1}, g_{t-1}, and c_{t-1} are the fossil fuel prices previously described. Even in the univariate case, we used both the frequentist estimation procedure and the Bayesian estimation procedure.

To support the multivariate formulation, we run 24 univariate models with dummies, lags of y_t, and fundamentals lagged same-hour prices, adding also the first lag of all other remaining hours; that is,

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

Then, from the resulting 24 residual series of each model, $\hat{\varepsilon}_{h,t}$, the variance–covariance matrix was computed. Uncorrelated residuals make the multivariate VAR specification unnecessary; however, we found evidence of large correlations across all markets studied.[12] Therefore, a VAR model with full covariance matrix Σ seems more appropriate to estimate this covariance structure and it should result in increased density forecast accuracy.

[10] We also performed the forecasting exercises including the lags (1, 2, and 7) for exogenous variables, but the results were unchanged although computationally intensive and time demanding. For these reasons, and to have proper forecasts, we prefer to use the former models without lagged exogenous variables.

[11] We performed the analysis using both a standard Minnesota prior and the normal–Wishart priors, and the results were similar. Therefore, because of lack of space, we report only the results for the latter. Details on the prior information and posterior distribution are reported in Section 8 in the supplementary material.

[12] These results have been omitted for lack of space, but they are available on request.

4.3. Forecast assessment

We assessed the goodness of our forecasts using different point and density metrics. Considering the accuracy of point forecasts, we used the root-mean-square errors (RMSEs) for each of the hourly prices, as well as the RMSEs of the daily average and of an average restricted only to central hours, as specified below. The RMSE for $h = 1, \ldots, 24$ hourly prices is computed as

$$\mathrm{RMSE}_h = \sqrt{\frac{1}{T-R}\sum_{t=R}^{T-1}\left(\hat{y}_{h,t+1|t} - y_{h,t+1}\right)^2}, \tag{4}$$

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. In addition, we analysed the average RMSEs for all 24 h ($\mathrm{RMSE}_{\mathrm{Avg}}$) and for the hours from 8 a.m. to 8 p.m. (peak hours, $\mathrm{RMSE}^P_{\mathrm{Avg}}$), computed as follows:

$$\mathrm{RMSE}_{\mathrm{Avg}} = \frac{1}{24}\sum_{h=1}^{24}\mathrm{RMSE}_h, \tag{5}$$

$$\mathrm{RMSE}^P_{\mathrm{Avg}} = \frac{1}{13}\sum_{h=8}^{20}\mathrm{RMSE}_h. \tag{6}$$

To evaluate density forecasts, we used the average continuous ranked probability score (CRPS).[13]

As indicated in Gneiting and Raftery (2007) and Gneiting and Ranjan (2011), some researchers view the CRPS as having advantages over the log score. In particular, the CRPS does a better job of rewarding values from the predictive density that are close to – but not equal to – the outcome, and it is less sensitive to outlier outcomes. The CRPS, defined such that a lower number is a better score, is given by

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

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. In the same way we can construct the average CRPS over the 24 h and over peak hours on day $t+1$.

More specifically, we report the RMSEs and the average CRPS for all the univariate and multivariate models and for every third hour.[14]

In addition, we apply Diebold–Mariano t tests Diebold and Mariano (1995) for equality of the average loss (with loss defined as the squared error or CRPS) to compare predictions of alternative models with the benchmark for a given horizon h.[15] 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 with use of the prewhitened 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 underrejection of the null hypothesis of equal predictability. We report p values based on one-sided tests, taking the AR (VAR) as the null and the other current models as the alternative.

Finally, we applied the model confidence set procedure of Hansen, Lunde, and Nason (2011) across models for a fixed horizon to jointly compare their predictive power without disentangling univariate and multivariate models. The R package MCS detailed in Bernardi and Catania (2016) was used, and the differences were tested separately for each hour and model, with the full process repeated across all countries. The results are discussed in the following section.

5. Results

Our results are based on a one-step-ahead forecasting process with a rolling window approach of four years for Germany and Denmark and of two years for Italy and Spain. We recall that we have two estimation samples: 1 January 2011 to 31 December 2014 for Germany and Denmark and 13 June 2014 to 13 June 2016 for Italy and Spain. We have two forecast evaluation periods: 1 January 2015 to 31 December 2016 for Germany and Denmark (for a total of 731 observations) and 14 June 2016 to 13 June 2017 for Italy and Spain (hence only 365 observations).

Before we evaluate the out-of-sample results, our in-sample evidence provides statistically significant coefficients for the RES variables in all markets, hence confirming the empirical findings in previous literature on univariate models augmented with RES variables and extending similar conclusions also to multivariate models. In particular, the coefficients for forecasted wind and forecasted solar power are negative in Germany, Italy, and Spain. Also in Denmark, wind power has a negative coefficient.[16] These results confirm that RES are significantly connected to and reduce electricity prices. Therefore, we continue our analysis by investigating whether these relationships can result in forecast gains.

To this end, our results show the performance of our different univariate and multivariate models, from the simplest ones (with only dummy variables, the benchmarks) to more complex ones containing gas, coal, CO_2, and forecasts for demand, wind power, and solar power. Alternative formulations referred to subsets of drivers are described in Table 3 in the supplementary material, and the results are summarized in Tables 1 and 2 across all

[13] We also found that the log predictive score and results were similar; hence, they have not been reported.

[14] Tables with all hours are available in Sections 10–13 in the supplementary material.

[15] In our application for testing density forecasts, we used equal weights without adopting a weighting scheme, as in Amisano and Giacomini (2007).

[16] Detailed in-sample results are available on request.

Table 1
RMSE values for AR (VAR) benchmark models, and RMSE ratios for other models.

	Hour									
	1	4	7	10	13	16	19	22	Avg	Avg_{8-20}
Germany										
AR	7.240	7.387	8.027	8.905	9.214	9.669	8.692	6.277	8.259	9.333
ARX (FD+RES+Fuels)	5.336***	5.939**	6.430***	6.928***	6.662***	7.068***	6.867***	4.871***	6.326	7.065
BAR	7.226**	7.387	8.011**	8.887***	9.214	9.659***	8.666***	6.258***	8.251	9.314
BARX (FD+RES+Fuels)	5.329***	5.932**	6.430***	6.768***	6.597***	7.068***	6.728***	4.821***	6.260	6.972
VAR	4.278	4.944	6.271	6.905	7.934	8.350	8.290	6.164	6.839	7.993
VARX (FD+RES+Fuels)	4.492	5.404	6.396	6.083***	6.315***	7.039***	6.715***	4.654***	5.964	6.698
BVAR	4.282	4.939	6.271	6.912	7.934	8.342	8.290	6.158	6.839	7.993
BVARX (FD+RES+Fuels)	4.445	5.414	6.415	6.035***	6.276***	6.964***	6.591***	4.629***	5.923	6.642
Denmark										
AR	5.850	6.566	6.857	13.162	8.226	8.029	10.300	6.012	8.468	10.465
ARX (FD+RES+Fuels)	5.376	6.211	6.082***	9.661***	6.893***	6.544***	8.240***	5.345*	6.969	8.121
BAR	5.838**	6.553***	6.843***	13.096***	8.210***	8.021***	10.279***	6.000***	8.451	10.444
BARX (FD+RES+Fuels)	5.382	6.211	6.089***	9.635***	6.885***	6.552***	8.219***	5.339*	6.961	8.110
VAR	3.413	4.131	5.159	10.913	7.607	7.466	10.441	5.897	7.197	9.328
VARX (FD+RES+Fuels)	4.386	5.684	5.690	10.553	6.778**	6.421***	8.750***	5.213***	6.974	8.460
BVAR	3.413	4.135	5.164	10.913	7.599***	7.451***	10.431**	5.891**	7.190	9.319
BVARX (FD+RES+Fuels)	4.386	5.680	5.690	10.553	6.755***	6.421***	8.760***	5.201***	6.974	8.451
Italy										
AR	4.560	4.405	5.162	9.107	6.389	8.122	9.835	7.104	6.838	8.331
ARX (FD+RES+Fuels)	4.410	4.185	4.966	8.561	5.776***	7.513**	9.638	6.862	6.469	7.806
BAR	4.551	4.401	5.157	9.098**	6.389	8.122*	9.835	7.111	6.831	8.331
BARX (FD+RES+Fuels)	4.432	4.238	4.961*	8.524**	5.756***	7.488***	9.560*	6.870	6.455	7.764
VAR	3.884	4.105	4.753	8.569	6.144	7.650	9.582	6.893	6.507	7.901
VARX (FD+RES+Fuels)	4.094	4.228	4.881	8.098	5.511***	7.107**	9.668	6.996	6.357	7.530
BVAR	3.880	4.101	4.753	8.569	6.150	7.650	9.572**	6.893*	6.507	7.901
BVARX (FD+RES+Fuels)	4.051	4.142	4.853	8.106	5.523***	7.061**	9.649	6.934	6.325	7.506
Spain										
AR	7.036	6.741	7.066	6.421	6.140	6.873	5.574	4.628	6.299	6.389
ARX (FD+RES+Fuels)	5.692***	5.251***	5.483***	5.162***	4.992***	5.471***	4.922**	4.721	5.197	5.201
BAR	7.029*	6.728***	7.045***	6.382***	6.122***	6.852***	5.541***	4.609***	6.274	6.363
BARX (FD+RES+Fuels)	5.755***	5.379***	5.688***	5.419***	5.145***	5.601***	5.033**	4.813	5.335	5.360
VAR	3.943	4.638	5.227	4.761	5.018	5.908	5.363	4.823	5.110	5.317
VARX (FD+RES+Fuels)	3.947	4.360	4.406***	4.071***	4.140***	4.443***	4.242***	4.138**	4.262	4.259
BVAR	3.943	4.638	5.227	4.751	5.013	5.890**	5.325***	4.794***	5.100	5.301
BVARX (FD+RES+Fuels)	3.982	4.415*	4.474***	4.094***	4.160***	4.472***	4.274***	4.148**	4.298	4.286

FD, forecasted demand.
Notes:
[1]Forecast errors are calculated using a rolling window estimation. 'Avg' and 'Avg_{8-20}' stand for RMSEs computed as in Eqs. (5) and (6).
[2]See Section 4 for details on model formulations. 'X' indicates models with exogenous variables, while 'B' represents Bayesian conjugate normal–Wishart priors.
[3]***, **, and * indicate RMSE ratios are significantly different from 1 at the 1%, 5%, and 10% significance levels according to the Diebold–Mariano test.
[4]Grey cells indicate those models that belong to the superior set of models resulting from the model confidence set procedure at a confidence level of 10%.

markets at every third hour, whereas extensive comparisons are shown in Tables 4–11 in the supplementary material.

Recalling the main objectives of this analysis, we have shown clearly the superiority of multivariate models when the full structure of 24 h is considered. Multivariate VAR models outperform simple AR models when only seasonality is included. This holds true systematically across all countries, and according to both point metrics and density metrics. For instance, in Germany, the average RMSE changes from 8.259 €/MWh in the univariate case to 6.839 €/MWh in the multivariate case. In Spain, it changes from 6.299 €/MWh to 5.110 €/MWh. The case is similar for the CRPSs, for which we observe substantial reductions of almost 18% (from 4.427 to 3.643) in Germany, 13% (from 4.901 to 4.273) in Denmark, 5% (from 3.658 to 3.469) in Italy, and 20% (from 3.517 to 2.831) in Spain for average values computed over the 24 h. The AR models are included in the model confidence set in only ten cases over 64 horizons in Tables 1 and 2 for both metrics and mainly for the Italian market. VAR models have a much higher frequency of inclusion. Hence, this supports our expectations of more efficient forecasts obtained by considering the interrelationships among the whole 24 h, as suggested by Stock and Watson (2002) and anticipated by Raviv et al. (2015).

Moreover, considering the most important fact, which is the forecasting improvements resulting from the inclusion of RES and/or a subset of drivers, our results show that the Bayesian multivariate models with forecasted RES-E and fuels exhibit substantial improvements generally in all markets. The average reductions in the loss function are similar for both metrics and range from 10% to 20% in Germany and Spain and from 1% to 5% in Denmark and Italy. Forecast gains increase in the peak hours, as shown in the columns with header 'Avg_{8-20}'. When

Table 2
Average CRPS for AR (VAR) benchmark models, and CRPS ratios for other models.

	Hour									
	1	4	7	10	13	16	19	22	Avg	Avg_{8-20}
Germany										
AR	3.770	4.062	4.467	4.942	4.962	4.970	4.792	3.423	4.427	4.964
ARX (FD+RES+Fuels)	2.926***	3.420***	3.672***	3.781***	3.563***	3.593***	3.786***	2.663***	3.422	3.733
BAR	3.770	4.062	4.458*	4.927***	4.957*	4.960**	4.768***	3.409***	4.418	4.954
BARX (FD+RES+Fuels)	2.926***	3.420***	3.672***	3.702***	3.528***	3.598***	3.719***	2.646***	3.391	3.688
VAR	2.261	2.901	3.443	3.772	4.185	4.208	4.525	3.347	3.643	4.173
VARX (FD+RES+Fuels)	2.390	3.040	3.443	3.316***	3.294***	3.509***	3.692***	2.570***	3.169	3.497
BVAR	2.254**	2.886***	3.426***	3.764**	4.177**	4.191***	4.507***	3.334***	3.632	4.160
BVARX (FD+RES+Fuels)	2.358	3.037	3.426	3.282***	3.260***	3.472***	3.615***	2.554***	3.137	3.451
Denmark										
AR	3.236	3.690	3.896	8.844	4.400	4.156	5.379	3.188	4.901	6.199
ARX (FD+RES+Fuels)	2.997***	3.446***	3.471***	7.517***	3.670***	3.387***	4.255***	2.866***	4.210	5.151
BAR	3.230**	3.683**	3.892	8.817***	4.387***	4.148**	5.368**	3.182**	4.891	6.187
BARX (FD+RES+Fuels)	2.997***	3.443***	3.475***	7.509***	3.665***	3.387***	4.249***	2.866***	4.210	5.151
VAR	2.019	2.644	3.035	7.829	3.925	3.761	5.329	3.100	4.273	5.606
VARX (FD+RES+Fuels)	2.314	3.093	3.199	7.727*	3.556***	3.295***	4.487***	2.737***	4.119	5.219
BVAR	2.007***	2.631***	3.020***	7.790***	3.901***	3.731***	5.308***	3.088***	4.247	5.572
BVARX (FD+RES+Fuels)	2.300	3.067	3.178	7.743*	3.532***	3.268***	4.460***	2.725***	4.106	5.208
Italy										
AR	2.547	2.460	2.851	4.772	3.494	4.390	5.037	3.616	3.657	4.413
ARX (FD+RES+Fuels)	2.476	2.362*	2.731**	4.514**	3.141***	4.100***	4.992	3.522**	3.481	4.157
BAR	2.539**	2.455	2.854	4.767	3.491	4.386	5.032	3.609**	3.653	4.409
BARX (FD+RES+Fuels)	2.488	2.386	2.720**	4.457***	3.113***	4.039***	4.881**	3.511**	3.449	4.095
VAR	2.147	2.265	2.588	4.509	3.342	4.095	4.954	3.569	3.469	4.177
VARX (FD+RES+Fuels)	2.291	2.378	2.694	4.356	3.008***	3.812**	5.177	3.744	3.455	4.043
BVAR	2.145	2.263	2.580*	4.482***	3.329**	4.075***	4.924***	3.551***	3.452	4.156
BVARX (FD+RES+Fuels)	2.259	2.310	2.653	4.338	2.991***	3.763***	5.157	3.708	3.417	4.014
Spain										
AR	3.914	3.757	3.948	3.555	3.402	3.844	3.139	2.669	3.517	3.556
ARX (FD+RES+Fuels)	3.112***	2.915***	3.064***	2.901***	2.786***	3.037***	2.734***	2.613	2.891	2.895
BAR	3.910	3.746	3.952	3.541	3.388	3.832	3.117	2.658	3.503	3.538
BARX (FD+RES+Fuels)	3.151***	2.987***	3.190***	3.040***	2.858***	3.106***	2.787***	2.666	2.961	2.973
VAR	2.128	2.536	2.897	2.644	2.743	3.267	3.019	2.745	2.831	2.945
VARX (FD+RES+Fuels)	2.145	2.381*	2.422***	2.239***	2.271***	2.447***	2.364***	2.306***	2.350	2.350
BVAR	2.122	2.526	2.885	2.631	2.724	3.244	2.989	2.720	2.814	2.924
BVARX (FD+RES+Fuels)	2.164	2.414	2.471***	2.250***	2.263***	2.440***	2.373***	2.306***	2.361	2.353

FD, forecasted demand.
Notes:
[1]Forecast errors are calculated by a rolling window estimation. 'Avg' and 'Avg_{8-20}' stand for average CRPS for average values.
[2]See Section 4 for details on model formulations. 'X' indicates models with exogenous variables, while 'B' represents Bayesian conjugate normal–Wishart priors.
[3]***, **, and * indicate average score ratios are significantly different from 1 at the 1%, 5%, and 10% significance levels according to the Diebold–Mariano test.
[4]Grey cells indicate those models that belong to the superior set of models resulting from the model confidence set procedure at a confidence level of 10%.

we focus on each individual hour, BVARX models statistically outperform VAR models, and they are included in the model confidence set in most of the cases, and almost always for late morning, afternoon, and evening hours. VARX models also perform accurately, but give some economically smaller gain than do BVARX models.

Going into detail and exploring the forecasting ability of several models with different combinations of variables to inspect their individual contribution, we first find evidence of forecasting improvements when demand and all RES are included. Moreover, the BVAR model with only forecasted power wind generation (besides forecasted demand and fuels) leads to better forecasts than those obtained with the inclusion of only forecasted solar power generation (besides forecasted demand and fuels), especially for point forecasts over hours 8–24 in Germany, Italy, and Spain (not performed in Denmark because there is no available solar power). From comparison of the ability of the BVAR model with forecasted demand and RES with that containing forecasted demand and fuels, the former is found to perform better. However, there are further gains when all these exogenous regressors are considered simultaneously.[17]

[17] This may be due to the contribution of individual fuels. For instance, in an additional analysis within the frequentist approach, we observed that models with selected fuels, forecasted demand, and RES show slight improvements in the RMSEs: in Germany, the inclusion of both CO_2 and coal increases the forecast accuracy over hours 8–24; in Denmark, the inclusion of only gas improves the forecast accuracy during hours 8–12, whereas the inclusion of coal increases the forecast accuracy over hours 13–24. In Italy, coal and gas together are important during ramp-up and ramp-down hours (9–10 and 18–19), whereas only gas is important during hours 11–17; this is consistent with Italy's dependence on thermal generation (and so on traditional fuels), given the still marginal penetration of RES (compared with the other countries studied). In Spain, the inclusion of coal slightly and generally

6. Conclusions

We have compared the forecasting performances of linear univariate and multivariate models with enlarged specifications. Our set of models include AR and VAR models with only dummy variables for seasonality, which are used as a baseline for the corresponding formulations enlarged by including also fuels, forecasted demand, and forecasted RES, analysed from both the frequentist preservative and the Bayesian perspective.

Our results indicate that models with demand, RES, and fuels dominate those without fuels and forecasted RES in terms of both point forecasting and density forecasting. In particular, the first important finding is that the multivariate models outperform the univariate ones, given that they allow for interrelationships among different hours of the day. Secondly, the Bayesian approach leads to further forecasting improvements. Thirdly, and for the first time since the increase in RES penetration, we show that the models with only forecasted wind power generation perform better than those with solar power generation only. Their simultaneous inclusion further improves the performance.

We also provide strong empirical evidence for the influence of renewable power generation during the day, and consistently with the country intraday profiles. During the first hours of the day, the models without forecasted RES-E are more accurate than those with it, and again with errors from multivariate models being lower than those from univariate ones. In contrast the increasing amount of RES-E during the day leads to more accurate forecasts from augmented models. Furthermore, our results are consistent across all scoring rules used, such as the RMSE and the CRPS.

From an energy forecasting perspective these linear multivariate AR models with forecasted RES, forecasted demand, and fuels seem to have interesting and important advantages over the widely used univariate ones. It is worth emphasizing the increasing relevance of density forecasting since in recent years market operators have been exploring opportunistic bidding across market sessions, as emphasized by Bunn et al. (2018). Indeed, forecasting the day-ahead prices is important for market operators and traders to plan their strategy. For example, arbitrage opportunities can be explored by deciding on in which market session to bid according to the forecasted day-ahead prices. For this reason, energy regulatory authorities are trying to formulate optimal pricing rules to avoid these market inefficiencies. Agents operating units responsible for balancing are exposed to economic consequences from differentials between day-ahead and balancing prices, which are used to evaluate the actual unit imbalance according to the sign of the system imbalance. In simple words, if one unit is short imbalanced when the market is long (or long imbalanced when the market is short), it receives profits for relieving the system (which are computed on the basis of price differentials). Otherwise, if the unit and the system have signs in agreement, the unit receives penalties because it increases the system imbalance.

All these considerations clearly show the extreme relevance of both point forecasting and density forecasting for these day-ahead electricity prices, and our results highlight that the Bayesian multivariate models with the drivers considered improve them substantially.

increases the forecast accuracy, which is, however, comparable with the model with all fuels at selected hours (12–15 and 23–24). In all these cases, adding the omitted fuels results in only a very small reduction in the performances, hence supporting the conclusion of the overall importance of all fossil fuels when one is forecasting day-ahead electricity prices. These results are omitted for lack of space, but they are available on request.

Acknowledgments

The authors thank the co-editor, and seminar and conference participants at Ca' Foscari University of Venice, the EU Joint Research Centre in Ispra, the University of Warwick, the 26th Annual Symposium of the Society for Nonlinear Dynamics and Econometrics in Tokyo, the 15th Conference on Computational Management Science in Trondheim, the 12th Annual RCEA Bayesian Econometric Workshop in Rimini, and the 8th Energy Finance Christmas Workshops in Bozen for useful comments and suggestions. This research used the SCSCF multiprocessor cluster system at Ca' Foscari University of Venice. Europe Energy S.p.A. is acknowledged for funding this research project. In addition, Angelica Gianfreda acknowledges the RTDcall2017 support for the project *Forecasting and Monitoring Electricity Prices, Volumes and Market Mechanisms*, funded by the Free University of Bozen-Bolzano, Italy. Luca Rossini acknowledges financial support from the EU Horizon 2020 programme under the Marie Skłodowska-Curie scheme (grant agreement no. 796902).

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.ijforecast.2019.11.002.

References

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

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

Bernardi, M., & Catania, L. (2016). Portfolio optimisation under flexible dynamic dependence modelling.

Bunn, D. W., Gianfreda, A., & Kermer, S. (2018). A trading-based evaluation of density forecasts in a real-time electricity market. *Energies*, *11*(10), 2658–2670.

Chen, D., & Bunn, D. (2014). The forecasting performance of a finite mixture regime-switching model for daily electricity prices. *Journal of Forecasting*, *33*(5), 364–375.

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

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

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

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

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

García-Martos, C., Rodríguez, J., & Sánchez, M. J. (2007). Mixed models for short-run forecasting of electricity prices: Application for the Spanish market. *IEEE Transactions on Power Systems*, *22*(2), 544–552.

Gelabert, L., Labandeira, X., & Linares, P. (2011). An ex-post analysis of the effect of renewables and cogeneration on spanish electricity prices. In *Supplemental issue: Fourth Atlantic workshop in energy and environmental economics Energy Economics*, *33*(Suppl. 1), S59–S65

Gianfreda, A., & Bunn, D. (2018). A stochastic latent moment model for electricity price formation. *Operations Research*, *66*, 1189–1456.

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

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

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

Gianfreda, A., Parisio, L., & Pelagatti, M. (2019). The RES-induced switching effect across fossil fuels: An analysis of day-ahead and balancing prices. *The Energy Journal*, *40*(The New Era of Energy Transition), 365–386.

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

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

Hansen, P. R., Lunde, A., & Nason, J. M. (2011). The model confidence set. *Econometrica*, *79*, 453–497.

Hirth, L., & Ziegenhagen, I. (2015). Balancing power and variable renewables: Three links. *Renewable & Sustainable Energy Reviews*, *50*, 1035–1051.

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

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

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

Karakatsani, N. V., & Bunn, D. W. (2008). Forecasting electricity prices: The impact of fundamentals and time-varying coefficients. *International Journal of Forecasting*, *24*(4), 764–785.

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

Kilian, L., & Lutkepohl, H. (2017). *Structural vector autoregressive analysis*. Cambridge University Press, Themes in modern econometrics.

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

Koopman, S. J., Ooms, M., & 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.

Lisi, F., & Edoli, E. (2018). Analyzing and forecasting zonal unbalancing signs in the Italian electricity market. *The Energy Journal*, *39*(5), 1–19.

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

Maciejowska, K., & Weron, R. (2015). Forecasting of daily electricity prices with factor models: utilizing intra-day and inter-zone relationships. *Computational Statistics*, *30*(3), 805–819.

Maciejowska, K., & Weron, R. (2016). Short- and mid-term forecasting of baseload electricity prices in the U.K.: The impact of intra-day price relationships and market fundamentals. *IEEE Transactions on Power Systems*, *31*(2), 994–1005.

Martinez-Anido, C. B., Brinkman, G., & Hodge, B.-M. (2016). The impact of wind power on electricity prices. *Renewable Energy*, *94*(Suppl. C), 474–487.

Mauritzen, J. (2013). Dead battery? Wind power, the spot market, and hydropower interaction in the Nordic electricity market. *Energy Journal*, *34*(1), 103–123.

Misiorek, A., Trueck, S., & Weron, R. (2006). Point and interval forecasting of spot electricity prices: linear vs. non-linear time series models. *Studies in Nonlinear Dynamics & Econometrics*, *10*(2), 1–36.

Nowotarski, J., & Weron, R. (2018). Recent advances in electricity price forecasting: A review of probabilistic forecasting. *Renewable & Sustainable Energy Reviews*, *81*(Part 1), 1548–1568.

Panagiotelis, A., & 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., & Pietsch, R. (2014). The impact of renewable energies on EEX day-ahead electricity prices. *Energy Policy*, *73*, 196–210.

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

Poplavskaya, K., & de Vries, L. (2019). Distributed energy resources and the organized balancing market: A symbiosis yet? Case of three European balancing markets. *Energy Policy*, *126*, 264–276.

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

REN21 (2017). *Renewables 2017 - global status report*, Paris: REN21 Secretariat, ISBN: 978-3-9818107-6-9.

Rintamäki, T., Siddiqui, A. S., & Salo, A. (2017). Does renewable energy generation decrease the volatility of electricity prices? An analysis of Denmark and Germany. *Energy Economics*, *62*, 270–282.

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

Uniejewski, B., Nowotarski, J., & Weron, R. (2016). Automated variable selection and shrinkage for day-ahead electricity price forecasting. *Energies*, *9*(621), 1–22.

Weron, R. (2014). Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International Journal of Forecasting*, *30*(4), 1030–1081.

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

Woo, C., Horowitz, I., Moore, J., & Pacheco, A. (2011). The impact of wind generation on the electricity spot-market price level and variance: The Texas experience. *Energy Policy*, *39*(7), 3939–3944, Special Section: Renewable energy policy and development.

Ziel, F., & Weron, R. (2018). Day-ahead electricity price forecasting with high-dimensional structures: Univariate vs. multivariate modeling frameworks. *Energy Economics*, *70*, 396–420.