



Forecasting electricity spot prices using time-series models with a double temporal segmentation

Marie Bessec, Julien Fouquau, Sophie Méritet

► To cite this version:

Marie Bessec, Julien Fouquau, Sophie Méritet. Forecasting electricity spot prices using time-series models with a double temporal segmentation. 2nd International Symposium on Energy and Finance Issues (ISEFI-2014), Mar 2014, Paris, France. pp.34. hal-01502835

HAL Id: hal-01502835

<https://hal.science/hal-01502835>

Submitted on 6 Apr 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Forecasting electricity spot prices using time-series models with a double temporal segmentation

MARIE BESSEC* JULIEN FOUQUAU† SOPHIE MERITET‡

May 21, 2014

Abstract: The French wholesale market is set to expand in the next few years under European pressure and national decisions. In this paper, we assess the forecasting ability of several classes of time series models for electricity wholesale spot prices at a day-ahead horizon in France. Electricity spot prices display a strong seasonal pattern, particularly in France given the high share of electric heating in housing during winter time. To deal with this pattern, we implement a double temporal segmentation of the data. For each trading period and season, we use a large number of specifications based on market fundamentals: linear regressions, Markov-switching models, threshold models with a smooth transition. An extensive evaluation on French data shows that modeling each season independently leads to better results. Among non-linear models, MS models designed to capture the sudden and fast-reverting spikes in the price dynamics yield more accurate forecasts. Finally, pooling forecasts gives more reliable results.

Keywords: Electricity spot prices, forecasting, regime-switching.

JEL classification: C22, C24, Q47.

*Corresponding author: LEDa-CGEMP Université Paris Dauphine, Place du Maréchal de Lattre de Tassigny 75016 Paris France, Tel.: 33 1 44 05 44 64, email: marie.bessec@dauphine.fr.

†NEOMA Business School and LEDa-CGEMP Université Paris Dauphine, email: julien.fouquau@neoma-bs.fr.

‡LEDa-CGEMP Université Paris Dauphine, email: sophie.meritet@dauphine.fr.

Introduction

In Europe, the reorganization of electricity markets is an ongoing project, with increasing market integration and competitive wholesale markets. The central-western European market (covering Benelux, France and Germany) is at the heart of power trading in continental Europe and it provides a blueprint for plans to create a single European electricity market.

The French wholesale market is the third largest in Europe after the German and British ones, and is about to change drastically. In France, it is more than challenging for new entrants to produce electricity given the very low cost of production of nuclear power plants managed by the historic utility. To open up the French market to more competition, the new national electricity law, passed in December 2010, compels Electricité de France (EDF), the biggest producer of electric power in Europe and former domestic vertically integrated public monopoly, to sell a proportion of its nuclear output to its rivals. The French government also gives alternative suppliers access to hydroelectric generation and modifies regulated rates for wholesale nuclear power. Moreover, following a European Commission decision, the regulated rates for electricity will discontinue in 2016 for small businesses on the retail market. Therefore, the number of electricity contracts offered by suppliers with market prices will increase. That is why the French wholesale market should become more liquid with more participants in the near future.

It is highly relevant to study the wholesale electricity prices in France at this stage. This paper deals more specifically with forecasting French spot prices at a daily horizon. This is a particularly challenging issue due to the specific properties of electricity prices. First, since electricity is not storable, it has to be delivered as soon as it is generated. Excess demand at a given time cannot be compensated by excess supply a few hours before. This characteristic is responsible for price spikes and strong volatility; it can even lead to negative prices. Second, spot prices display a strong seasonal pattern at daily, weekly, and yearly levels. This pattern is particularly important in France due to the high thermal sensitivity of electricity consumption.

In France, electricity consumption is highly dependent on temperature. This dependence has increased significantly over the past decade. It is now estimated at 2 300 MW/°C at 7 pm, the peak hour of daily consumption in winter. The thermal sensitivity is much higher in France than in other European countries. For instance, the German electricity consumption increases only by 500 MW/°C at 7 pm in winter (RTE [2012]). This phenomenon is partly explained by the high rate of electric heating equipment in French housing. Electric heating has widely developed in France since the 70s. Since 2005, more

than 50% of new houses are equipped with electric heating due to the low off-peak tariffs offered and in 2009, this rate even reached 70%. In 2010, 35% of the total stock of housing (new and old) was heated with electricity, i.e. about 9.5 million homes (RTE [2012]).

In the literature, there has been considerable work on modeling and forecasting spot prices. Basically, three groups of methods are used for this purpose: equilibrium or game theory approach, simulation methods, and times series forecasting methods. The latter approach uses the past dynamics of prices and eventually some exogenous variables to forecast electricity prices. For this purpose, linear and non-linear models (Markov-switching, threshold, time-varying parameter regressions) are used. Many authors rely only on the past dynamics of prices, while others also include fundamental variables in the model such as climatic variables, forecasted demand and capacity margin. In addition, a strand of this literature allows a time-varying conditional volatility using the class of GARCH models. See Liu and Shi [2013] for a recent survey.

In this paper, we further explore in an extensive way the relevance of time series models for day-ahead forecasting of power prices in France. In particular, we address several issues which are still open when modeling and forecasting electricity spot prices.

A first point concerns the treatment of seasonality in the data. This question is particularly important for the French market given the high sensitivity of electricity consumption to temperature, as explained above. In this paper, we implement a double temporal segmentation of the data, to account for the intra- and inter-day seasonality. Prices display a distinct pattern depending on the hour of the day. Moreover, their level and volatility are much higher in fall and in winter. Therefore, the empirical analysis is conducted separately for each trading hour and for each season. The hourly segmentation is now common in literature (e.g. Crespo Cuaresma, Hlouskova, Kossmeier and Obersteiner [2004], Misiorek, Trueck and Weron [2006], Karakatsani and Bunn [2008], Bordignon, Bunn, Lisi and Nan [2013] in the context of power price forecasting). However, to our knowledge, seasonal segmentation has not been applied for power price modeling and forecasting. Other more common approaches to deal with the long-run seasonality are described in Nowotarski, Tomczyk and Weron [2013b] and include the use of dummy variables, sinusoidal functions and wavelets.

Moreover, we investigate whether nonlinear models perform better than the linear ones for price forecasting, and if so, whether latent variable methods are better than threshold methods. We use regime switching models which are more likely to capture the sudden and fast reverting spikes in the price dynamics. We consider two widely used classes of non-linear models: threshold regressive models with a smooth transition

and Markov-switching models with fixed and time-varying transition probabilities. We employ various specifications to capture the jumps of prices and a possible modification of the relationship between prices and explanatory variables in the extreme regimes. In the threshold and MS models with time-varying probabilities, the transitional variable is selected for each trading hour and each season according to the results of the linearity test. A few related papers provide evidence in favor of regime-switching models for power price forecasting. Karakatsani and Bunn [2008] find that Markov-switching models and models with time-varying coefficients have a better predictive performance in the British market, while the threshold class of models outperforms the other specifications for the forecasting of Californian spot prices in Misiorek et al. [2006]. Kosater and Mosler [2006] also obtain better long-run forecasts of German prices with Markov-switching models.

Finally, we explore the benefit of combining forecasts over individual predictors. We compare the performance of individual models, as well as combinations of forecasts with various weighting schemes. Despite a long tradition in the use of forecast combination (Bates and Granger [1969] and Newbold and Granger [1974]) and a renewed interest in the forecasting literature in the last decade, the use of combination of forecasts is recent in the context of electricity prices and gives positive results in this area too. The existing papers provide evidence in favor of combinations for several power markets: the UKPX power exchange (Bordignon et al. [2013]), Nord Pool (Raviv, Bouwman and van Dijk [2012]), the European Energy Exchange and the Pennsylvania-New Jersey-Maryland Interconnection (Nowotarski, Raviv, Trueck and Weron [2013a]).¹ In this paper, we provide additional evidence for various weighting schemes and for a new geographical area.

We conduct an out-of-sample evaluation of the forecasting performance of the various models on a day-ahead basis. We consider forecasts of French wholesale prices over 2012 and assess the performance of the models or of their combinations for each trading hour and each season. We calculate several forecast evaluation criteria and test for a significant difference between them with the Diebold and Mariano [1995] test, the Giacomini and White [2006] test and the test for encompassing of Harvey, Leybourne and Newbold [1997]. The evaluation is extensive, since we consider 2880 individual models and 2304 model combinations. The results show that the double temporal segmentation improves the results, except in summer. The gain is stronger if we focus on linear models and on specifications without fundamental variables. Among the non-linear specifications, the three-state Markov-switching model leads to more accurate results. However, the performance of individual models remains unstable across hours and seasons. Hence,

¹See also de Menezes, Bunn and Taylor [2000] for a survey and an application to electricity load.

pooling the results provides more reliable forecasts.

The paper proceeds as follows. In the first section, we present the French electricity markets with a focus on the wholesale one. The second section describes the data. The third section presents the estimation and forecast design. The fourth section provides an assessment of the forecasting performance of the various specifications. The last section concludes.

1 The market design

1.1 French energy situation

The energy sector in France has changed dramatically since the first oil shock in 1973 and the government's decision in March 1974 to increase the nuclear power capacity rapidly. Following this decision, the use of coal and oil declined markedly, substituted mainly by the growth in electricity and natural gas consumption. Today, France is the eighth largest producer of electricity in the world and the second in term of nuclear production.

Contrary to several European countries which benefit from an abundance of raw materials, France is poor in energy resources. To limit imports of oil and natural gas and therefore dependence, French electricity consumption has grown twice as fast as the overall energy consumption between 1973 and 2012. In this period, electricity consumption almost tripled despite a minor decline in the years 2009 and 2011 (due to the economic crisis and weather conditions). Without fossil fuel reserves, the nuclear power program led to a substitution in electricity generation. The increase in nuclear power generation (15 TWh in 1973 to 405 TWh in 2012) is therefore accompanied by a reduction of the use of classic thermal units. Today, France represents 17% of the nuclear power generation in the world with 58 reactors.² EDF, the former vertical integrated public monopoly, controls 80% of the country's power generating capacity including all atomic reactors, two-thirds of thermal capacity, 81% of hydroelectric capacity and about a third of renewable output.

The French energy mix points to a large consumption of oil (used especially in transport) and nuclear power. In 2012, 75% of electricity was generated by nuclear power plants and 12% by hydro plants which are used for base load in the merit order. The costs of electricity production are among the lowest in Europe. The residential sector is the country's largest electricity user, currently representing about two thirds of the total. Next in importance is the industrial sector.³

²The French nuclear electricity production amounts to 63 130 MW in 2012; the rest is 6 160 MW for

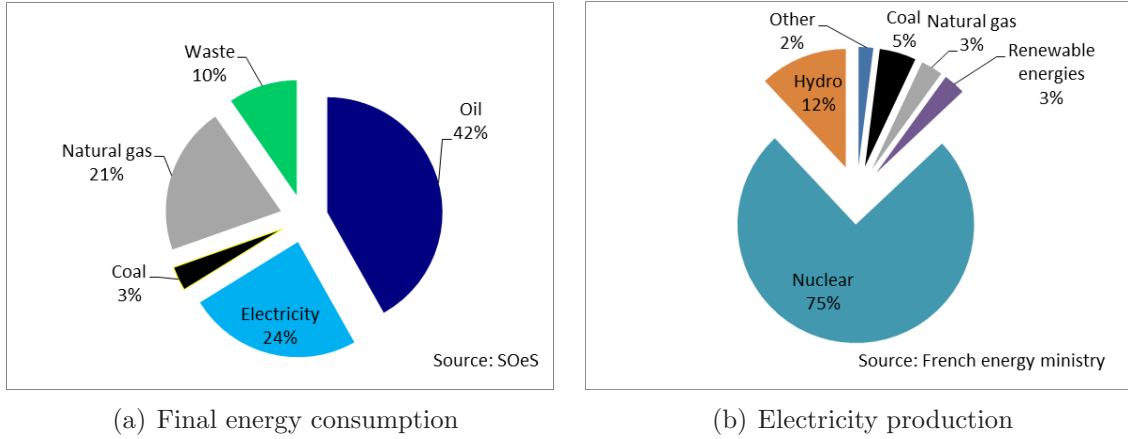


Figure 1: French energy situation in 2012

While French consumption is the second largest in Europe after Germany (in terms of number of consumers and in output), its power trading market is not as dynamic due to the limited number of market players. It can also be particularly constrained in winter, since French demand is very sensitive to temperature. In 2012, for example, the level of electricity consumption reached more than 102 000 MW on February 8, the highest recorded level in France. As mentioned previously, a 1°C drop in the temperature is estimated to increase the 7 pm electricity demand peak of up to around 2 300 MW. When temperatures are very low, France may not have enough installed capacity. During the limited number of days in the year that the country does not have enough capacity, it has to find another way to secure its extreme peak power needs in order to guarantee security of supply. That usually comes from importing power from the surrounding countries. One key partner is Germany which has a better reserve margin in winter due to its combined heat and power plants that are running, so that more capacity is provided during that season.

1.2 Organization of the French electricity market

Electricity is traded either on the wholesale market or on the retail market. The empirical analysis will deal with the wholesale exchanges made on the EPEX spot. It currently represents 17% of the electricity produced and sold in 2012 in France but this share should increase with the ongoing reorganization of electricity markets in Europe.⁴

Retail electricity market

coal plants, 5 009 MW for natural gas, 6 945 MW for fuel and peak load, and 24 301 MW for hydro.

³In 2011, the demand was divided up as follows: residential and services 67.3%, industries 28.1%, transport 2.8% and agriculture 1.7% (source: SOeS).

⁴See also Vassilopoulos [2010] for a recent discussion on the French wholesale market.

Retail transactions involve the final sale of power to end-use customers such as residential houses, commercial properties and manufacturing firms. In 2012, the French retail market represented 25.8 million sites which accounted for 423 TWh of annual consumption.

Clients have the choice between two contracts, either under regulated tariffs (offered mostly by the incumbent supplier) or at market prices (offered by all suppliers). The regulated tariffs are set by the Ministries of the Economy and Energy, on a proposal from the CRE (the French regulatory authority on energy). They differ according to the type of customer (individual, small business, large companies). In December 2012, the share of contracts at market prices represented 31% of the total with a share of 66% for large nonresidential sites.

The NOME Law of 2010 (Nouvelle Organisation du Marché de l'Electricité - The New Organization of the electricity market) provides that regulated tariffs will disappear in 2016 for small businesses and will only exist for residential customers. 21.8% of electricity used by businesses is purchased under a contract at market prices with an alternative supplier. Therefore, this share will evolve significantly and the transactions on the wholesale market will also increase in volume and number.

Wholesale market

On the wholesale market, electricity is traded (bought and sold) before being delivered to the network to end users (individuals or companies).

The participants are the electricity suppliers most of whom (but not all) own generating plants (EDF, GDF Suez, SNET, Poweo, Direct Energy, Enercoop...), brokers and traders. In 2012, 565 TWh were injected into grids: 449 TWh from generating units, 31 TWh from VPP generation (virtual power plant), 56 TWh from ARENH (Regulated Access to Incumbent Nuclear Electricity), and 28 TWh from imports.

- EDF offers access to 5400 MW of its domestic capacity via quarterly auctions (virtual power plants). This capacity can be obtained by producers, suppliers and traders already active in the domestic market or wishing to enter through the quarterly auction mechanism. EDF sells this production capacity in the form of power purchase contracts specifying both capacity and prices.
- The NOME law established the ARENH system (Regulated Access to Incumbent Nuclear Electricity). This system entitles suppliers to purchase electricity from EDF at a regulated price, in volumes determined by the French energy regulator. One issue with the deregulation of the electricity market in France lies in the fact that

the incumbent supplier EDF could offer tariffs related to its fleet of inexpensive generation (dependent on 90% of the cost of nuclear and hydro). It is not the case for the alternative suppliers who make offers based on the European wholesale prices. To reduce this difference, the solution chosen is to allow alternative suppliers to repurchase a share of the EDF nuclear generation at the ARENH rate.⁵

- France also imports electricity when the cost of electricity is more attractive at a given time. The electricity interconnections enable importing less expensive electricity at certain times of the day than that produced by the national production units. This is the case in France at peak times (especially in the evening in winter), when thermal power plants are used. It is usually more profitable to import electricity at these times and to export when domestic demand decreases.

These 565 TWh injected on the grids in 2012 were withdrawn from grids through: 463 TWh for end user consumption, 31 TWh grid losses purchased on markets and 62 TWh for exports.

An important part of the electricity wholesale market activity takes place on power exchange markets: France Epex Spot for spot products, based in Paris and EEX Power Derivatives for future products in Leipzig.⁶ Epex Spot was created in 2008 through the merger of the power spot activities of the French energy exchanges Powernext and the German one EEX. The Epex Spot trades contracts for the physical delivery of electricity in Austria, Germany, France and Switzerland.

The reference price for the spot trade is the price of the day-ahead product on Epex Spot. Epex spot prices are negotiated the day before the delivery by an auction mechanism. In the morning, buyers and sellers (consumers and power generators) submit their orders (price/quantity combination) for each hour (or block of hours) of the forthcoming day. The market is closed at 12:00 noon for France and Epex Spot computes the demand and supply curves and the equilibrium price and volume for each hour of the forthcoming day. The results are published as soon as they are available from 12:40 pm for France.⁷

⁵In April 2011, the government set the ARENH at 40 euros per MWh between July and December 2011 and 42 euros per MWh after January 2012.

⁶The rest of exchanges takes place over-the-counter (OTC), through direct transactions or through intermediaries (brokers and trading platforms). The total volume of OTC transactions is not public.

⁷Epex Spot also operates intraday power markets for Germany, France and Austria. Intra-day contracts on this market are intended to respond to unpredictable temporary physical requirements during the day (contracts can be traded up to 45 minutes before physical fulfillment). 2012 trading volumes on Epex Spot accounted for 321 TWh in the day-ahead auction against 18 TWh on the intraday markets.

2 Data

We focus on electricity spot prices in euros per megawatt hour (€/MWh) for France from the European Exchange market (EPEX).

We consider several price drivers as explanatory variables. The variables are published at a daily or hourly frequency and must be available before the release of spot prices for day t . Widely accepted market fundamentals include the following indicators of supply, demand and risk:

- Demand forecast (in MW): The 24 hourly day-ahead forecasts for continental France are released by RTE (the French transmission grid operator) at 0:00 in $t-1$.
- Capacity margin (in MW): The forecasted margin is published by RTE at 8 p.m. in $t-1$. It represents ‘the volume of capacity available for RTE over and above the operating schedule capable of being used to cope with generation or consumption contingencies’. We take the margin of the morning peak available all over the year.
- Volatility of spot prices (in €/MWh): We use the coefficient of variation (standard deviation/mean) of the hourly prices over the last five days (weekly rolling window).
- Past values of spot prices (in €/MWh): The price of the day before (lag-1) and of the same day the week before (lag-5) and the weighted average price over the previous day are particularly relevant.
- Gas price (in €/MWh): We consider the Dutch day-ahead price of Natural Gas (TTF) which is a reference price in Europe.⁸ To avoid an endogeneity problem, we use lag-1 price.
- Forecasted balance of exchange programs with Germany (in MW): RTE provides figures for the total volume of exports from France and imports into France for each hour as it is known each day at the end of the afternoon, for the following day.

The power prices, demand, volumes of exchange have an hourly frequency, while the capacity margin and the gas price are available at a daily frequency. With the exception of cross-border trade, these variables have already been considered as price drivers in the literature (see for instance, Karakatsani and Bunn [2008] and Bordignon et al. [2013]).⁹

⁸The Title Transfer Facility (TTF) is a virtual trading point for natural gas in the Netherlands, set up in 2003. It became the biggest gas trading platform in Europe.

⁹Many authors only include the (forecasted) demand in addition to lagged prices in the forecasting equation. See for instance Nogales, Contreras and Conejo [2002], Conejo, Contreras, Espinola and Plazas [2005], Misiorek et al. [2006], Weron and Misiorek [2008].

In this paper, we also take into account the cross-border exchanges in electricity. As described in the previous section, the French electricity transmission network is interconnected with neighboring countries. The growth of cross-border trade leads us to account for these exchanges in our models.¹⁰ We focus on the exchanges with the historical partner Germany. We omit temperature and other climatic variables, already taken into account in the demand forecast by RTE.

In contrast with many papers, we take into account the particular timing of the release of the spot prices, as depicted in Figure 2. The market is closed at 12:00 noon and the spot prices for the 24 hours of the next day are published as from 12:40 pm. Following the market practice, we only use information available up to noon before the market clearing. This means that all variables involved in the forecasting must be available before noon one day before the reference day. This is the case of the forecasted demand released at 0:00 the day before the reference day and the lagged gas price. However, the forecasted capacity margin and exchange volumes with Germany for the next day are published by RTE only at the end of the afternoon. For this reason, we use the lagged values of these two variables in our models. By doing so, we ensure that our results are relevant for practitioners.¹¹

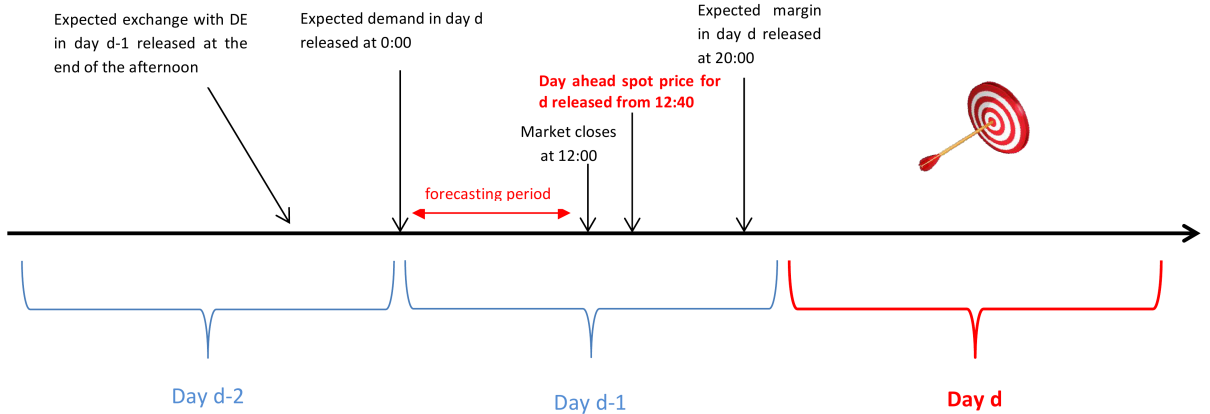


Figure 2: Daily updating scheme of data releases in Epex Spot

The whole sample spans from January 2, 2008 to December 31, 2012, yielding 1263

¹⁰In November 2010, the Central Western Europe Market Coupling (CWE) was launched. CWE is a cooperation of Transmission System Operators and power exchanges covering the Belgian, Dutch, French and German electricity markets.

¹¹Given the delay in the publication of the explanatory variables and the early publication of spot prices, the forecasts performed here are useful over a short window, from 0:00 to noon. Market participants use this window to optimally elaborate their bidding strategies.

daily observations for each trading hour. We consider several ranges of data to forecast some prices in each season of 2012 (see below). Note that it is particularly challenging to forecast prices in 2012. A historical peak was reached in February 2012 during an exceptionally cold wave, while the lowest consumption level observed for five years occurred in August 2012. The prices, demand and margin are taken in logarithm given their high volatility. Like Karakatsani and Bunn [2008] and Bordinon et al. [2013], we remove weekends and holidays from the data and consider working days only. We also remove two days in 2012 (12/26/2012 and 01/02/2012), where prices are negative at 4 am and 5 am.¹² We apply several unit root tests to each variable in logarithm. All series are found stationary at the usual significance levels, except the forecasted demand and the gas price (the results are available upon request). In the following, we will consider the latter ones in difference.

3 Modeling and forecasting procedures

3.1 Temporal segmentation

Due to the strong seasonality present both in the intra-day and inter-day dynamics of electricity prices, especially in France given the high share of electricity in residential heating, we implement a double temporal segmentation of the data.

As already mentioned in the literature, prices display a distinct pattern depending on the hour of the day. This variation is mainly a result of the evolution of demand during the day. According to RTE, the intra-day profile of French consumption is characterized by four phases: the night trough, which is the minimum consumption of the day, the morning peak, the afternoon trough and the evening peak. The maximum consumption is reached at the morning peak at 13:00 in summer and at the evening peak around 19:00 in winter (heavy use of electrical appliances, public transport, lighting, heating ...).

As done in Garcia-Martos, Rodriguez and Sanchez [2011], Figure 3(a) shows the box plot of prices between 2008 and 2012 at each trading hour.¹³ In line with the demand profile, peaks in level occur around 1 pm and 7 pm, while prices are lower from 2 am to 5 am and 3 pm to 6 pm. Prices display a larger volatility between 3 am and 5 am and between 6 pm and 8 pm. To allow a variation of the coefficients of the forecasting

¹²Negative price can occur in situations of overcapacity. For technical or economic reasons, some types of generators, such as nuclear, hydroelectric, or wind power plants, cannot or prefer not to decrease their production for short periods of time of lower demand. We decided to discard these observations to be able to use the logarithmic transformation.

¹³293 extreme values out of the 24792 observed prices have been removed for the readability of the graph.

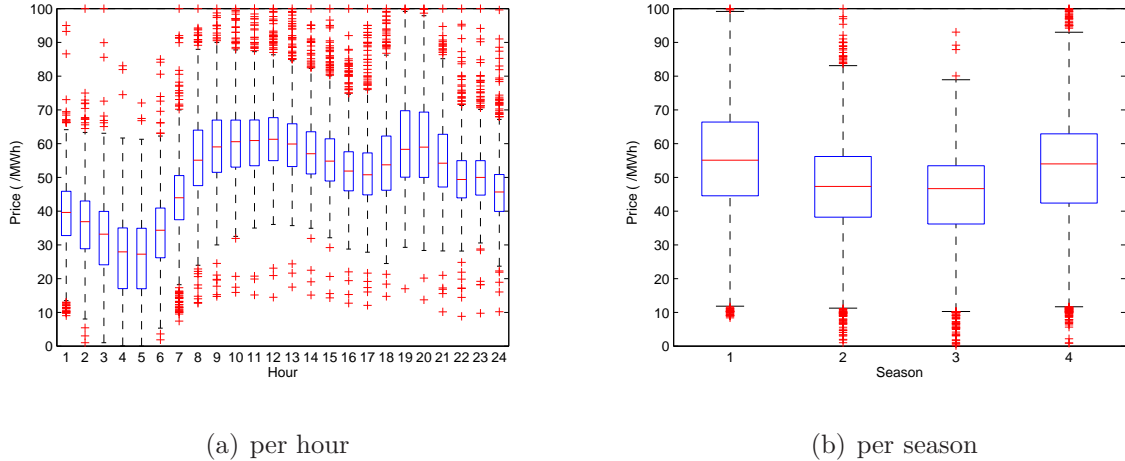


Figure 3: Boxplot of spot prices (2008-2012)

model during the day, we model separately each hour of the day, as often done in the literature (see, for recent examples in forecasting, Crespo Cuaresma et al. [2004], Misiorek et al. [2006], Weron and Misiorek [2008], Karakatsani and Bunn [2008], Bordignon et al. [2013]).¹⁴

Spot prices also display a strong seasonal pattern during the year. Their level and volatility are much higher in fall and winter, as shown by the box plot of power prices in each season (Figure 3(b)). Indeed, consumption is seasonal and the price per MWh varies seasonally: summer demand is lower and the MWh is cheaper. It is produced from nuclear and hydropower. In winter, the demand is higher and the MWh is more expensive. Therefore, the last power plant "called" has to be able to quickly meet demand. These plants can meet this constraint by using fossil fuels at a higher cost. For example in France, about 30 GW are needed in August during the day while up to 102 GW have been called in February at 7.00 pm in a cold winter. To also allow a variation of the coefficients during the year, we model each season independently.

As mentioned above, the hourly segmentation is now common in literature. However, to the best of our knowledge, seasonal segmentation has not been applied for power price modeling. Other more common approaches to deal with the long-run seasonality are described in Nowotarski et al. [2013b]: dummies, sinusoidal functions, wavelets. The usual approach employing monthly dummy variables (allowing generally a variation of the intercept only) or sinusoidal functions is less flexible and/or less parsimonious. The

¹⁴Alternative approaches include the use of dummies for the trading periods in a unique equation (Popova [2004]). However this method is far less flexible or parsimonious. Alternatively, multivariate approaches are employed by Huisman, Huurman and Mahieu [2007] with a panel framework or by Panagiotelis and Smith [2008] taking into account the hourly heterogeneity in a VAR model.

use of wavelet decomposition for forecasting requires arbitrary choices about the wavelet family, the vanishing moment in many of them, the number of octaves. Moreover, the use of fundamental variables is less straightforward with the wavelet decomposition. In a close context, Taylor [2010] uses seasonal ARMA methods or smoothing approaches to capture the intraday, intraweek and intrayear seasonality in load data. In this paper, the intraweek component is omitted, since we have removed the weekends and holidays from the data.

The variation of the coefficients of a linear regression of the price on the drivers depicted in Figure 4 across the two dimensions (trading hours and seasons) supports this double segmentation. Figure 4 shows the estimated coefficients of a linear model relating the spot prices to the aforementioned fundamental variables. This regression is conducted either on 2012 data or on seasonal data (i.e. on data-sets consisting only of observations of the same season). This will be explained further in the next section. We note a significant variation of the coefficients across hours (variation of the coefficients along each curve) and seasons (variation between the curves).

3.2 Linear versus non linear models

We consider linear and nonlinear specifications of the relationship between spot prices and market fundamentals. A summary of all specifications with their code is given in Table 1a. In the following, $p_t^{(h)}$ denotes the log-price and $x_t^{(h)}$ the fundamentals for hour h , $h = 1, \dots, 24$ of day t .¹⁵

The linear specifications include a basic autoregressive model, an exponential model and autoregressive regressions with exogenous variables. The first two models do not include fundamental regressors and will be useful to assess the contribution of exogenous day-ahead information to forecasting. The autoregressive (AR) model is simply given by:

$$p_t^{(h)} = \alpha^{(h)} + \sum_{i=1}^p \phi_i^{(h)} p_{t-i}^{(h)} + \varepsilon_t^{(h)}, \quad t = 1, \dots, T \quad (1)$$

where $\alpha^{(h)}$, $\phi_i^{(h)}$, $i = 1, \dots, p$ and $h = 1, \dots, 24$, are constant coefficients estimated by least squares and $\varepsilon_t^{(h)}$ is the error term. The order p is selected with the Akaike Information Criterion (AIC) for a maximum lag length fixed to 5. In this framework, the optimal

¹⁵We have assessed the existence of a break in the forecasting equations due to the implementation of market coupling in November 2010 (see footnote 9). To this aim, we have performed rolling regressions from January 2008 to December 2012. The results do not show any major change in the relationship between spot prices and the price drivers in 2011 and 2012. Therefore, we do not incorporate a structural change in the forecasting equations in November 2010.

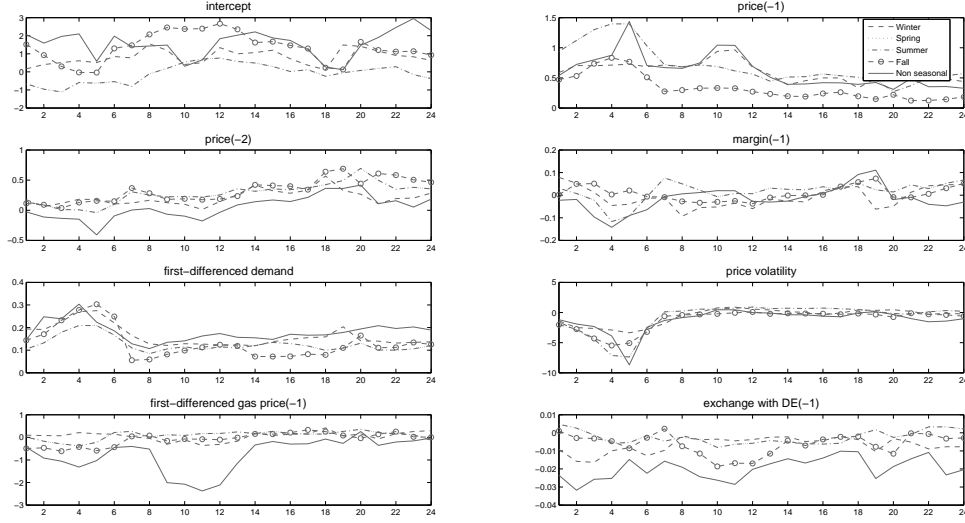


Figure 4: Coefficients of the drivers in linear models

predictor $\hat{p}_T(1) = E(p_{T+1}|p_T, p_{T-1}, \dots)$ of the price in $T+1$ based on the information available up to time T is given by:

$$\hat{p}_T^{(h)}(1) = \alpha^{(h)} + \sum_{i=1}^p \phi_i^{(h)} p_{T+1-i}^{(h)} \quad (2)$$

Forecasts from equation (1) perform well as documented in the literature.

To forecast stationary data, another usual solution relies on exponential smoothing technique. This approach consists in assigning exponentially decreasing weights over time to the past observations:

$$\begin{aligned} s_t^{(h)} &= \alpha^{(h)} p_t^{(h)} + (1 - \alpha^{(h)}) s_{t-1}^{(h)} \quad \text{for } t > 0 \\ \text{with } s_0^{(h)} &= p_0^{(h)} \end{aligned} \quad (3)$$

where $s_t^{(h)}$ is the smoothed value and $\alpha^{(h)} \in [0, 1]$ is the smoothing factor. The parameter $\alpha^{(h)}$ is estimated by minimizing the mean square error (with $e_t^{(h)} = p_t^{(h)} - s_{t-1}^{(h)}$). Then, the optimal predictor $\hat{p}_T(1)$ can be written as follows:

$$\hat{p}_T^{(h)}(1) = s_T^{(h)} \quad (4)$$

Third, the autoregressive model with exogenous variables (AR-X) is specified as:

$$p_t^{(h)} = \alpha^{(h)} + \beta^{(h)'} x_t^{(h)} + \varepsilon_t^{(h)}, \quad t = 1, \dots, T \quad (5)$$

where $\beta^{(h)}$ is a $k \times 1$ vector of constant parameters, $x_t^{(h)} = (p_{t-1}^{(h)}, \dots, p_{t-p}^{(h)}, z_t^{(h)})$ is the $k \times 1$ vector of regressors consisting of autoregressive terms and fundamentals $z_t^{(h)}$ in day t and

$\varepsilon_t^{(h)}$ is the error term. The model is estimated by least squares. In this case, the optimal predictor $\hat{p}_T(1) = E(p_{T+1}|z_{T+1}, p_T, p_{T-1}, \dots)$ is given by:

$$\hat{p}_T^{(h)}(1) = \alpha^{(h)} + \beta^{(h)'} x_{T+1}^{(h)} \quad (6)$$

We also consider nonlinear models in order to capture infrequent and extreme observations typical of electricity prices (spikes and drops) and the possible change in the relationship between spot prices and their fundamentals during these abnormal periods.

First, we consider Markov-switching models popularized by Hamilton [1989]. In these models, the parameters of equation (5) depend on an unobservable discrete variable S_t :

$$p_t^{(h)} = \alpha^{(h)}(S_t^{(h)}) + \beta^{(h)'}(S_t^{(h)})x_t^{(h)} + \varepsilon_t^{(h)}(S_t^{(h)}), \quad t = 1, \dots, T \quad (7)$$

where $\varepsilon_t^{(h)}(S_t^{(h)}) \rightarrow NID(0, \sigma^2(S_t^{(h)}))$. The variable $S_t^{(h)} = 1, 2, \dots, M$ represents the state that the process is in at time t . This variable is assumed to follow a first-order Markov chain defined by the following transition probabilities:

$$q_{ij}^{(h)} = P(S_t^{(h)} = j | S_{t-1}^{(h)} = i) \quad \forall t \quad (8)$$

where $\sum_{j=1}^M q_{ij}^{(h)} = 1, \forall i, j = 1, 2, \dots, M$. The forecast of the chain $S_t^{(h)}$ at horizon 1 is given by $\hat{\xi}_{T+1|T}^{(h)} = P^{(h)} \hat{\xi}_{T|T}^{(h)}$, where $P^{(h)}$ is the transition matrix of the Markov chain and $\hat{\xi}_{T|T}^{(h)}$ the vector of smoothed probabilities in T . The optimal predictor $\hat{p}_T(1)$ is derived by weighting each estimated regime with the predicted probabilities of the states in $T + 1$, $\hat{\xi}_{T+1|T}^{(h)}$.

Following Diebold, Lee and Weinbach [1994], we extend this specification with time-varying transition probabilities. In this case, the transition probabilities are not time invariant. Instead, they depend on an exogenous variable $q_t^{(h)}$:

$$P(S_t^{(h)} = i | S_{t-1}^{(h)} = i, q_t) = \Phi(a_i^{(h)} + b_i^{(h)} q_t^{(h)}) \quad i = 1, 2 \quad (9)$$

where Φ is the cumulative distribution function of the standard normal distribution, $a_i^{(h)}$ and $b_i^{(h)}$ are unknown parameters for regime $S_t = i$ and $q_t^{(h)}$ is an exogenous variable (possibly included in $x_t^{(h)}$). The forecast of the chain $S_t^{(h)}$ at horizon 1 is given by $\hat{\xi}_{T+1|T}^{(h)} = P_{T+1}^{(h)} \hat{\xi}_{T|T}^{(h)}$, where $P_{T+1}^{(h)}$ is the transition matrix of the Markov chain in $T + 1$.

In the following, we only consider the case of two and three regimes ($M = 2, 3$). For $M = 2$, we consider three versions of the MS-ARX model with fixed or time-varying transition probabilities. The first ones allow a variation of the intercept and the variance only, while the second ones also allow a variation of the autoregressive coefficients. In

the third ones, all the parameters depend on S_t . For $M = 3$, we only consider the first two constrained specifications with fixed transition probabilities and do not consider the general form given the proliferation of parameters when the number of regimes increases.

The Markov-switching models are estimated via a numerical maximization of the maximum likelihood (we use the function *fminunc* in Matlab). The likelihood is derived in the Hamilton filter (1989). For the specifications with time-varying probabilities, the fixed transition probabilities are replaced by time-varying probabilities in the first step of the filter and the rest of the estimation procedure is similar. A large set of initial conditions is used to avoid any dependency of the results on the initialization. When a convergence problem was encountered in the maximization, the estimation of the model was re-run with a new set of starting values.

A second class of nonlinear models considered in this paper is the Smooth Threshold AutoRegressive models (Granger and Terasvirta [1993]) augmented with exogenous variables. The variable $p_t^{(h)}$ follows a STAR-X model if:

$$p_t^{(h)} = \alpha_0^{(h)} + \beta_0^{(h)'} x_t^{(h)} + \left(\alpha_1^{(h)} + \beta_1^{(h)'} x_t^{(h)} \right) G(q_t^{(h)}; \gamma^{(h)}, c^{(h)}) + \varepsilon_t^{(h)} \quad (10)$$

for $t = 1, \dots, T$. The residual $\varepsilon_t^{(h)}$ is assumed to be *i.i.d.* $(0, \sigma_\varepsilon^2)$. In our case, the mechanism of transition is defined as a logistic function of order 1:

$$G(q_t^{(h)}; \gamma^{(h)}, c^{(h)}) = [1 + \exp(-\gamma^{(h)}(q_t^{(h)} - c^{(h)}))]^{-1}, \gamma^{(h)} > 0 \quad (11)$$

where $c^{(h)}$ denotes a location parameter and $\gamma^{(h)}$ determines the slope of the transition function *i.e.* the speed of the transition from one regime to the other one. The logistic function is a continuous function with S shape bounded between 0 to 1. It represents a smooth transition from an inferior regime (coefficient $\beta_0^{(h)}$) to a superior one (coefficient $\beta_0^{(h)} + \beta_1^{(h)}$) with an infinite number of intermediate regimes in between. The estimation of the parameters in a STR model consists in applying non-linear least squares. The optimal predictor $\hat{p}_T(1)$ is given by:

$$\hat{p}_T^{(h)}(1) = \alpha_0^{(h)} + \beta_0^{(h)'} x_{T+1}^{(h)} + \left(\alpha_1^{(h)} + \beta_1^{(h)'} x_{T+1}^{(h)} \right) G(q_{T+1}^{(h)}; \gamma^{(h)}, c^{(h)}) \quad (12)$$

Again, we consider three versions of this model: the first one with a variation of the intercept only, the second one with a variation of the intercept and the AR coefficients and the last one with a variation of the coefficients of all regressors.

With respect to the regressors' selection, we consider a linear model with explanatory variables selected with a stepwise procedure among the set of market fundamentals described in the previous section. To get a fair comparison and make the convergence of the

estimation algorithms easier for the nonlinear models, we also consider all the specifications with the same reduced structure: two autoregressive terms - the weighted average price of the preceding day¹⁶ and the price at the same hour two days before - and demand in first-difference. At this level, we choose the most significant regressors over all data ranges and trading hours. The forecasted demand is found to be more influential in our models since it refers to the day of interest, while the other regressors are lagged because of their delay of publication.

As suggested by Chen and Bunn [2010], the transitional variable can vary according to the season and the trading hour. To take this result into account, we select the transitional variable of the threshold and MS models with time-varying probabilities for each hour, each season and each treatment of seasonality. The variable is chosen among the whole set of explanatory variables (lagged price, lagged margin, first-difference forecast demand, first-difference of lagged gas price, forecasted power exchange with Germany, past volatility of spot prices). For each regime-switching model with a time-varying transition, we conduct a linearity test with each potential variable and select the one which provides the strongest rejection of the null of linearity.

3.3 Individual versus pooled forecasts

In the previous subsection, we have proposed several linear or nonlinear candidate methods for forecasting electricity prices. However, an alternative solution used by Bordignon et al. [2013], Raviv et al. [2012] and Nowotarski et al. [2013a] in the context of electricity market consists in combining forecasts. The objective highlighted by Bordignon et al. [2013] is to obtain accurate predictions when a wide variety of models exists and no predominant one can be selected.

We consider several methodologies for combining forecasts. The easiest way to obtain the combined prediction \hat{p}_t^c is to consider the mean of the individual predictions $\hat{p}_t^{(k)}$ of the k individual models considered (for $k = 1, \dots, K$):

$$\hat{p}_t^c = \frac{1}{K} \sum_{k=1}^K \hat{p}_t^{(k)} \quad (13)$$

An alternative approach is to weight the predictions based on a selected criterion:

$$\hat{p}_t^c = \sum_{k=1}^K \omega_{t,k} \hat{p}_t^{(k)} \quad (14)$$

¹⁶The lagged weighted average price is more significant than the price at the same hour the day before. The weighted average price is also smoother and it makes the estimation of nonlinear models easier.

with $\omega_{t,k}$, the weight given at time t to the predictor k , $k = 1, \dots, K$. Like Bordignon et al. [2013], we adopt the usual solution proposed by Bates and Granger [1969] which consists in assigning weights depending on the inverse of means square prediction errors of the l most recent observations:

$$\omega_{t,k} = \frac{(\sum_{\tau=t-l}^{t-1} e_{\tau,k}^2)^{-1}}{\sum_{j=1}^K (\sum_{\tau=t-l}^{t-1} e_{\tau,j}^2)^{-1}} \quad (15)$$

where $e_{\tau,k} = p_{\tau} - \hat{p}_{\tau}^{(k)}$, $\omega_{1,k} = 1/K$. Accordingly, $0 \leq \omega_{t,k} \leq 1$, $\sum_{k=1}^K \omega_{t,k} = 1$. The adaptive weights are larger if the individual models yield more accurate forecasts. The third solution is the Aggregated Forecast Through Exponential Re-weighting, AFTER (Yang [2004]; Zou and Yang [2004]). The weights are obtained recursively according to:

$$\omega_{t,k} = \frac{\hat{v}_{t-1,k}^{-1/2} \exp[-\frac{e_{t-1,k}^2}{2\hat{v}_{t-1,k}}] \omega_{t-1,k}}{\sum_{i=1}^K \hat{v}_{t-1,i}^{-1/2} \exp[-\frac{e_{t-1,i}^2}{2\hat{v}_{t-1,i}}] \omega_{t-1,i}} \quad (16)$$

where $e_{t-1,k} = p_{t-1} - \hat{p}_{t-1}^{(k)}$, $\omega_{1,k} = 1/K$. Accordingly, $0 \leq \omega_{t,k} \leq 1$, $\sum_{k=1}^K \omega_{t,k} = 1$ and $\hat{v}_{t-1,k} = \frac{1}{t-1} \sum_{\tau=1}^{t-1} e_{\tau,k}^2$ is the estimation of the prediction variance.

A last solution adopted in this paper relies on the predictive likelihood model average (PLMA) proposed by Kapetanios, Labhard and Price [2006]. The weights are assigned depending on AIC information criteria built on a sample of the l previous forecast errors:

$$\omega_{t,k} = \frac{\exp(-1/2\Psi_{t,k})}{\sum_{j=1}^K \exp(-1/2\Psi_{t,j})} \quad (17)$$

for $k = 1, \dots, K$ where $\Psi_{t,k} = AIC_{t,k} - \min_j AIC_{t,j}$, $0 \leq \omega_{t,k} \leq 1$ and $\sum_{k=1}^K \omega_{t,k} = 1$. The AIC criterion is not calculated on the basis of the estimated log likelihood (which is derived in-sample from observed data). Instead, this criterion is based on a predictive likelihood derived from out-of-sample forecast errors. This construction has the advantage to combine a measure of forecast accuracy and a penalty term for model complexity. A summary of all possible combinations with their code is given in Table 1b.

4 Day-ahead forecasting

4.1 Empirical design

We compare the various approaches described above for modeling French spot prices through their forecasting performance.

The evaluation is conducted in real conditions, i.e., out of the estimation period. We remove the last observations of our sample. We estimate the 15 specifications over a

rolling window and run a day-ahead forecast of the variable $p_t^{(h)}$ for each trading hour $h = 1, \dots, 24$. We repeat these calculations up to the last observation of the out-of-sample period. The out-of-sample period consists of the last 35 observations of each season in 2012. The estimation set includes the preceding weekdays either of the entire past year, or of the same season (hereafter referred to as *non seasonal* and *seasonal* ranges respectively). To have a comparable number of observations (around 255 weekdays) across simulations, the seasonal range consists of four seasons. The alternative data-sets used to forecast the last observations of each season are described in Figure 5.

The choice of the specifications is done in real-time. In the case of the AR process, the lag length is chosen with the AIC criterion at each iteration of the out-of-sample exercise. Similarly, the choice of the regressors in the linear model with the stepwise procedure is done at each iteration on the basis of the available information at the time of the forecast. The threshold variable in the STR models and the variable entering the transition probabilities in the MS models are selected on the sample excluding the forecasting period at the beginning of the iterations.¹⁷ As suggested by Chen and Bunn [2010], the optimal transitional variables vary a lot depending on the time period, the season as well as the specification. The forecasted balance of power exchange, the demand forecast and the past volatility of spot prices are chosen more often.¹⁸ At each period of the evaluation period, we use a large set of initial values to avoid any dependency of the results on the initial values for the estimation of the MS and STR models¹⁹.

To mimic the practice of forecasters, we use an automatic insanity filter (Stock and Watson [1999]). This filter removes unrealistic forecasts and replaces them with some reasonable values. Given the existence of spikes in the price series, this filter must be used with parsimony. We only discard values found outside a wide interval given by the historical range of prices (over all hours) enlarged by plus/minus the average standard deviation of hourly prices. This rule requires discarding 0.45% of forecasts only. These insane forecasts are replaced by the last observed price at the same hour (naive forecast).

The out-of-sample evaluation is conducted for each model or combination of models in

¹⁷The linearity tests are not performed at each recursion because it is too time consuming. However, we exclude the 35 forecasted observations to have a real time experience.

¹⁸For sake of parsimony, the results are not reported in the paper but they are available upon request.

¹⁹For MS models, the initial vector of parameters was reinitialized every five observations of the evaluation period. For STR models, we apply a dense grid search to get good starting values for (γ, c) at the first iteration only. We use the values of the threshold variable q_t between the first and the third quartile for the location parameter and 50 positive values (close to zero) for the smoothness parameter. Given the values of (γ, c) , the estimation of the regressors coefficients proceeds by ordinary least squares. Finally, we select the values of the parameters minimizing the residual sum of squares and keep them as starting values for the rest of the iterations.

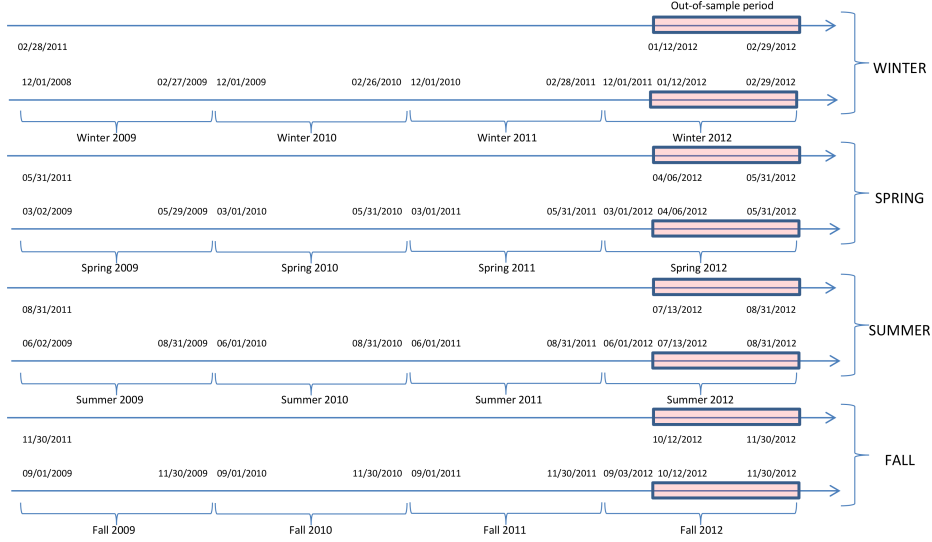


Figure 5: Experimental design - Estimation and forecasting windows

each season and trading hour. We consider a total of 2880 individual models and obtain a set of 35 forecasts with each of them. Pooled forecasts are also obtained by combining these individual forecasts with the four combination methods described in the previous section. In the adaptive rules, the weights are computed from the last $l = 10$ more recent observations. In the case of the last weighting scheme, the AIC criteria are computed with the number of parameters estimated recursively for each hour and each season.²⁰ At this level, we also distinguish combinations from linear and non-linear specifications. A total of 2304 combinations is obtained.

Several comparisons are done with these results: First, we contrast the forecasting accuracy of regime switching models (MS and STR) to linear specifications (AR, exponential and AR-X). Second, the benefit of combining forecasts to achieve more accurate predictions is studied by comparing the individual forecasts of the 15 models versus pooled forecasts. To this end, we compare the forecasts of the models estimated over an entire year or on data belonging to the same season, as described above. Finally, we assess the gain due to the double temporal segmentation of the data.

The forecast accuracy of each model or combination of models is measured with the usual root mean square error (RMSE), mean absolute error (MAE) and mean absolute percent error (MAPE). Using these criteria, we provide a ranking of the models according to their RMSE/MAE and report an average rank over the 24 trading hours for each season and the full year. This evaluation is qualitative since it relies on averages of rankings. To take into account quantitatively the forecast accuracy of each model, we also compute

²⁰The number of parameters in the AR and LSTEP models vary at each iteration and it is fixed in the other specifications.

MAPE for each model over all hours for each season and the whole year and rank the models according to the results (see Nowotarski et al. [2013b] for a similar approach).

To compare the predictive ability of the competing approaches, we implement several tests. First, to assess if there is a significant difference between the MSE or MAE criteria, we perform the test of Diebold and Mariano [1995] with the correction of Harvey et al. [1997] in small samples and the unconditional test of Giacomini and White [2006]. The latter has the advantage of being implementable even if the models are nested (for example, this is the case of linear and regime-switching models) and it captures the effects of estimation uncertainty on the relative forecast performance. We also conduct the encompassing test of Harvey et al. [1997]. If the null hypothesis of this test is rejected, the competing predictor contains useful information not present in the model.

4.2 Out-of-sample results

The evaluation criteria and the rankings of the models are provided in Tables 2-4. The results of the tests appear in Table 5.

To have a detailed view by trading hour and season, Tables 2a-2d provide the evaluation criteria of individual models for peak (H9, H14 and H20) and off-peak (H3) hours. MSE and MAE display higher values at 3 am, which is probably due to the strong volatility of prices between 3 and 5 in the morning (in winter, the worst results are obtained at 9 am). Regarding the results per season, winter is the most difficult period to forecast. The largest MSE and MAEs are found in winter (and in fall between 3 am and 5 am), while the smallest values appear in summer. In winter, the proportion of correct direction is also rather low, whereas specifications give more accurate signals about the sign of price variations in the other seasons.

To compare the performance of individual models, Table 3 gives the number of times over the 24 hours of each season that a given model provides the best forecast in terms of MSE, MAE and MAPE. Table 4 provides the average rank per season for each model and the global rank over the entire year. We provide a qualitative ranking based on the RMSE and MAE criteria and a quantitative ranking relying on the aggregate MAPE of the models. Table 5 reports the results of the tests by class of models. We provide the percentage of times that each approach gives lower MSE or MAE and the percentage of times that these criteria are significantly lower according to the Diebold-Mariano and Giacomini-White tests. The last block of results gives the percentage of rejection of the null hypothesis for the encompassing test of Harvey et al. [1997]. The results of the tests are given at a 10% significance level. Several results are worth commenting on.

The comparison of the models in Table 3 shows that the inclusion of fundamental variables such as the forecasted demand to the specification improves the forecast accuracy of the models: Models using only the past dynamics of prices display better forecasts in only 14% of cases in terms of MSE (in 19% of cases in terms of MAE or MAPE). This result is consistent with the findings by Misiorek et al. [2006] for the California power exchange, Karakatsani and Bunn [2008] and Bordignon et al. [2013] for the UK market and Conejo et al. [2005] for the PJM interconnection. The good performance of LSTEP also shows that the restriction of the set of fundamentals to the forecasted demand can be detrimental, since LSTEP outperforms the other individual models in 23% of cases in terms of MSE (and 18-19% of cases in terms of MAE or MAPE). Therefore, it could be interesting to enlarge the comparison of the various specifications with a larger set of regressors.

With respect to the comparison of linear and non-linear models²¹, Tables 3 and 4 show the good performance of Markov-switching models, especially the three-state model MS3_1 allowing a variation of the intercept and the variance. This model often leads to the best MSE, MAE and MAPE in Table 3 and it ranks first among the individual specifications in Table 4 over all seasons with the three criteria. The good performance of MS3_1 seems due to its ability to catch the spikes and drops (or less extreme values) of power prices. To illustrate this point, Figure 6 depicts the smoothed probabilities of the three regimes obtained on the whole period for a couple of trading hours in each season. By contrast, the three threshold specifications rarely outperform the other models in Table 3 and appear among the worse specifications in Table 4.

For this reason, the results of the tests reported in Table 5 are not clear-cut when considering all non-linear models against the linear ones. The results are more supportive of non-linear models when we focus on MS models, especially the specification MS3_1 with three regimes. This model outperforms the linear specification in terms of MSE or MAE in 54.9 to 79.9% of cases (except in winter on the seasonal range). The results of the Diebold-Mariano and the Giacomini-White tests are consistent with these findings with a more frequent rejection of the null hypothesis in favor of a better forecast accuracy of MS3_1 than the linear specifications. The null hypothesis that the linear models contain all the information provided by the nonlinear specifications is also more frequently rejected with the encompassing test of Harvey et al. [1997] in the particular case of MS3_1.

Despite the good performance of MS3_1, it is clear from the results in Table 3 that no

²¹To get a fair comparison, the specification LSTEP (which contains more regressors than the non-linear models) is excluded in the comparison of the linear and non-linear models.

individual model (including MS3_1) exhibits a superior performance for all trading hours and seasons. This instability is favorable to combinations, as shown by the results in Tables 4-5. Table 4 provides the average rank of each model or combination of models. The results are striking. The pooled forecasts always appear among the first best competitors. This result holds over all seasons, all seasonal treatments and for all the three criteria. The combinations over linear models and over all models rank first. As far as the weighting rule is concerned, the Bates and Granger weights and the simple average are more successful, while more sophisticated rules give larger RMSE. This result is consistent with the previous literature on the subject (see Bordignon et al. [2013] and Nowotarski et al. [2013a] in the context of power price forecasting). The results of the tests in Table 5 also give evidence in favour of combination. The pooled forecasts yield lower RMSE and MAE in 61.9% to 74.6% of cases and the difference is statistically significant in a significant proportion of cases according to the Diebold-Mariano and Giacomini-White tests. The encompassing test rejects the null hypothesis that the individual forecasts contain all the information in the combined forecasts in a large proportion of cases, especially in summer and in fall.

Finally, the results in Table 5 support estimations on data belonging to the same season rather than to an entire year. Indeed, most seasonal models yield better forecasts than non-seasonal ones, except in summer. More accurate forecasts are obtained with a model or a combination of models estimated on a seasonal range in almost 70 to 80% of cases in winter and in spring and in 55% of cases in fall. The difference is significant in a part of the cases according to the Diebold-Mariano and Giacomini-White tests. In spring and in fall, the encompassing test shows a high rejection of the null hypothesis that the seasonal models do not contain information in addition to the information already included in the non-seasonal models. The results are more disappointing in summer with only 30.4% of cases favorable to the seasonal segmentation and a low rejection rate with the Diebold-Mariano and Giacomini-White tests for a lower or equal forecast accuracy of the non-seasonal models.²² However, the magnitude of forecast errors is much smaller in both cases than in the other seasons (see Tables 2).

Interestingly, the gain of the seasonal segmentation is more evident if we restrict the comparison to models without fundamental regressors (the autoregressive and exponential models, AR and EXPO). In this case, the seasonal range leads to lower MSE or MAE

²²This result is probably due to a drop in prices in summer 2009 included in the seasonal dataset. This phenomenon reflects the fall in industrial activity during the 2008-09 recession in France. This is less visible in other seasons where electricity demand is mainly driven by the residential sector (heating effect).

than the non-seasonal specifications in 74% of cases over all seasons (versus 58% when we consider all specifications). The improvement is particularly strong in winter and in fall (95.8% versus 69.1% and 70.8% versus 54.5%). Similarly, the gain of the seasonal range is stronger when we focus on the linear specifications (with or without fundamentals). The linear non-seasonal models are beaten by the seasonal ones in 67.7% of cases. Again, the difference is larger in the colder seasons (89.6% in winter and 64.6% in fall). This means that a part of the seasonality is captured by the fundamental variables and by the change of regimes. Overall, our simple treatment of the long-run seasonality present in electricity prices appears to improve forecast accuracy, especially when we use simple specifications for the price forecasting.

Conclusion

The ongoing reorganization of the French electricity market will lead to an increasing volume of trade on the wholesale market. This being the case, developing forecasting tools of electricity spot prices is a key issue for academics and practitioners alike.

In this paper, we have investigated the forecasting ability of several classes of time series models for day-ahead spot prices in France. We have estimated these models with a double temporal segmentation of the data. Our out-of-sample evaluation indicates that this novel approach improves the forecasting ability of the models. Among the non-linear models, a three-state Markov-switching model designed to capture the sudden and fast-reverting spikes in the price dynamics yields better forecasts. However, the forecast accuracy of combinations of models is more stable across all the hours and seasons.

There are a number of potential extensions to this paper. In particular, it could be interesting to explore how revised weather forecasts available in the morning before the auction could add extra value beyond the midnight release of demand forecasts. Given the importance of the demand as a regime driver and the high thermal sensitivity of the French consumption, this appears as a promising way to improve forecasts.

References

- Bates, J.M. and C.W.J. Granger, The combination of forecasts, *Operational Research*, 1969, *20*, 451–468.
- Bordignon, S., D.W. Bunn, F. Lisi, and F. Nan, Combining day-ahead forecasts for British electricity prices, *Energy Economics*, 2013, *35*, 88–103.

- Chen, D. and D.W. Bunn, Analysis of the nonlinear response of electricity prices to fundamental and strategic factors, *IEEE Trans. Power Syst.*, 2010, *25*, 595–606.
- Conejo, A.J., J. Contreras, R. Espinola, and M.A. Plazas, Forecasting electricity prices for a day-ahead pool-based electric energy market, *International Journal of Forecasting*, 2005, *21*, 435–462.
- Crespo Cuaresma, J., J. Hlouskova, S. Kossmeier, and M. Obersteiner, Forecasting electricity spot-prices using linear univariate time-series models, *Applied Energy*, 2004, *77*, 87–106.
- de Menezes, L.M., D.W. Bunn, and J.W. Taylor, Review of guidelines for the use of combined forecasts, *European Journal of Operational Research*, 2000, *120*, 190–204.
- Diebold, F.X. and R.S. Mariano, Comparing Predictive Accuracy, *Journal of Business & Economic Statistics*, 1995, *13*, 253–263.
- , J.H. Lee, and G.C. Weinbach, Regime switching with time varying transition probabilities, in C.P. Hargreaves, editor, *Nonstationary Time Series Analysis and Cointegration*, Oxford University Press: Oxford (UK), 1994, pp. 283–302.
- Garcia-Martos, C., C. Rodriguez, and M.J. Sanchez, Forecasting electricity prices and their volatilities using Unobserved Components, *Energy Economics*, 2011, *33*, 1227–1239.
- Giacomini, R. and H. White, Tests of conditional predictive ability, *Econometrica*, 2006, *74* (6), 1545–1578.
- Granger, C.W.J. and T. Terasvirta, Modelling nonlinear economic relationships, in Oxford University Press, Oxford 1993.
- Hamilton, J., A new approach to the economic analysis of nonstationary time series and the business cycle, *Econometrica*, 1989, *57*, 357–384.
- Harvey, D., S. Leybourne, and P. Newbold, Testing the equality of prediction mean squared errors, *International Journal of Forecasting*, 1997, *13*, 281–291.
- Huisman, R., C. Huurman, and R. Mahieu, Hourly electricity prices in day-ahead markets, *Energy Economics*, 2007, *29*, 240–248.
- Kapetanios, G., V. Labhard, and S. Price, Forecasting using predictive likelihood model averaging, *Economics Letters*, 2006, *91*, 373–379.

- Karakatsani, N.V. and D.W. Bunn, Forecasting electricity prices: The impact of fundamentals and time-varying coefficients, *International Journal of Forecasting*, 2008, *24*, 764–785.
- Kosater, P. and K. Mosler, Can Markov regime-switching models improve power-price forecasts? Evidence from German daily power prices, *Applied energy*, 2006, *83*, 943–958.
- Liu, H. and J. Shi, Applying ARMA-GARCH approaches to forecasting short-term electricity prices, *Energy Economics*, 2013, *37*, 152–166.
- Misiorek, A., S. Trueck, and R. Weron, Point and interval forecasting of spot electricity prices: Linear vs. non-linear time series models, *Studies in nonlinear dynamics and econometrics*, 2006, *10* (3).
- Newbold, P. and C.W.J. Granger, Experience with forecasting univariate time series and the combination of forecasts (with discussion), *Journal of the Royal Statistical Society A*, 1974, *137*, 131–164.
- Nogales, F.J., J. Contreras, and A.J. Conejo, Forecasting next-day electricity prices by time series models, *IEEE transactions on power systems*, 2002, *17* (2), 342–348.
- Nowotarski, J., E. Raviv, S. Trueck, and R. Weron, *An empirical comparison of alternate schemes for combining electricity spot price forecasts*, HSC Research Reports HSC/13/07, Wroclaw University of Technology 2013.
- , J. Tomczyk, and R. Weron, Robust estimation and forecasting of the long-term seasonal component of electricity spot prices, *Energy Economics*, 2013, *39*, 13–27.
- Panagiotelis, A. and M. Smith, Bayesian density forecasting of intraday electricity prices using multivariate skew t distribution, *International Journal of Forecasting*, 2008, *24*, 710–727.
- Popova, J., *Spatial pattern in modelling electricity prices: Evidence from the PJM market*, West Virginia University 2004.
- Raviv, E., K.E. Bouwman, and D. van Dijk, *Forecasting day-ahead electricity prices: utilizing hourly prices*, Erasmus University Rotterdam, Tinbergen Institute 2012.
- RTE, *La vague de froid de février 2012*, Document de la Direction Economie Prospective et Transparence 2012.

- Stock, J.H. and M.W. Watson, A comparison of linear and nonlinear univariate models for forecasting macroeconomic time series, in R.F. Engle and H. White, editors, *Cointegration, causality and forecasting. A festschrift in honour of Clive W.J. Granger*, Oxford University Press: Oxford (UK), 1999.
- Taylor, J.W., Triple seasonal methods for short-term electricity demand forecasting, *European Journal of Operational Research*, 2010, *204*, 139152.
- Vassilopoulos, P., Price Signals in "Energy-only" Wholesale Electricity Markets: An Empirical Analysis of the Price Signal in France, *The Energy Journal*, 2010, *31* (3), 83–112.
- Weron, R. and A. Misiorek, Forecasting spot electricity prices: A comparison of parametric and semiparametric time series models, *International Journal of Forecasting*, 2008, *24*, 744–763.
- Yang, Y., Combining forecasting procedures: some theoretical results, *Econometric Theory*, 2004, *20*, 176–222.
- Zou, H. and Y. Yang, 2004. Combining time series models for forecasting, *International journal of Forecasting*, 2004, *20*, 6984.

Table 1a: Individual models

Model	Code
Autoregressive model	AR
Exponential model	EXPO
AR model augmented with forecasted demand	AR-X
Linear regression with regressors selected with a stepwise procedure	LSTEP
MS model, 2 regimes, FTP, variation of intercept and variance	MS_1
MS model, 2 regimes, FTP, variation of intercept, AR coefficients and variance	MS_2
MS model, 2 regimes, FTP, variation of all coefficients	MS_3
MS model, 2 regimes, TVTP, variation of intercept and variance	MSV_1
MS model, 2 regimes, TVTP, variation of intercept, AR coefficients and variance	MSV_2
MS model, 2 regimes, TVTP, variation of all coefficients	MSV_3
MS model, 3 regimes, FTP, variation of intercept and variance	MS3_1
MS model, 3 regimes, FTP, variation of intercept, AR coefficients and variance	MS3_2
STR model, variation of intercept	STR_1
STR model, variation of intercept and AR coefficients	STR_2
STR model, variation of all coefficients	STR_3

Table 1b: Combination of models

Model combinations	Code
Combination by average, all models	C1
Combination by average, linear models	C1_L
Combination by average, non linear models	C1_NL
Combination with Bates and Granger weights, all models	C2
Combination with Bates and Granger weights, linear models	C2_L
Combination with Bates and Granger weights, non linear models	C2_NL
Combination with AFTER weights, all models	C3
Combination with AFTER weights, linear models	C3_L
Combination with AFTER weights, non linear models	C3_NL
Combination with PLMA weights, all models	C4
Combination with PLMA weights, linear models	C4_L
Combination with PLMA weights, non linear models	C4_NL

Table 2a: Forecast evaluation - Winter

	SEASONAL								NON SEASONAL							
	RMSE				MAE				RMSE				MAE			
	3	9	14	20	3	9	14	20	3	9	14	20	3	9	14	20
AR	0.192	0.499	0.153	0.224	0.144	0.253	0.101	0.152	0.214	0.519	0.165	0.244	0.168	0.305	0.103	0.169
EXPO	0.179	0.456	0.162	0.231	0.139	0.241	0.107	0.165	0.213	0.464	0.180	0.243	0.171	0.280	0.119	0.170
AR-X	0.253	0.458	0.183	0.261	0.178	0.224	0.106	0.169	0.521	0.465	0.166	0.260	0.351	0.218	0.115	0.168
LSTEP	0.183	0.530	0.202	0.239	0.142	0.256	0.111	0.164	0.329	0.503	0.170	0.253	0.222	0.258	0.117	0.181
MS_1	0.197	0.435	0.180	0.255	0.131	0.222	0.111	0.167	0.213	0.428	0.175	0.247	0.179	0.206	0.120	0.167
MS_2	0.233	0.458	0.192	0.289	0.144	0.225	0.109	0.195	0.423	0.538	0.170	0.326	0.271	0.246	0.123	0.195
MS_3	0.254	0.429	0.188	0.277	0.175	0.215	0.109	0.177	0.520	0.535	0.172	0.263	0.347	0.244	0.121	0.171
MSV_1	0.213	0.436	0.175	0.250	0.143	0.221	0.104	0.170	0.270	0.429	0.172	0.256	0.224	0.210	0.121	0.171
MSV_2	0.261	0.455	0.149	0.264	0.161	0.219	0.096	0.173	0.521	0.499	0.167	0.238	0.351	0.241	0.117	0.151
MSV_3	0.253	0.550	0.184	0.268	0.177	0.254	0.110	0.176	0.521	0.470	0.166	0.281	0.351	0.224	0.116	0.173
MS3_1	0.212	0.532	0.166	0.269	0.133	0.268	0.106	0.172	0.199	0.468	0.169	0.206	0.165	0.214	0.124	0.139
MS3_2	0.253	0.462	0.196	0.281	0.179	0.229	0.114	0.188	0.341	0.517	0.202	0.337	0.216	0.266	0.146	0.193
STR_1	0.237	0.463	0.179	0.272	0.151	0.222	0.107	0.168	0.378	0.482	0.194	0.246	0.248	0.227	0.119	0.162
STR_2	0.288	0.518	0.172	0.278	0.177	0.232	0.110	0.172	0.555	0.419	0.158	0.304	0.272	0.200	0.104	0.182
STR_3	0.300	0.563	0.220	0.285	0.181	0.232	0.123	0.174	0.575	0.417	0.161	0.259	0.274	0.197	0.104	0.180
C1	0.209	0.472	0.172	0.254	0.140	0.230	0.104	0.166	0.320	0.449	0.161	0.245	0.217	0.219	0.110	0.163
C1_L	0.209	0.476	0.174	0.254	0.141	0.230	0.104	0.166	0.270	0.449	0.161	0.246	0.188	0.216	0.109	0.166
C1_NL	0.240	0.477	0.193	0.252	0.151	0.237	0.114	0.173	0.248	0.465	0.177	0.237	0.202	0.225	0.117	0.165
C2	0.189	0.471	0.168	0.244	0.132	0.229	0.102	0.161	0.246	0.455	0.159	0.244	0.175	0.232	0.109	0.165
C2_L	0.161	0.478	0.161	0.229	0.122	0.235	0.099	0.156	0.211	0.476	0.162	0.242	0.163	0.261	0.110	0.167
C2_NL	0.170	0.478	0.167	0.229	0.126	0.234	0.101	0.157	0.191	0.475	0.163	0.243	0.153	0.258	0.110	0.168
C3	0.196	0.466	0.194	0.232	0.142	0.243	0.115	0.164	0.226	0.469	0.181	0.257	0.169	0.230	0.118	0.178
C3_L	0.164	0.473	0.161	0.227	0.124	0.234	0.099	0.156	0.185	0.476	0.163	0.242	0.149	0.261	0.110	0.167
C3_NL	0.237	0.471	0.179	0.267	0.153	0.228	0.107	0.171	0.366	0.443	0.165	0.248	0.242	0.208	0.112	0.162
C4	0.233	0.477	0.179	0.267	0.150	0.229	0.107	0.171	0.313	0.444	0.164	0.250	0.222	0.206	0.110	0.165
C4_L	0.242	0.520	0.183	0.269	0.149	0.243	0.107	0.172	0.251	0.465	0.170	0.236	0.207	0.224	0.112	0.163
C4_NL	0.234	0.473	0.179	0.267	0.150	0.227	0.107	0.170	0.336	0.443	0.164	0.250	0.230	0.209	0.109	0.163

Notes: This table reports the RMSE and MAE criteria for hours 3, 9, 14 and 20. The left panel provides the forecast-evaluation criteria for the seasonal models and the right panel the results for the non seasonal specifications.

Table 2b: Forecast evaluation - Spring

	SEASONAL								NON SEASONAL							
	RMSE				MAE				RMSE				MAE			
	3	9	14	20	3	9	14	20	3	9	14	20	3	9	14	20
AR	0.250	0.150	0.154	0.121	0.170	0.115	0.119	0.094	0.245	0.146	0.164	0.128	0.171	0.114	0.125	0.097
EXPO	0.244	0.157	0.153	0.123	0.173	0.122	0.118	0.097	0.248	0.149	0.158	0.125	0.177	0.118	0.123	0.099
AR-X	0.181	0.111	0.127	0.119	0.129	0.095	0.096	0.101	0.192	0.140	0.145	0.123	0.149	0.111	0.106	0.098
LSTEP	0.192	0.117	0.130	0.113	0.131	0.097	0.102	0.092	0.167	0.136	0.153	0.109	0.121	0.114	0.107	0.089
MS_1	0.169	0.112	0.129	0.115	0.113	0.096	0.099	0.098	0.202	0.138	0.145	0.117	0.147	0.106	0.107	0.092
MS_2	0.157	0.111	0.127	0.120	0.101	0.095	0.097	0.102	0.204	0.132	0.147	0.122	0.153	0.106	0.107	0.097
MS_3	0.168	0.111	0.128	0.118	0.108	0.096	0.096	0.101	0.189	0.135	0.150	0.123	0.138	0.107	0.110	0.098
MSV_1	0.158	0.111	0.129	0.121	0.107	0.096	0.098	0.102	0.186	0.136	0.145	0.118	0.133	0.105	0.109	0.096
MSV_2	0.165	0.110	0.127	0.120	0.117	0.095	0.097	0.100	0.193	0.153	0.147	0.121	0.145	0.119	0.107	0.098
MSV_3	0.161	0.111	0.130	0.121	0.105	0.096	0.097	0.101	0.183	0.138	0.150	0.126	0.133	0.110	0.105	0.100
MS3_1	0.178	0.107	0.125	0.110	0.117	0.090	0.103	0.087	0.200	0.141	0.134	0.107	0.140	0.112	0.100	0.084
MS3_2	0.222	0.110	0.127	0.160	0.138	0.093	0.094	0.114	0.181	0.139	0.161	0.122	0.130	0.112	0.117	0.093
STR_1	0.173	0.110	0.127	0.121	0.123	0.096	0.097	0.103	0.203	0.139	0.149	0.107	0.160	0.109	0.103	0.089
STR_2	0.169	0.116	0.125	0.129	0.123	0.097	0.097	0.109	0.222	0.166	0.147	0.106	0.178	0.129	0.119	0.090
STR_3	0.170	0.121	0.129	0.124	0.122	0.102	0.099	0.107	0.222	0.179	0.144	0.118	0.178	0.139	0.117	0.097
C1	0.162	0.113	0.127	0.111	0.109	0.095	0.097	0.092	0.177	0.138	0.144	0.111	0.131	0.109	0.105	0.087
C1_L	0.160	0.113	0.127	0.111	0.107	0.096	0.098	0.092	0.177	0.138	0.144	0.110	0.130	0.109	0.105	0.087
C1_NL	0.167	0.112	0.127	0.115	0.110	0.096	0.098	0.095	0.188	0.138	0.143	0.111	0.128	0.110	0.106	0.093
C2	0.165	0.117	0.129	0.110	0.112	0.097	0.098	0.090	0.181	0.137	0.145	0.109	0.135	0.108	0.107	0.086
C2_L	0.188	0.128	0.135	0.110	0.131	0.103	0.105	0.088	0.196	0.135	0.151	0.112	0.145	0.106	0.113	0.088
C2_NL	0.180	0.124	0.135	0.109	0.128	0.101	0.104	0.087	0.188	0.137	0.151	0.110	0.142	0.108	0.113	0.087
C3	0.201	0.119	0.130	0.116	0.134	0.099	0.101	0.095	0.192	0.138	0.149	0.115	0.134	0.112	0.109	0.092
C3_L	0.186	0.129	0.137	0.110	0.132	0.104	0.105	0.088	0.198	0.137	0.152	0.112	0.147	0.108	0.114	0.088
C3_NL	0.158	0.111	0.126	0.114	0.106	0.095	0.097	0.096	0.174	0.141	0.143	0.114	0.127	0.112	0.104	0.090
C4	0.157	0.111	0.126	0.115	0.104	0.095	0.097	0.096	0.175	0.141	0.143	0.113	0.126	0.111	0.104	0.090
C4_L	0.162	0.111	0.127	0.115	0.107	0.095	0.098	0.096	0.192	0.139	0.142	0.111	0.133	0.110	0.105	0.091
C4_NL	0.157	0.111	0.126	0.114	0.105	0.096	0.097	0.096	0.175	0.141	0.143	0.113	0.126	0.112	0.104	0.090

Notes: See Table (2a).

Table 2c: Forecast evaluation - Summer

	SEASONAL								NON SEASONAL							
	RMSE				MAE				RMSE				MAE			
	3	9	14	20	3	9	14	20	3	9	14	20	3	9	14	20
AR	0.236	0.068	0.067	0.076	0.202	0.058	0.050	0.058	0.200	0.076	0.066	0.073	0.137	0.059	0.049	0.058
EXPO	0.200	0.067	0.067	0.075	0.148	0.055	0.053	0.059	0.206	0.069	0.067	0.076	0.148	0.056	0.053	0.061
AR-X	0.266	0.059	0.088	0.069	0.238	0.044	0.072	0.056	0.175	0.091	0.065	0.074	0.145	0.074	0.053	0.059
LSTEP	0.182	0.061	0.070	0.072	0.147	0.051	0.053	0.059	0.175	0.106	0.074	0.077	0.134	0.087	0.062	0.062
MS ₁	0.240	0.063	0.092	0.064	0.219	0.047	0.075	0.049	0.177	0.085	0.066	0.067	0.137	0.067	0.053	0.055
MS ₂	0.237	0.064	0.090	0.070	0.215	0.048	0.073	0.056	0.182	0.086	0.064	0.070	0.141	0.068	0.051	0.058
MS ₃	0.266	0.060	0.087	0.067	0.238	0.045	0.071	0.053	0.175	0.125	0.065	0.074	0.145	0.108	0.051	0.059
MSV ₁	0.227	0.057	0.096	0.065	0.204	0.044	0.078	0.053	0.172	0.096	0.067	0.067	0.134	0.079	0.055	0.055
MSV ₂	0.258	0.062	0.089	0.075	0.218	0.049	0.073	0.060	0.178	0.088	0.064	0.079	0.139	0.071	0.052	0.064
MSV ₃	0.266	0.058	0.089	0.069	0.238	0.043	0.073	0.056	0.175	0.112	0.067	0.074	0.145	0.094	0.054	0.059
MS3 ₁	0.226	0.064	0.072	0.069	0.199	0.049	0.057	0.051	0.175	0.084	0.071	0.063	0.127	0.067	0.055	0.050
MS3 ₂	0.264	0.063	0.090	0.072	0.237	0.054	0.073	0.058	0.172	0.102	0.073	0.067	0.138	0.085	0.058	0.057
STR ₁	0.272	0.065	0.076	0.094	0.234	0.051	0.062	0.066	0.174	0.103	0.075	0.094	0.150	0.084	0.062	0.075
STR ₂	0.271	0.064	0.077	0.076	0.240	0.046	0.062	0.060	0.173	0.101	0.073	0.096	0.143	0.086	0.058	0.073
STR ₃	0.255	0.076	0.079	0.075	0.202	0.051	0.064	0.060	0.176	0.101	0.079	0.095	0.147	0.086	0.064	0.071
C1	0.222	0.057	0.077	0.068	0.200	0.043	0.062	0.054	0.170	0.087	0.064	0.070	0.136	0.069	0.051	0.058
C1 _L	0.217	0.058	0.074	0.067	0.193	0.043	0.059	0.054	0.171	0.083	0.064	0.068	0.138	0.065	0.051	0.056
C1 _{NL}	0.208	0.060	0.072	0.067	0.168	0.046	0.054	0.054	0.175	0.074	0.067	0.069	0.141	0.060	0.052	0.058
C2	0.206	0.057	0.071	0.068	0.184	0.043	0.056	0.054	0.171	0.078	0.063	0.068	0.136	0.060	0.050	0.056
C2 _L	0.185	0.058	0.066	0.069	0.162	0.047	0.050	0.055	0.176	0.076	0.063	0.068	0.131	0.058	0.049	0.056
C2 _{NL}	0.178	0.058	0.066	0.069	0.151	0.046	0.050	0.055	0.172	0.074	0.063	0.069	0.133	0.058	0.049	0.055
C3	0.192	0.060	0.070	0.071	0.151	0.046	0.052	0.056	0.178	0.072	0.067	0.074	0.142	0.058	0.052	0.061
C3 _L	0.183	0.058	0.066	0.069	0.158	0.047	0.050	0.055	0.176	0.073	0.063	0.068	0.132	0.057	0.049	0.055
C3 _{NL}	0.241	0.059	0.083	0.069	0.218	0.043	0.068	0.055	0.170	0.094	0.066	0.072	0.138	0.078	0.053	0.059
C4	0.243	0.059	0.082	0.068	0.218	0.043	0.067	0.055	0.171	0.092	0.065	0.072	0.140	0.076	0.052	0.058
C4 _L	0.241	0.061	0.076	0.067	0.213	0.046	0.061	0.053	0.174	0.087	0.066	0.068	0.141	0.070	0.052	0.055
C4 _{NL}	0.244	0.059	0.082	0.069	0.218	0.043	0.067	0.055	0.171	0.093	0.065	0.074	0.140	0.077	0.052	0.060

Notes: See Table (2a).

Table 2d: Forecast evaluation - Fall

	SEASONAL								NON SEASONAL							
	RMSE				MAE				RMSE				MAE			
	3	9	14	20	3	9	14	20	3	9	14	20	3	9	14	20
AR	0.303	0.072	0.073	0.093	0.193	0.049	0.058	0.071	0.308	0.080	0.071	0.104	0.196	0.059	0.054	0.079
EXPO	0.310	0.076	0.071	0.090	0.200	0.055	0.055	0.070	0.329	0.081	0.070	0.096	0.209	0.060	0.055	0.077
AR-X	0.250	0.064	0.064	0.083	0.185	0.052	0.049	0.066	0.263	0.080	0.061	0.091	0.167	0.063	0.052	0.067
LSTEP	0.251	0.091	0.064	0.078	0.173	0.065	0.052	0.062	0.240	0.108	0.081	0.085	0.158	0.083	0.068	0.068
MS ₁	0.266	0.069	0.062	0.091	0.178	0.055	0.049	0.068	0.263	0.070	0.063	0.108	0.163	0.054	0.055	0.092
MS ₂	0.263	0.067	0.064	0.092	0.167	0.054	0.049	0.073	0.271	0.072	0.063	0.094	0.171	0.055	0.053	0.073
MS ₃	0.286	0.064	0.064	0.113	0.186	0.051	0.050	0.083	0.263	0.085	0.061	0.102	0.167	0.066	0.051	0.084
MSV ₁	0.279	0.071	0.071	0.092	0.173	0.051	0.053	0.067	0.252	0.074	0.064	0.113	0.164	0.057	0.053	0.098
MSV ₂	0.255	0.068	0.063	0.090	0.156	0.056	0.050	0.068	0.260	0.073	0.061	0.089	0.171	0.055	0.049	0.070
MSV ₃	0.237	0.063	0.063	0.146	0.166	0.053	0.047	0.104	0.263	0.093	0.063	0.106	0.167	0.073	0.053	0.086
MS3 ₁	0.266	0.068	0.063	0.092	0.184	0.053	0.050	0.069	0.249	0.072	0.060	0.093	0.152	0.053	0.051	0.076
MS3 ₂	0.266	0.065	0.064	0.100	0.172	0.052	0.049	0.079	0.263	0.082	0.068	0.106	0.165	0.059	0.055	0.086
STR ₁	0.297	0.081	0.071	0.093	0.217	0.058	0.058	0.070	0.282	0.088	0.084	0.083	0.195	0.069	0.071	0.064
STR ₂	0.240	0.113	0.067	0.100	0.169	0.077	0.053	0.078	0.431	0.108	0.087	0.084	0.269	0.088	0.075	0.065
STR ₃	0.276	0.106	0.071	0.088	0.207	0.073	0.055	0.064	0.449	0.123	0.087	0.087	0.294	0.093	0.074	0.065
C1	0.254	0.068	0.063	0.086	0.159	0.052	0.050	0.066	0.263	0.079	0.065	0.088	0.158	0.060	0.055	0.069
C1 _L	0.255	0.067	0.063	0.085	0.157	0.051	0.050	0.066	0.260	0.076	0.064	0.087	0.160	0.057	0.054	0.068
C1 _{NL}	0.264	0.067	0.064	0.086	0.163	0.052	0.051	0.067	0.261	0.074	0.063	0.087	0.160	0.054	0.054	0.067
C2	0.259	0.068	0.063	0.083	0.158	0.050	0.050	0.064	0.266	0.077	0.065	0.086	0.157	0.056	0.054	0.066
C2 _L	0.263	0.070	0.063	0.079	0.166	0.048	0.050	0.061	0.271	0.080	0.066	0.087	0.156	0.056	0.053	0.066
C2 _{NL}	0.268	0.069	0.063	0.078	0.163	0.048	0.050	0.061	0.268	0.078	0.066	0.086	0.158	0.056	0.053	0.067
C3	0.278	0.069	0.064	0.081	0.174	0.049	0.052	0.065	0.266	0.079	0.065	0.087	0.167	0.060	0.054	0.067
C3 _L	0.269	0.070	0.064	0.080	0.166	0.048	0.050	0.061	0.276	0.078	0.066	0.087	0.159	0.056	0.053	0.067
C3 _{NL}	0.252	0.069	0.064	0.090	0.159	0.054	0.051	0.069	0.263	0.081	0.066	0.090	0.163	0.062	0.056	0.071
C4	0.253	0.067	0.064	0.090	0.159	0.053	0.051	0.069	0.260	0.077	0.064	0.089	0.163	0.059	0.055	0.070
C4 _L	0.259	0.066	0.064	0.090	0.159	0.053	0.051	0.068	0.260	0.074	0.063	0.087	0.160	0.055	0.054	0.068
C4 _{NL}	0.254	0.070	0.065	0.089	0.161	0.054	0.051	0.069	0.261	0.079	0.067	0.088	0.162	0.061	0.057	0.070

Notes: See Table (2a).

Table 3: Best models in terms of RMSE, MAE and MAPE

	in terms of RMSE										in terms of MAE										in terms of MAPE									
Model	Winter		Spring		Summer		Fall		All	Winter		Spring		Summer		Fall		All	Winter		Spring		Summer		Fall		All			
	SA	NS	SA	NS	SA	NS	SA	NS	all	SA	NS	SA	NS	SA	NS	SA	NS	all	SA	NS	SA	NS	SA	NS	SA	NS	all			
Autoreg	14	5	1	0	5	1	0	1	27	13	5	1	0	7	4	3	3	36	13	7	1	0	6	4	3	3	37			
AR	9	4	0	0	3	0	0	0	16	9	5	0	0	4	3	2	1	24	9	7	0	0	4	3	2	1	26			
expo	5	1	1	0	2	1	0	1	11	4	0	1	0	3	1	1	2	12	4	0	1	0	2	1	1	2	11			
Linear	0	4	5	8	7	5	10	7	46	0	2	4	5	7	9	7	4	38	0	1	4	5	8	9	7	5	39			
AR-X	0	1	0	0	0	0	0	0	1	0	1	0	0	0	2	0	0	3	0	1	0	0	0	2	0	0	3			
LSTEP	0	3	5	8	7	5	10	7	45	0	1	4	5	7	7	7	4	35	0	0	4	5	8	7	7	5	36			
MS	8	10	13	11	9	15	11	15	92	7	13	15	15	9	10	11	15	95	8	13	15	15	9	10	11	14	95			
MS_1	2	2	1	1	2	1	1	3	13	4	3	0	2	2	0	1	2	14	4	3	0	2	2	0	1	2	14			
MS_2	0	0	4	1	1	1	0	1	8	0	1	3	0	1	0	1	2	8	0	1	3	1	1	0	1	2	9			
MS_3	1	0	0	0	1	1	4	0	7	1	0	1	1	1	0	2	0	6	1	0	1	1	1	0	2	0	6			
MSV_1	1	1	0	4	2	3	1	2	14	0	1	1	3	2	0	1	0	8	1	1	1	2	2	0	1	0	8			
MSV_2	1	0	0	1	0	2	0	2	6	1	0	2	1	0	2	2	2	10	2	0	2	1	0	3	2	2	12			
MSV_3	1	3	2	0	0	1	2	1	10	0	3	2	0	1	0	1	0	7	0	3	2	0	1	0	1	0	7			
MS3_1	0	4	5	4	3	4	3	6	29	0	5	5	7	1	6	3	8	35	0	5	5	7	1	6	3	8	35			
MS3_2	2	0	1	0	0	2	0	0	5	1	0	1	1	1	2	0	1	7	0	0	1	1	1	1	0	0	4			
STR	2	5	5	5	3	3	3	1	27	4	4	4	4	1	1	3	2	23	3	3	4	4	1	1	3	2	21			
STR_1	0	0	4	0	2	1	2	1	10	3	0	3	0	1	1	2	2	12	3	0	3	0	1	1	2	2	12			
STR_2	2	3	1	3	1	1	1	0	12	1	2	1	0	0	0	0	0	4	0	1	1	0	0	0	0	0	2			
STR_3	0	2	0	2	0	1	0	0	5	0	2	0	4	0	0	1	0	7	0	2	0	4	0	0	1	0	7			
Total	24	24	24	24	24	24	24	24	192	24	24	24	24	24	24	24	24	192	24	24	24	24	24	24	24	24	192			

Notes: This table provides the number of times over the 24 hours of each season that each specification yields the lowest RMSE (first block of columns), MAE (second block) and MAPE (last block). The columns SA contain the results for the seasonal range and the columns NS the results for the non-seasonal specifications.

Table 4: Rank of individual and pooled forecasts in terms of RMSE, MAE and MAPE

Rank with RMSE										Rank with of MAE										Rank with of MAPE									
Model	Winter		Spring		Summer		Fall		All	Winter		Spring		Summer		Fall		All	Winter		Spring		Summer		Fall		All		
	SA	NS	SA	NS	SA	NS	SA	NS		SA	NS	SA	NS	SA	NS	SA	NS		SA	NS	SA	NS	SA	NS	SA	NS		SA	NS
AR	6	8	27	27	15	24	27	24	26	8	7	27	27	14	18	25	23	22	8	7	27	27	14	18	26	23	22		
EXPO	3	6	26	26	12	27	26	25	23	7	15	26	26	9	24	24	24	23	7	15	26	26	10	24	24	24	24		
AR-X	17	23	14	23	25	17	9	18	18	19	20	13	24	25	20	17	20	20	19	18	12	24	25	19	16	20	19		
LSTEP	13	17	13	9	9	18	2	12	13	13	19	15	10	10	17	9	18	14	11	18	15	10	9	17	9	18	14		
MS_1	10	14	23	21	18	16	17	8	16	10	18	23	15	18	16	20	10	16	10	16	23	17	18	16	20	11	16		
MS_2	25	24	14	20	24	20	20	11	20	23	24	14	17	21	19	21	13	18	23	24	14	18	20	20	21	13	18		
MS_3	23	27	9	22	20	19	11	22	19	24	26	6	23	22	21	12	22	19	24	26	6	23	21	21	13	22	20		
MSV_1	11	18	20	11	14	9	21	13	15	11	23	20	12	13	9	18	19	15	13	22	21	11	13	9	18	19	15		
MSV_2	18	22	9	18	27	15	19	6	17	15	25	9	21	26	15	19	13	17	15	25	8	20	26	15	19	14	17		
MSV_3	22	20	11	24	26	22	18	20	21	25	21	8	22	27	23	16	21	21	25	21	9	22	27	23	16	21	21		
MS3_1	12	9	12	10	11	11	16	2	9	11	10	12	9	12	7	15	3	10	14	10	13	9	12	7	15	2	10		
MS3_2	20	26	18	25	22	21	24	16	22	22	27	17	25	24	22	26	16	25	22	27	18	25	24	22	25	16	25		
STR_1	21	25	16	17	21	25	23	23	24	21	17	16	15	19	27	22	25	24	21	20	16	14	19	27	23	25	23		
STR_2	26	19	19	16	19	23	22	26	25	26	16	21	19	23	26	23	26	26	26	16	20	19	23	25	22	26	26		
STR_3	27	21	25	15	22	26	25	27	27	27	22	24	18	20	25	27	27	27	26	23	24	16	22	26	27	27	27		
C1	7	5	5	1	7	3	4	9	3	6	6	3	2	7	5	6	7	3	6	6	4	2	8	4	5	7	3		
C1_L	9	7	3	4	5	1	1	4	2	5	3	2	4	5	2	4	2	2	5	2	2	5	5	3	3	3	2		
C1_NL	14	13	7	3	6	5	7	1	6	14	9	11	3	7	8	7	1	7	12	8	11	3	7	8	7	1	7		
C2	5	3	8	7	4	2	3	7	1	4	1	10	8	3	1	1	5	1	3	1	10	8	3	1	1	4	1		
C2_L	2	2	21	13	1	7	6	19	5	2	4	22	11	4	4	3	12	5	2	4	22	12	4	5	4	12	5		
C2_NL	4	4	17	12	2	4	5	17	4	3	5	17	13	1	3	2	6	4	4	5	17	13	1	2	2	6	4		
C3	8	15	22	14	8	14	10	10	14	9	14	19	14	6	11	8	11	11	9	13	19	15	6	11	8	9	11		
C3_L	1	1	24	19	3	10	14	21	7	1	2	25	20	2	6	5	9	6	1	3	25	21	2	6	6	10	6		
C3_NL	15	10	2	5	15	8	15	14	11	17	12	4	5	16	13	14	17	12	17	12	3	4	16	13	14	17	12		
C4	16	11	1	8	13	5	8	5	8	16	8	1	6	15	10	11	8	9	16	9	1	6	15	10	11	8	9		
C4_L	23	15	6	2	10	13	12	3	10	20	11	7	1	11	12	10	4	8	19	11	7	1	11	12	10	4	8		
C4_NL	19	12	3	6	17	12	13	15	12	18	13	5	7	17	14	13	15	13	18	14	5	7	17	14	12	15	13		

This table reports the rank of each model in terms of RMSE, MAE and MAPE at each season and over the entire year. The columns SA contain the results for the seasonal range and the columns NS the results for the non-seasonal sample. For instance, C3.L displays the lowest RMSE in winter (seasonal range). The last columns of each panel rank the models according to their average rank over the 4 seasons and the two seasonal treatments.

Table 5: Comparison of non-linear, pooled, seasonal models (in %)

	Winter		Spring		Summer		Fall	
	SA	NS	SA	NS	SA	NS	SA	NS
<i>Comparison of the loss functions (MSE and MAE)</i>								
Non-linear beats linear	31.8	36.1	72.6	67.7	42.0	58.5	55.4	57.3
MS3_1 beats linear	41.7	54.9	77.8	77.8	55.6	77.8	69.4	79.9
Combined beats individual	67.3	70.4	61.9	70.9	74.6	73.7	71.9	66.6
Seasonal beats non seasonal	69.1		77.6		30.4		54.5	
Seasonal beats non seasonal (AR & EXPO)	95.8		77.1		52.1		70.8	
Seasonal linear beats non seasonal (linear models)	89.6		80.9		36.1		64.6	
<i>Significance of difference with D-M test</i>								
Non-linear beats linear	9.2	11.7	41.2	29.2	15.6	21.2	20.6	29.7
MS3_1 beats linear	14.6	13.9	52.1	47.2	27.1	27.1	20.8	52.8
Combined beats individual	25.2	26.9	27.1	36.4	43.9	38.9	35.2	37.5
Seasonal beats non seasonal	32.4		42.7		13.4		24.5	
Seasonal beats non seasonal (AR & EXPO)	50.0		51.0		20.8		41.7	
Seasonal linear beats non seasonal (linear models)	46.9		50.4		16.7		33.0	
<i>Significance of difference with G-W test</i>								
Non-linear beats linear	3.7	6.9	30.8	16.7	8.1	10.2	7.5	19.1
MS3_1 beats linear	7.6	5.6	38.9	30.6	13.9	15.3	3.5	33.3
Combined beats individual	11.8	16.5	16.3	19.4	32.5	26.2	21.2	29.4
Seasonal beats non seasonal	21.2		24.6		8.6		15.4	
Seasonal beats non seasonal (AR & EXPO)	38.5		26.0		4.2		14.6	
Seasonal linear beats non seasonal (linear models)	34.7		30.6		7.3		16.3	
<i>Encompassing test</i>								
H0: linear encompasses non-linear	34.7	40.0	71.6	68.2	61.2	64.4	57.2	61.9
H0: linear encompasses MS3_1	38.9	47.2	73.6	76.4	76.4	75.0	65.3	79.2
H0: individual encompasses combined	29.3	39.7	43.3	56.3	66.8	59.4	54.4	52.5
H0: NS encompasses SA	34.1		66.5		31.3		60.7	
H0: NS encompasses SA (AR & EXPO)	52.1		56.3		33.3		50.0	
H0: NS linear encompasses SA (linear models)	45.8		66.0		27.8		54.9	

Notes: The first block reports the percentage of times over the 24 hours that (a) yields a lower loss than (b) with MSE or MAE loss functions. For example, non-linear models beat the linear specifications in 72.6% of cases in spring (seasonal range). The second block provides the results of the one-sided Diebold and Mariano test for the null hypothesis that (a) has the same forecast accuracy than (b) (with MSE or MAE loss functions). Each entry represents the percentage of times that (a) is significantly better than (b). The third block gives the results of the Giacomini and White test. Each entry represents the percentage of times that (a) is significantly better than (b) (with MSE or MAE loss functions). The last block provides the results of the encompassing test of Harvey, Leybourne and Newbold for the null hypothesis that (a) encompasses (b). Each entry represents the percentage of rejection of the null hypothesis. The results of the test are reported at a 10% significance level.

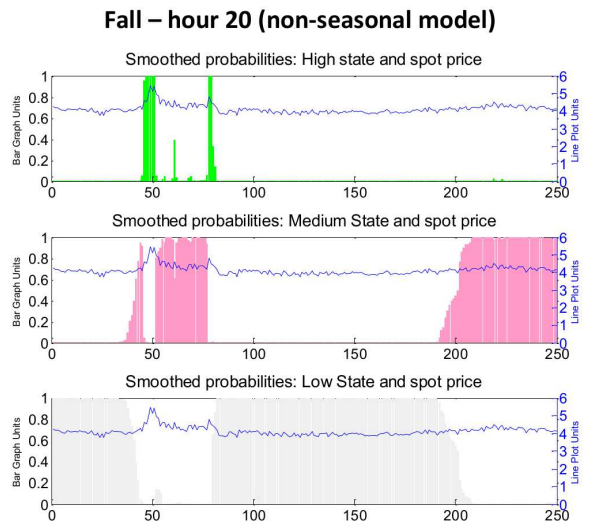
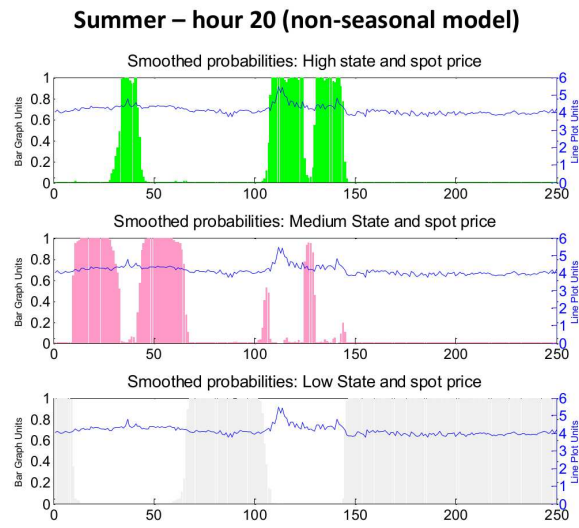
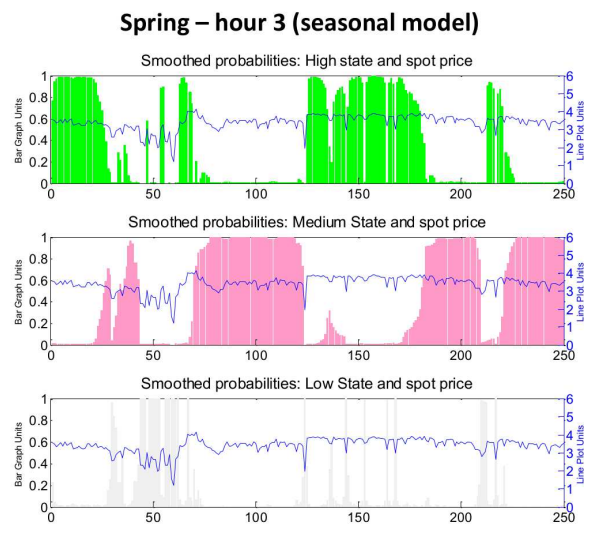
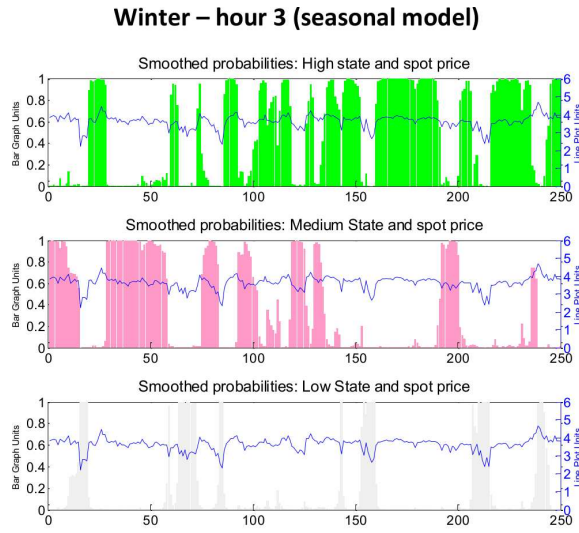


Figure 6: MS3_1 model - Smoothed probabilities