Electric Power Systems Research 167 (2019) 240–251

Contents lists available at ScienceDirect

Electric Power Systems Research

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

Electricity price forecasting in the short term hybridising fundamental and econometric modelling☆

Rodrigo A. de Marcos*, Antonio Bello, Javier Reneses

Institute for Research in Technology, Technical School of Engineering (ICAI), Universidad Pontificia Comillas, 28015 Madrid, Spain

ARTICLE INFO

Keywords:
Econometric models
Electricity markets
Fundamental models
Hybrid models
Short-term forecasting

ABSTRACT

Traders and practitioners in diverse power exchanges are nowadays being most exposed to uncertainty than ever. The combination of several factors such as renewable generation and regulatory changes calls for suitable electricity price forecasting models that can deal with complex and unusual market conditions. Several authors have proposed combining fundamental approaches with econometric models in order to cover all relevant aspects for electricity price forecasting. This combination has shown positive results for medium-term horizons. However, this approach has rarely been carried out for short-term applications. Moreover, several day-to-day applications in electricity markets require fast responsiveness and accurate forecasts. All of these facts encourage this work's short-term hybrid electricity price forecasting model, which combines a cost-production optimisation (fundamental) model with an artificial neural network (econometric) model. In order to validate the advantages and contributions of the proposed model, it has been applied to a real-size power exchange with complex price dynamics, such as the Iberian electricity market. Moreover, its forecasting performance has been compared with those of the two individual components of the hybrid model as well as other well-recognised methods. The results of this comparison prove that the proposed forecasting model outperforms the benchmark models, especially in uncommon market circumstances.

1. Introduction and literature review

As of today, electricity market agents and participants are increasingly being exposed to price uncertainty and high market volatility, especially since the deregulation and liberalisation of power exchanges. Moreover, this uncertainty is further heightened by the growing penetration of renewable energy sources and global financial instability. Furthermore, regulatory and market structural changes are intensifying the complexity of the system. Therefore, the combination of these facts strongly encourages the use of electricity price forecasting models.

Other uses of electricity price forecasting models among traders and practitioners include risk management, speculation and strategic purposes. Moreover, these traders and practitioners usually take part in day-to-day applications and decision making, which call for fast forecasting approaches. Therefore, the fact that electricity market price forecasting models are highly demanded is not in question.

One way of classifying these models is by paying attention to their applications, purposes and aims, which are mostly related to the planning horizon, i.e. short- medium- and long-term. Short-term forecasting model uses regularly include statistical and econometric methods, while longer term applications generally involve fundamental modelling of the market dynamics [1].

The work presented in this manuscript focuses on electricity market price forecasting in the short term (horizons ranging from a few hours up to one week in advance), which plays a very important role in day-to-day market operations. Furthermore, a great number of research works confirm that electricity market prices exhibit volatile and non-stationary behaviour, making price forecasting a highly challenging task. Therefore, the current literature encompasses numerous forecasting methods in order to yield accurate and adequate forecasts, as can be seen in Ref. [2].

Some of the most traditional electricity price forecasting models in short-term applications are statistical/econometric methods, which involve time-series methods (e.g. ARIMA, GARCH [3–7]), Artificial Intelligence (AI) techniques (e.g. Support Vector Machine, Radial Basis Function Networks [8–13]) and combinations of both types of models [14–21].

These combinations have recently received a wide acceptance due to their proficiency at modelling linear (time series) as well as non-linear (AI) trends and patterns in datasets [22]. Furthermore, these

☆ This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

* Corresponding author.
E-mail addresses: demarcos@comillas.edu (R.A. de Marcos), abello@comillas.edu (A. Bello), javierr@comillas.edu (J. Reneses).

https://doi.org/10.1016/j.epsr.2018.10.034
Received 6 March 2018; Received in revised form 23 October 2018; Accepted 29 October 2018

models are capable of capturing the revealed behavioural aspects of market participants, such as strategic and speculative behaviour [23].

However, statistical models may not be able to fully incorporate market dynamics and operation to their electricity price forecasts. Furthermore, these models tend to have difficulties when it comes to representing regulatory and market structural changes. Moreover, another weakness of econometric models lies in the assumption that history repeats itself, which is not suitable in most cases for today's power markets, which are more volatile and complex than ever.

To cover these aspects, market clearing prices are estimated by means of fundamental methods, which are aimed at thoroughly modelling the power market, including all generation units and their technical features, such as production costs. Nevertheless, market clearing models are not usually resorted for short-term price forecasting applications, often due to their poor performance at capturing short-term price dynamics, as mentioned in Ref. [1].

Therefore, fundamental models are sometimes combined with other approaches, such as econometric methods, in order to improve their predictive performance. However, recent works prove that a great deal of work has been carried out in the context of these hybrid models for medium-term horizons, not only from a point-forecasting perspectives but also from probabilistic points of view [1,24,25].

This fundamental-econometric combination will be referred to as hybrid models for the remainder of the manuscript. One of the main advantages of these medium-term hybrid models is the capability to consider most relevant, economic drivers of electricity market prices, such as supply, demand, unit commitment, dispatch and technical constraints [23,24,26,27]. By doing so, the behaviour and operation of the power market is successfully incorporated to the electricity price forecasts, which is of great interest for its participants.

Fundamental procedures are increasingly being resorted to by traders and practitioners considering the ongoing structural and regulatory changes that most power exchanges are experiencing (e.g. new taxes, incentives for specific generation technologies, CO_2 emission allowances, etc.), which do have an effect on electricity market prices [28].

However, according to the work presented in Ref. [23], these models tend to provide less volatile and more flat predictions than those ultimately observed. Moreover, fundamental approaches that aim to thoroughly model power exchanges in hourly resolution are commonly avoided in short-term price forecasting applications due to extremely large size and resolution times in real power systems. This issue calls for simplification methods that are usually targeted at the temporal resolution or the structure of the system, such as the generation technology aggregation that has been carried out in Ref. [26]. This simplification provides a reduction of the volume of the input data, which may also be effectively reduced by means of variable selection and shrinkage tools.

It is well known that electricity prices are affected by a plethora of factors, ranging from weather-related variables (e.g. wind speed, rainfall amounts) to energy generation parameters (e.g. fuel prices, maintenance costs). Variable selection tools have been considered in economic and price forecasting frameworks in Refs. [27,29–31] with the purpose of screening out the least useful and most noisy predictors. These works cover several variable selection methods, including traditional approaches such as stepwise regression, least squares and principal components; as well as more complex approaches such as least absolute shrinkage selection operator (LASSO, [32]), ridge regression [33] and elastic nets (combination of LASSO and ridge regression, [34]). Some of these methods have been recently applied to electricity price forecasting in Ref. [35], whose results suggest utilising an elastic net above the other approaches due to its beneficial outcomes regarding day ahead electricity price forecasting accuracy by means of linear regression.

The above paragraphs have pointed out some insufficiencies and scarcities in the context of short-term electricity price forecasting that encourage motivation for the proposed model of this manuscript. Several hybrid forecasting methods have been mainly employed in medium- and long-term horizons, and thus their adequacy in the short term is either poor or untested.

Moreover, price forecasting models that combine fundamental and econometric approaches have proved beneficial in medium-term horizons, as these take several aspects into account such as market dynamics and structural/regulatory market changes (fundamentals), as well as strategic/speculative behaviour and linear/non-linear modelling capabilities (econometrics). However, the literature regarding this kind of hybrid models in short-term contexts is relatively scarce and thus it would be interesting to determine if the same advantages can be attained for the short term.

Therefore, the main objective of this work is to propose a novel short-term hybrid electricity price forecasting model, which takes advantage of the combination of fundamental and econometric approaches in order to capture not only the effects caused by structural and regulatory market changes, but also the strategic and speculative behaviour exhibited by market agents, as discussed in Ref. [23]. Thus, fundamental approaches are able to model changes in market agent bidding strategies, which are frequent in the short-term, as opposed to medium-term contexts where this effect can be ignored without any drawbacks.

The only work in the literature that considers the same kind of fundamental-econometric hybrids for short-term electricity price forecasting is the one presented in Ref. [26], although it yields daily average values for forecasting periods of one month instead of monthly values for periods of one day to one week (as considered in this manuscript). Furthermore, while the fundamental component of Ref. [26] represents a simplified bidding curve in which generation technologies are stacked together and an estimated generation price is obtained as a linear function of the different generation costs, the authors of this manuscript intend to consider a higher level of detail on their fundamental model by taking advantage of other elements, such as the operation of the power system, technical features of generation units, demand levels and interconnections with other markets. Moreover, the authors of Ref. [26] then utilise the estimated generation cost in a non-linear regression model, although it would be interesting to consider an AI approach, such as neural networks, on this stage of the hybrid model. However, there are no other similar hybrid short-term approaches in the literature, which thoroughly model the power exchange with hourly precision, to the best knowledge of the authors.

The main contributions of the paper are summarised as follows:

1. A novel hourly short-term hybrid forecasting tool is proposed and developed, which is based on a cost-production optimisation model that is linked to a neural network model. The computational efficiency of each model is enhanced individually in order to adapt the resulting hybrid forecasting model to the short-term electricity price forecasting contexts.
2. The forecasting performance of the proposed hybrid model is put to the test on a real-size market with complex price dynamics: the Iberian (Portuguese and Spanish) electricity market.
3. Benchmark models, including well-recognised methods and both individual components of the hybrid model, were tested on the same case studies so as to compare their forecasting performance with that of the proposed forecasting methodology.

The remainder of this manuscript is organised as follows: Section 2 describes this work's methodology. Section 3 presents the case studies in which the proposed forecasting method has been tested, as well as a comparison with other forecasting procedures. Section 4 contains the conclusions that were drawn in this work, including the suggestions for extensions and future developments of the proposed methodology. Additionally, Appendix A contains the main details of the fundamental model that was used in this work's hybrid forecasting approach.

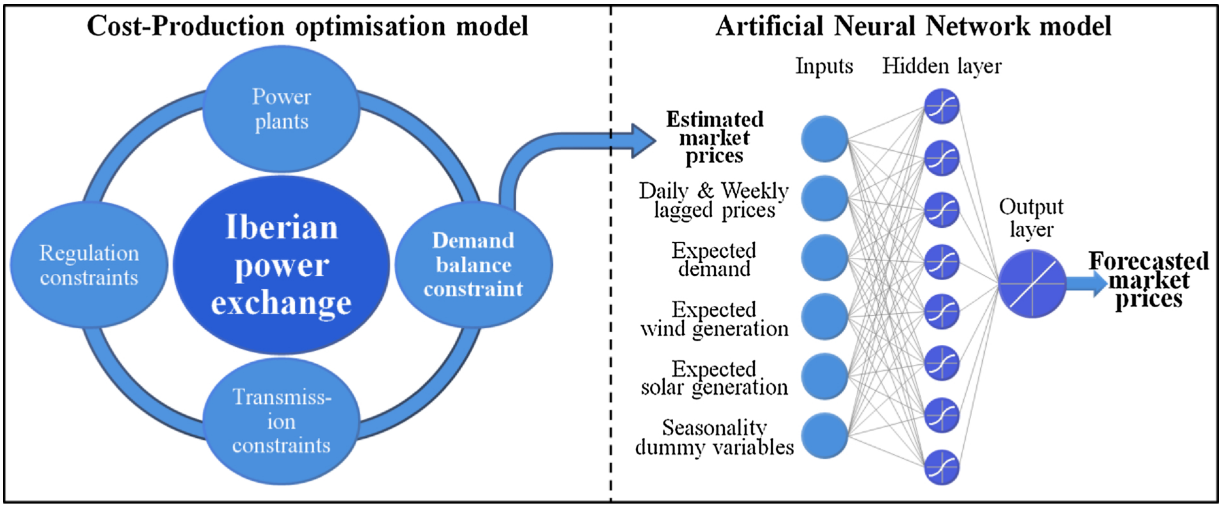

Fig. 1. Proposed hybrid forecasting model.

2. Proposed methodology

The general objective of this work is to create a new short-term hybrid electricity price forecasting model and put it to the test on a real, full-scale and complex electricity market, such as the Iberian power exchange. Forecasts have been considered for both one day and one week horizons so as to agree with the two ends of the short-term framework. The proposed fundamental-econometric hybrid model is displayed in Fig. 1.

The first half of the model, displayed at the left part of Fig. 1, is composed of a cost-production optimisation model that aims to simulate the market-clearing process by minimising the total system costs, which are constrained by generation unit technical features, regulation limits, transmission limits and the demand vs. generation balance. Thus, in this model, the estimated electricity market price can be obtained as the dual variable of the demand balance constraint.

The above optimisation model, which is similar to the one presented in Ref. [23], has been applied to the Iberian power exchange. Nevertheless, in order to decrease its runtime, certain simplifications were carried out. This simplification consisted of an aggregation of similar power plants, which are owned by the same market agent and have identical cost functions and other technical features. As a result, the optimisation problem size was significantly reduced and thus the estimation time of electricity market prices was reduced from a few minutes to a few seconds. Specifically, the simplified optimisation problem for a forecasting period of one week consists of 12,440 equations and 71,024 variables with a total runtime of 3.91 s and a maximum RAM usage of 76 MB on a 64-bit Windows 7 PC with 16 GB installed RAM and the following processor: Intel® Core™ i7-3770 CPU@ 3.40 GHz of 4 cores and 8 logical processors. Further details of this fundamental model can be found at Appendix A.

However, even though other variables were simultaneously calculated (e.g. transmitted power, emissions, generation outputs), only the estimated price was utilised for the second half of the proposed hybrid model. Nevertheless, the authors do not deny the possibility that considering other outputs in the statistic model may prove beneficial. Moreover, an additional term may be incorporated to the objective function in order to consider agent strategic behaviour, which is defined in Ref. [23] as the “conjectured-price response”. Nevertheless, the authors discarded this possibility due to the fact that agent strategic behaviour can also be incorporated to the forecasts by means of econometric methods, such as the second half of the proposed hybrid approach.

As shown in Fig. 1, the estimated market price is used as an additional input variable to the neural network model, which is trained alongside other inputs in order to produce short-term electricity price forecasts. The other input variables that were considered are as follows:

- Lagged electricity market prices. The following lags were considered:
  - One day
  - Two days
  - One week
  - Two weeks
- Expected system demand, expected wind generation and expected solar generation.
- Dummy variables indicating if a day belongs to one of the following day types:
  - Business day or non-business day, i.e. Sundays and holidays
  - Saturday or non-Saturday.

These variables are effortlessly obtainable from the Spanish ISO information website [36], and are appropriately handled by neural network models. Moreover, in order to bear in mind several levels of autoregression and seasonalities, four lagged prices were taken into account with the following delays: one day, two days, one week and two weeks. It is important to bear in mind that the expected demand, wind and solar generation have been also used in the fundamental model, and these affect the values of other unit generation levels (e.g. coal and CCGT) via the demand balance constraint. However, neural network models take into account linear and non-linear trends and statistical patterns between electricity prices and the aforementioned “expected” variables. Therefore, even if the same data was used in both models, these were not handled and treated alike.

The neural network model configuration is shown in Fig. 1, which is comprised of a hidden layer and an output layer. According to Ref. [37], experience shows that one hidden layer is suitable for most applications. The hyperbolic tangent sigmoid was utilised as the activation function of the hidden layer’s neurons, whereas a pure linear transfer function was selected for the output layer.

The neural network was trained as per the standard Levenberg-Marquardt algorithm, which is one of the most popular neural network training methods utilised in electricity price forecasting applications, such as [8] and [38]. Therefore, different numbers of neurons were utilised on the hidden layer in order to test different levels of complexity and avoid overfitting.

The number of neurons is chosen as soon as the training algorithm reaches its performance optimality conditions by computing the Mean

Squared Error (MSE) on the validation period (a time span that is placed between training and test/forecast periods). The network with the number of neurons on its hidden layer that yields the lower MSE was then utilised to ascertain the electricity market price for the forecast period. The MSE is calculated as per the following formula with the conventional notation ($\hat{Y}_i$ represent the forecasted values for a certain period of N hours, whereas Y_i are the real values pertaining to the same period):

$$MSE = \frac{1}{N}\sum_{i=1}^{N} (\hat{Y}_i - Y_i)^2 \tag{1}$$

Moreover, in order to take into account the variability and the randomness of the initialisation of the neural network's weights at the start of the training process, a high number of replications of the described neural network procedure were carried out. This is also done bearing in mind that neural networks are, to some extent, prone to getting stuck on local minima on the training process, and thus the global optimum may not be ascertained. The final step involves calculating the mean of the obtained forecasts in every replication in order to yield the forecast of this work's proposed hybrid model.

However, the training data, which consists of 10 variables, were modified with the intention of increasing efficiency as well as reducing overfitting occurrences. The training, validation and forecasting periods have been organised according to the timeline of Fig. 2.

The training set is comprised of three time spans of d_1 days each: T_1, T_2 and T_3. The last one is placed right before the validation period V and contains the most updated information for the neural network to train on. This information reinforced by the data pertaining to T_1, which occurs one year before T_3. Furthermore, T_2 contains a possible evolution of electricity prices and all related variables, which is what happened a year before the forecasting period. This training set arrangement is more efficient and better reflects the behaviour of electricity prices on the forecasting period F than utilising three times d_1 days immediately prior to the validation period V. Furthermore, both V and F periods are set to be of the same duration, which is of d_2 days (i.e. the forecasting horizon).

Not only the arrangement of the input data was modified, but also a test was carried out in order to assess variable importance. Additionally, it would be useful to increase the parsimoniousness of the model, i.e. reducing model complexity as well as increasing predictive accuracy, and thus reducing possible overfitting occurrences.

To this end, the variables were tested in a backward-elimination manner, i.e. evaluating all ten factors at once and discarding one by one the most noisy and redundant. However, this is more straightforward in linear regression cases than in neural network applications. Therefore, the performance of the neural network model was assessed for several combinations of variables and numbers of neurons in the hidden layer (comparing its validation set MSE) using a backward-elimination procedure until only one variable was considered.

First of all, the neural network forecast MSE was calculated considering all 10 input variables. Then, the process was repeated for every combination of 9 input variables and the resulting MSE was compared to the one obtained with 10 variables. The variable that was not considered in the combination that yielded the greatest reduction in MSE was considered discarded. This method was then replicated for 8 input variables, considering the lowest MSE obtained for 9 input variables, and so on. If the MSE error could not be further reduced by removing one additional variable, the test reached its end and thus all remaining variables were not discarded. This procedure was carried out several times for numerous days. As a result, the variables that have been mostly discarded were: two-day lagged electricity prices and the Saturday dummy variable.

The estimated price from the fundamental model has not been discarded, which suggests that the underlying information within these prices is useful to the hybrid model (e.g. coal costs, CO_2 emission allowances, maintenance schedules). Moreover, elastic nets were also used for the same purpose, although this method is more appropriate for linear regression contexts, and the results were similar for a certain tolerance level (elastic net parameters). However, the authors preferred the backward-elimination procedure because it is a non-parametric approach and thus does not require further studies in order to ascertain additional information.

Therefore, the volume of the input data was reduced by 20% and, as a result, the model's performance was enhanced in terms of runtime and forecasting accuracy. The forecasting performance is tested as per some of the most used error metrics in the literature, e.g. Ref. [13], which are: Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), and Root-Mean-Square Error (RMSE). These error measures are computed as follows:

$$MAPE = \frac{100}{N}\sum_{i=1}^{N} \left|\frac{\hat{Y}_i - Y_i}{Y_i}\right| \tag{2}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N} |\hat{Y}_i - Y_i| \tag{3}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (\hat{Y}_i - Y_i)^2} \tag{4}$$

However, it is important to bear in mind that prices in the Iberian power exchange may go to zero, and thus may result in infinite MAPE. Nevertheless, this has not happened throughout all hours in the considered case studies. Furthermore, it is also common in the literature, e.g. Refs. [35,39], to provide statistically significant conclusions regarding forecasting performance comparisons via the Diebold-Mariano (DM) test [40], and thus it has also been applied to this work's case studies. A 10% significance level has been considered, an absolute error difference as the loss differential series, and a two-sided perspective, i.e. testing for both out- and underperformance.

Although the main objective of this work is focused on reducing the error of the mean of the forecasts, the proposed model's statistical performance has been evaluated so as to provide some insight related to risk analysis and worst case scenario evaluation. Therefore two different methods were carried out related to the percentiles of the forecasts of the proposed model, denoted by $\hat{Y}_{ia}$, with $a = 1, 2, \ldots, 99$. Firstly, the percentage of times that the percentile of the forecast is above the real value of the electricity price (i.e. Y_i) has been measured. This measure will be referred to as exceedance rate for the remainder of the paper. Ideally, given a percentile forecast $\hat{Y}_{ia}$, its exceedance rate

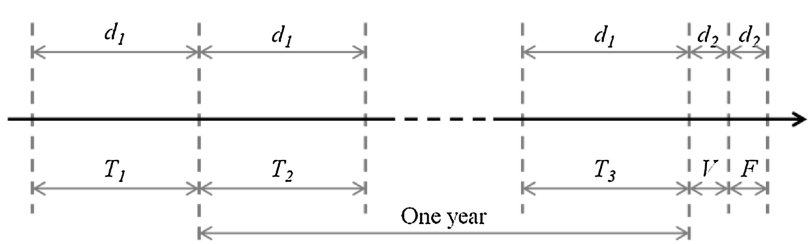

Fig. 2. Training, validation and test/forecast periods arrangement.

should be of %. Secondly, the proposed model was tested the pinball loss function (PLF) as done in several works in the current literature that focus on probabilistic analyses, such as Ref. [41]. The PLF for a certain hour is calculated as per the following equation:

$$PLF(a)_i = \begin{cases} (1 - \frac{a}{100})\cdot(\hat{Y}_{ia} - Y_i) & if\ \hat{Y}_{ia} > Y_i \\ (\frac{a}{100})\cdot(Y_i - \hat{Y}_{ia}) & if\ Y_i > \hat{Y}_{ia} \end{cases} \tag{5}$$

Lower values of the PLF score indicate that the forecasts are statistically superior and better reflect the probability of the occurrence of the price value associated with the target percentile.

3. Case studies, results and discussion

This section is composed of four parts. On the first subsection, seven case studies are presented. On the second subsection, the forecasts yielded by the proposed forecasting model are analysed whereas on the third subsection the resulting forecasts are compared with 5 benchmarks. Moreover, in order to further validate this work's proposed model, the same experiments have been conducted in a much broader case study, whose results are given on the last subsection.

3.1. Case study description and selection

Several periods in late 2016 have been studied for both forecasting horizons (one day and one week). Nevertheless, for the sake of simplicity and clarity, only the most representative cases that are characterised by different types of complexities are detailed in this section (shown on Table 1), all of which pertain to the period depicted in Fig. 3.

The most outstanding period is 21/Nov/2016, when prices collapsed due to unusually high wind generation in the Iberian power system, reaching 10.88€/MWh at early morning hours. Moreover, this day presents a range of 56.61€/MWh and a standard deviation of 15.07€/MWh. These values are significantly higher than those of its adjacent days, which present a range of 30.12€/MWh at most and a maximum standard deviation of 9.48€/MWh. Therefore, it would be interesting to analyse the models' forecasting performance on such a day (case C_2).

Furthermore, there is an apparent price level difference between November and December, which implies a slight market structural change. Moreover, coal plants were slightly less available in December than in November, whereas demand levels also increased in December [36]. For these reasons, a day and a week that do not present uncommon behaviours have been selected for both months in order to compare both market circumstances (cases C_1, C_3, C_5 and C_7).

Moreover, the most erratic weekly period in Fig. 3 takes place between 05/Dec/2016 and 11/Dec/2016, which contains two Spanish National holidays on the 6th and 8th of December (case C_6). Due to these holidays, electricity prices are lower than on their adjacent days.

Additionally, another uncommon day included in Fig. 3 is 26/Dec/2016 (case C_4), which is a Monday. However, the usual Christmas Day holiday happened on Sunday and thus the studied day was a holiday on most areas of Spain, so it cannot be considered a normal business day. These seven case studies put the proposed forecasting model to the test under diverse circumstances and challenges, all of which are analysed and discussed in the following subsections.

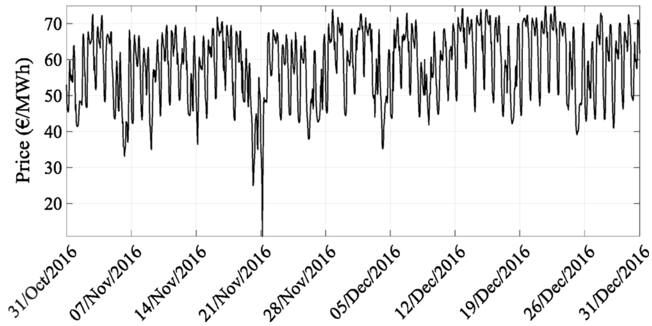

Fig. 3. Iberian electricity market prices for late 2016.

However, given that these specific cases may provide evidence as to how the considered forecasting models may perform under specific circumstances, the corresponding results cannot be adequately generalised. Therefore, in order to provide more statistically significant results and further proof as to how these models perform, the entire year 2017 was also used as a case study, whose results are located in subsection 3.4.

3.2. Proposed hybrid price forecasting model

First of all, the cost-production optimisation model has been run for the training and forecasting periods with hourly precision. In this test, regular circumstances in the Iberian power system were assumed regarding weather conditions (wind generation, hydro inflows, etc.). As a result, the estimated electricity market price was determined and used as an additional input to the neural network model.

The neural network model was run for seven different market circumstances in order to assess its adequacy for the short term. This includes horizons of one day and one week with hourly precision. Once the day or week to forecast was set, all periods according to the timeline of Fig. 2 can be set. In all cases, a training set arranged as per Fig. 2 was used with d_1 equal to 30 days, thus a total of 90 days were used as training data. Neural networks with this training set were trained considering different numbers of neurons (10 to 60 with a step of 5). The neural network with the lowest MSE on the validation set was later used to forecast the electricity market price on the test/forecast period. This procedure (from neural network training to forecasting) was carried out 300 times, whose mean was used as the resulting forecast of the proposed electricity market price forecasting model.

For the weekly cases, the authors previously tested this methodology by both using a standard 168-h horizon and a rolling window of a 24-h horizon, i.e. forecasting day by day up to one week. However, by using a rolling-window method on a neural network model, the authors

Table 1
Case study training, validation and forecast periods.

Case study	Training periods		Validation period *V*	Forecasting period *F*
	$T_1 \cup T_2$	T_3		
C_1	15/Oct/2015–13/Dec/2015	15/Oct/2016–13/Nov/2016	14/Nov/2016	15/Nov/2016
C_2	21/Oct/2015–19/Dec/2015	21/Oct/2016–19/Nov/2016	20/Nov/2016	21/Nov/2016
C_3	10/Nov/2015–09/Jan/2016	10/Nov/2016–10/Dec/2016	11/Dec/2016	12/Dec/2016
C_4	24/Nov/2015–23/Jan/2016	24/Nov/2016–24/Dec/2016	25/Dec/2016	26/Dec/2016
C_5	08/Oct/2015–06/Dec/2015	08/Oct/2016–06/Nov/2016	07/Nov/2016–13/Nov/2016	14/Nov/2016–20/Nov/2016
C_6	29/Oct/2015–26/Dec/2015	29/Oct/2016–26/Nov/2016	27/Nov/2016–04/Dec/2016	05/Dec/2016–11/Dec/2016
C_7	05/Nov/2015–03/Jan/2016	05/Nov/2016–04/Dec/2016	05/Dec/2016–11/Dec/2016	12/Dec/2016–18/Dec/2016

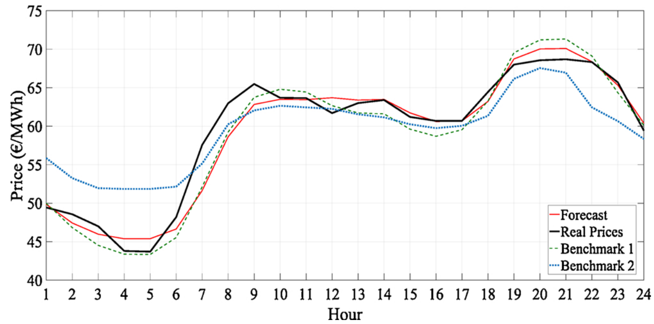

Fig. 4. Electricity price forecast for 15/Nov/2016 (C_1).

found that the forecasts given in every simulation had a considerably higher volatility than those yielded by the neural network model with a 168-h horizon, as well as a generally higher MAPE. Therefore, the rolling-window method was not utilised in these cases. However, one-day lagged prices were not included in the one-week forecasts, because, in reality, the price of the previous day becomes unknown if forecasting further than one day. Nevertheless, the estimated price from the fundamental model may contain, to some extent, the information of one-day lagged prices due to the chronological constraints (e.g. unit commitment and hydro reserve balance) that set links between current prices and other factors in the past.

The first case study (C_1) is displayed in Fig. 4, which shows the forecast for 15/Nov/2016 (Tuesday). The daily trend that electricity prices usually exhibit is successfully mimicked and the model yields a 2.179% MAPE, which mostly corresponds to the early morning hours, whereas the late hours are considerably accurate.

The dashed and dotted lines of Fig. 4 represent the forecasts of both components of the proposed hybrid model on their own, i.e. only the neural network model without the additional estimated price from the cost-minimisation model (Benchmark 1 or BM_1), and also this estimated price on its own (Benchmark 2 or BM_2). By analysing and comparing these results, the benefits of the hybridisation of both methods can be checked and verified.

The estimated price of BM_2 clearly lacks intraday dynamics and thus yields a lower accuracy (5.675% MAPE). Nevertheless, this forecast is somewhat centred on the average price level, which is of vital importance for the developed hybrid model. The daily behaviour exhibited by BM_1 better resembles electricity price patterns mainly thanks to its adaptability for non-linear trends, although its accuracy is considerably lower on the afternoon and evening (2.856% MAPE). The combination of the advantages of both models yields a suitable adaptive behaviour, and thus increases the accuracy of the hybrid model forecast.

Regarding the one-week case studies, the forecast for the week of 05/Dec/2016–11/Dec/2016 is displayed in Fig. 5. In this case study, the resulting MAPE is of 5.878%. Once again, it can be seen in Fig. 5 that the estimated price from the fundamental model (BM_2) fails to follow the intraday pattern (8.709% MAPE). The neural network model on its own (BM_1) shows an adequate performance (6.136% MAPE), although it seems to yield considerably lower values on the early hours of Thursday to Sunday, which may be caused by a slight underperformance of the neural network model as the forecasting horizon increases.

This performance decrease is somewhat diminished by the estimated price of BM_2, which provides the equilibrium price level even at longer horizons and thus results in a more accurate forecast for the hybrid model. The benefits of the proposed hybrid model are experienced yet again, which strongly supports the statement that the combination of both models' advantages is highly valuable. Furthermore, the fundamental model's ability to incorporate the effects of the reduction of availability in the system coal power plants provided a slight upward pressure on the hybrid model's price forecasts. This contribution of the fundamental model proved relevant and useful in this forecasting period.

The rest of the cases (C_2, C_3, C_4, C_5 and C_7) are discussed on the following subsection, including the performance of the other benchmark forecasting models.

3.3. Comparison with other forecasting models

The performance of this work's proposed hybrid forecasting model has been compared with five other electricity price forecasting models, two of which represent the split versions of the hybrid model (BM_1 and BM_2). The third benchmark (BM_3) is a slight modification of a linear regression model that was proposed in Ref. [42] and most recently applied to electricity price forecasting in Ref. [35]. This linear regression model can be represented as per the following equations in order to calculate the log-price $p_{d,h}$ at day and hour :

$$p_{d,h} = \beta_{h,1}p_{d-1,h} + \beta_{h,2}p_{d-2,h} + \beta_{h,3}p_{d-7,h} + \beta_{h,4}p_{d-1}^{min} + \beta_{h,5}z_{d,h} + \beta_{h,6}D_{Sat} + \beta_{h,7}D_{Sun} + \beta_{h,8}D_{Mon} + \varepsilon_{d,h} \quad (6)$$

$$p_{d,h} = log\,(P_{d,h}) - \frac{1}{T}\sum_{t=1}^{T} log\,(P_{d,h}) \quad (7)$$

The betas are the regressor coefficients, which respectively represent lagged log-prices (one, two and seven days), the minimum log-price of the 24 h in day minus one, the expected demand and three dummy variables indicating if day is Saturday, Sunday or Monday.

Furthermore, as mentioned before, a slight modification was carried out, which pertains to the logarithmic transform of Eq. (7), where refers

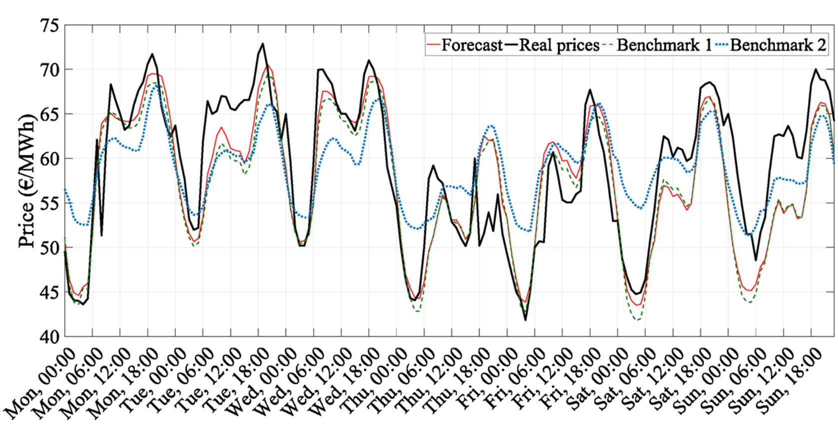

Fig. 5. Electricity price forecast from Mon, 05/Dec/2016 to Sun, 11/Dec/2016 (C_6).

to the training period. The mirror-log transform, recently applied to electricity price forecasting in Ref. [43], was applied due to the possibility of prices equal to zero in the Iberian power system, which is represented in the following equations:

$$n_{d,h} = \frac{(P_{d,h} - \mu_T)}{\sigma_T} \tag{8}$$

$$p_{d,h} = sgn(n_{d,h})\left[log\left(n_{d,h} + \frac{1}{c}\right) + log(c)\right] \tag{9}$$

First of all, according to Eq. (8), the prices were normalised by subtracting their mean in the training period (μ_T) and dividing by their standard deviation in the training period (σ_T). The parameter was set to 1/3 as done in Ref. [43].

The fourth benchmark (BM_4), based on ARIMA models, is more established than the previous two and well recognised. These models have been widely used in electricity price forecasting, including the Iberian electricity market [44]. In this case, the model consists of a transfer function with SARIMA noise, which has been developed according to the methodologies presented in the works of Refs. [45,46]. Electricity prices were stabilised (variance) by means of the Box–Cox transformation [47]. The BIC value of the fitted models was used as model selection criterion. The obtained SARIMA noise presents the following parameters with the standard notation: $SARIMA(1,0,0)_{168}(1,0,2)_{24}(1,0,0)_1$. The expected demand was used as an exogenous variable in this model that can be therefore also referred to as a SARIMAX model.

The last benchmark (BM_5) is a simple naïve approach in which the real electricity prices from the previous week are directly taken as the forecast:

$$P_{d,h} = P_{d-7,h} \tag{10}$$

The MAPE, MAE and RMSE of these benchmarks along with those yielded by the proposed electricity price forecasting model on the seven case studies are displayed in Tables 2–4, respectively. First of all, it is worth noting that the proposed model outperforms the others in all cases except C_2, in which the pure neural network model yields a higher accuracy. This is mostly due to the fact that, in early morning hours, BM_1's forecast is closer to the real value than the proposed model's forecast. Furthermore, in a case in which prices collapse to such a low value (10.88€/MWh), this difference is more apparent and noticeable.

This trend also happens in the more common days pertaining to cases C_1 (see Fig. 4) and C_3, although the proposed hybrid model's forecast in the rest of the hours of the day makes up for it more than enough, yielding a higher overall accuracy in terms of MAPE. This fact also confirms that the fundamental contribution enhances forecasting performance on late morning hours up to midnight, whereas on early morning hours it yields a reduced accuracy and this reduction is further heightened on uncommon, low-price situations such as C_2.

Moreover, the one-week cases (C_5, C_6 and C_7) comparison between the proposed model and BM_1 show similar accuracies. However, on early December (case C_6), when the overall price levels are beginning to increase, the difference is higher. This may suggest that whenever such a structural market evolution is underway, fundamental information should be taken into consideration.

Furthermore, case C_6 includes two Spanish National holidays, and thus it may imply that the proposed model is also the most proficient at forecasting prices on non-business days. The same conclusion can be reached from the results of case C_4, which is also a non-business day. This may also indicate that by considering the estimated price from the fundamental model, the bias effect from the previous week is lessened.

The results shown by the MAPE, MAE and RMSE values have shown some differences between the proposed model and the five benchmarks. In order to adequately analyse the statistical differences between these forecasting models, a DM with 10% significance level was carried out, whose results are displayed in Table 5.

It can be seen that the proposed model's forecasts generally yield statistically lower errors than those of its competitors. The only exception in which its forecasts are statistically underperforming is in case C_2, where BM_1 has proved to be more accurate according to the error metrics of the previous tables. However, three cases, C_1, C_4 and C_6, show significant differences in favour of the proposed model when compared to BM_1.

Table 2
Comparison of the proposed forecasting model with five benchmarks in terms of MAPE (%).

Model	C_1	C_2	C_3	C_4	C_5	C_6	C_7
Proposed	**2.179**	15.96	**3.726**	**3.242**	**6.146**	**5.878**	**4.119**
BM_1	2.856	**13.79**	4.075	3.825	6.172	6.136	4.166
BM_2	5.675	26.73	6.812	10.83	11.23	8.709	8.135
BM_3	7.522	27.84	5.421	13.63	11.95	10.84	7.746
BM_4	9.280	22.48	6.238	6.657	9.894	13.67	7.224
BM_5	7.595	31.91	8.544	19.57	12.84	14.11	10.85

Table 3
Comparison of the proposed forecasting model with five benchmarks in terms of MAE (€/MWh).

Model	C_1	C_2	C_3	C_4	C_5	C_6	C_7
Proposed	**1.250**	5.294	**2.414**	**1.767**	**3.062**	**3.462**	2.609
BM_1	1.690	**4.444**	2.715	2.045	3.123	3.634	**2.542**
BM_2	3.062	7.575	3.414	3.374	5.496	4.893	4.890
BM_3	4.646	7.935	3.422	7.226	5.811	6.173	4.750
BM_4	5.695	7.365	4.068	3.397	5.043	8.367	4.641
BM_5	4.727	9.658	5.378	10.51	6.456	7.870	6.623

Table 4
Comparison of the proposed forecasting model with five benchmarks in terms of RMSE (€/MWh).

Model	C_1	C_2	C_3	C_4	C_5	C_6	C_7
Proposed	**1.861**	7.166	**2.622**	**2.474**	**3.897**	**4.466**	3.213
BM_1	2.035	**5.893**	3.063	2.718	3.930	4.621	**3.149**
BM_2	3.817	10.85	4.018	3.771	6.986	5.619	5.534
BM_3	5.249	9.938	4.697	8.241	7.151	7.617	5.655
BM_4	6.202	8.105	4.407	3.743	6.163	9.704	5.552
BM_5	6.043	11.24	6.876	12.41	8.333	9.780	8.453

3.4. Results for the entire year 2017

This work's proposed model, as well as the five benchmarks, has also been tested for the entire year 2017 for both one-day and one-week forecasting horizons. The MAPE, MAE and RMSE forecasting error results are shown in Tables 6–8 respectively.

It is important to take into account that, during January 2017, the

Table 5
Results of the Diebold-Mariano test across all case studies.

Model comparison	C_1	C_2	C_3	C_4	C_5	C_6	C_7
Proposed vs. BM_1	1[a]	−1[b]	0[c]	1	0	1	0
Proposed vs. BM_2	1	0	1	1	1	1	1
Proposed vs. BM_3	1	1	1	1	1	1	1
Proposed vs. BM_4	1	1	1	1	1	1	1
Proposed vs. BM_5	1	1	1	1	1	1	1

[a] A value of 1 indicates significant outperformance of the proposed model forecasts.
[b] A value of −1 indicates significant underperformance of the proposed model forecasts.
[c] A value of 0 indicates no significant difference between model forecasts.

Table 6
MAPE comparison of the proposed model with five benchmarks for the year 2017 (%).

Model	One-day forecasting horizon					One-week forecasting horizon				
	Winter	Spring	Summer	Autumn	Average	Winter	Spring	Summer	Autumn	Average
Proposed	**12.83**	**8.840**	5.016	**6.764**	**8.341**	**17.08**	8.330	**5.120**	8.000	**9.633**
BM_1	12.97	9.018	**4.977**	6.815	8.424	17.31	**8.202**	5.305	**7.919**	9.684
BM_2	26.67	21.54	12.39	19.10	19.89	26.38	21.94	12.13	19.18	19.91
BM_3	16.79	13.58	7.153	10.51	11.99	17.69	13.24	7.050	10.56	12.13
BM_4	15.06	9.293	5.097	7.654	9.248	17.64	10.36	5.915	8.662	10.65
BM_5	25.93	17.55	9.343	12.82	16.37	26.17	17.02	9.462	12.91	16.39

Iberian power market was affected by an uncommon mix of events: unusually low temperatures, low renewable generation, and high natural gas prices; and thus the forecasting error on all models is significantly higher than in the other seasons of the year. Furthermore, January's monthly average price was of 71.49€/MWh (18.6% and 38.2% higher than that of the previous and the following month respectively) and its standard deviation was of 14.26€/MWh (52.4% and 23.9% higher than that of the previous and the following month respectively).

Although the overall mean of the proposed model's forecasting errors is the lowest during the entire year 2017, it slightly underperforms during some seasons. For example, for the one-day forecasting horizon cases, the proposed model's performance during the early morning hours of the day in summer is lower, which coincides with the aforementioned trend described in the previous subsection.

In order to check if there are any significant differences in predictive performances, a DM test has also been carried out for each model throughout the year 2017, whose results are displayed in Table 9 (with the same notation as in Table 5). The proposed model is not significantly outperformed in any of the cases and there seems to be a general outperformance in the entire year 2017 for the one-week horizon cases. Taking into account that neural network forecast accuracy is reduced for longer horizons, it can be deduced from this result that the contribution of the equilibrium price level that is provided by the estimated fundamental price is more notable, which coincides with case C_6's results that are explained in subsection 3.2. This is also why, in numerous works in the current literature (such as Refs. [23,24]), this effect has proven useful for longer forecasting windows (i.e. medium-term horizons). However, the results repeatedly suggest that the proposed fundamental-econometric hybrid model's performance needs to be improved in the early morning hours of the day.

Furthermore, it is also of interest to verify the statistical features of the forecasts that have been given by the proposed model, which can be done by analysing the percentiles of the forecasts. The exceedance rate of the percentile forecasts of the proposed model and BM_1 is displayed in Table 10. As mentioned before, the ideal exceedance rate for the percentile 1, 5, 95 and 99 forecasts are of 1%, 5%, 95% and 99% respectively.

First of all, it is natural that the longer horizon forecasts deviate more from these ideal values. The proposed model is generally closer to the ideal values than BM_1, especially the percentile 1 and percentile 5 values. This means that the proposed model's forecasts are generally more suitable for risk analyses such as extreme or worst case scenario evaluation. However, no notable improvement was achieved for the percentile 95 and percentile 99 cases. Furthermore, for both forecasting horizons, the percentile 95 and percentile 99 exceedance rates seem to be farther from their ideal values, which may be an indication that the probability distribution of the forecasts presents a positive skew and therefore is not symmetrical. This can be verified by calculating the PLF as per Eq. (5). of the proposed model's forecasts, whose results can be seen in Table 11.

The PLF results suggest that the proposed hybrid model yields overall superior probabilistic forecasts for percentiles 1 and 5 as opposed to percentiles 99 and 95 respectively, which indicates that the forecasts do not capture the probability of the occurrence of extremely high prices as well as the occurrence of extremely low prices. Therefore, this may call for a peak or extreme value detection procedure if this imbalance is to be solved.

4. Conclusions and future work

The novel methodology that has been proposed in this work is based on a hybrid model which consists of a cost-production optimisation model and a neural network model. Both models have been linked by using the cost-production optimisation model's estimated price as an additional input to the neural network model. Furthermore, the input data on both components of the proposed hybrid model were rearranged and modified in order to decrease computational burden and therefore increase efficiency, as well as reduce runtime and overfitting occurrences on the neural network model.

The proposed hybrid model has shown adequate performance in seven case studies, all of which have presented diverse circumstances and challenges. The benchmark models were outperformed by the proposed model in most case studies, especially the estimated price from the cost-production optimisation model, the linear regression model of Ref. [42], the SARIMAX model and the simple naïve approach that utilises the previous week's electricity prices as the forecast.

Furthermore, it can be concluded that the proposed hybrid forecasting model's accuracy is generally increased by the effect of the estimated price from the fundamental model. In addition, the non-linear

Table 7
MAE comparison of the proposed model with five benchmarks for the year 2017 (€/MWh).

Model	One-day forecasting horizon					One-week forecasting horizon				
	Winter	Spring	Summer	Autumn	Average	Winter	Spring	Summer	Autumn	Average
Proposed	**5.137**	3.068	2.359	**3.331**	**3.465**	7.369	**3.247**	**2.407**	4.044	**4.267**
BM_1	5.178	**3.059**	**2.335**	3.383	3.480	7.486	3.257	2.491	**4.028**	4.315
BM_2	14.96	9.413	5.949	11.03	10.31	14.68	9.667	5.797	11.12	10.32
BM_3	6.838	4.765	3.262	5.066	4.972	**7.213**	4.652	3.198	5.116	5.045
BM_4	8.113	4.150	2.473	4.454	4.780	9.382	4.694	2.844	5.051	5.493
BM_5	10.53	6.225	4.266	6.387	6.828	10.54	6.043	4.315	6.434	6.833

Table 8
RMSE comparison of the proposed model with five benchmarks for the year 2017 (€/MWh).

Model	One-day forecasting horizon					One-week forecasting horizon				
	Winter	Spring	Summer	Autumn	Average	Winter	Spring	Summer	Autumn	Average
Proposed	**5.921**	3.658	2.840	**4.003**	**4.096**	**8.927**	**4.074**	**3.065**	5.113	**5.295**
BM_1	5.953	**3.638**	**2.822**	4.089	4.115	9.046	4.115	3.164	**5.083**	5.352
BM_2	16.59	10.97	6.985	12.39	11.70	16.84	11.87	7.114	12.87	12.17
BM_3	7.809	5.552	3.885	6.055	5.814	9.010	5.944	4.105	6.612	6.418
BM_4	10.84	5.585	4.531	4.959	6.460	12.76	6.414	5.341	5.639	7.537
BM_5	11.48	7.092	5.030	7.567	7.773	12.32	7.775	5.458	8.298	8.464

Table 9
Results of the Diebold-Mariano test for the entire year 2017.

Model	One-day forecasting horizon					One-week forecasting horizon				
	Win.	Spr.	Sum.	Aut.	Avg.	Win.	Spr.	Sum.	Aut.	Avg.
Proposed vs. BM_1	0	0	0	1	0	1	0	1	0	1
Proposed vs. BM_2	1	1	1	1	1	1	1	1	1	1
Proposed vs. BM_3	1	1	1	1	1	0	1	1	1	1
Proposed vs. BM_4	1	1	0	1	1	1	1	1	1	1
Proposed vs. BM_5	1	1	1	1	1	1	1	1	1	1

Table 10
Exceedance rate of the percentile forecasts for the entire year 2017 (%).

Percentile	Model	One-day forecasting horizon					One-week forecasting horizon				
		Win.	Spr.	Sum.	Aut.	Avg.	Win.	Spr.	Sum.	Aut.	Avg.
P1	Proposed	**4.398**	**4.076**	**5.978**	5.952	**5.103**	6.227	**4.487**	**8.562**	**10.58**	**7.463**
	BM_1	4.583	6.024	7.428	**5.815**	5.970	**6.181**	5.037	10.12	11.86	8.299
P5	Proposed	**10.51**	**7.926**	**11.01**	12.13	**10.39**	**12,73**	**8,379**	**15,11**	**17,54**	**13,44**
	BM_1	11.06	9.149	12.55	**11.81**	11.14	13,74	9,432	15,84	17,99	14,25
P95	Proposed	88.94	**85.73**	**85.51**	**81.87**	**85.50**	**79,35**	**80,08**	**75,05**	**69,60**	**76,02**
	BM_1	**89.03**	84.92	84.65	81.04	84.90	77,66	80,91	71,61	69,51	74,92
P99	Proposed	94.68	93.39	**93.80**	**92.31**	**93.54**	**87.18**	89.88	**84.62**	80.45	**85.53**
	BM_1	**95.51**	**94.34**	91.39	91.07	93.07	86.49	**90.43**	80.91	**83.47**	85.32

Table 11
Pinball loss function score of this work's proposed model for the entire year 2017.

Percentile	One-day forecasting horizon					One-week forecasting horizon				
	Win.	Spr.	Sum.	Aut.	Avg.	Win.	Spr.	Sum.	Aut.	Avg.
P1	**0.259**	**0.199**	0.180	0.239	**0.219**	**0.415**	**0.162**	**0.162**	**0.341**	**0.270**
P99	0.324	0.212	**0.172**	**0.217**	0.231	0.711	0.287	0.399	0.516	0.478
P5	**0.813**	**0.503**	**0.408**	0.590	**0.577**	**1.141**	**0.455**	**0.391**	**0.748**	**0.684**
P95	0.816	0.515	0.428	**0.588**	0.586	1.410	0.656	0.672	1.031	0.942

patterns in electricity prices have been adequately dealt with by the neural network model. Moreover, the combination of the longer-term price level yielded by the fundamental model and the intraday pattern given by the econometric model has unquestionably proven to be advantageous, especially on uncommon market situations, such as holidays or increasing unit unavailableness.

Additionally, the case study of the entire year 2017 showed an overall lower forecasting error from this work's proposed forecasting model. This advantage is more notable on the one-week forecasting horizon results, which further indicates that the equilibrium price level of the fundamental model enhances predictive accuracy even if the forecasting horizon is longer. Moreover, even though this work's proposed model yields overall superior probabilistic forecasts, there seems to be room for improvement in order to obtain percentile forecasts that are closer to the ideal values regarding exceedance rates.

However, the results suggest that, on early morning hours, a combination technique with another pure statistical model, or a regime-switching model within a hybrid framework, may enhance the resulting model's accuracy. Furthermore, other variables from the fundamental model may be utilised as additional information for the econometric model, such as the unit generation levels for the different thermal technologies. Nevertheless, considering more input variables in this application may call for a more suitable sensitivity analysis or variable selection procedure than the one carried out in this work, although developing a computationally efficient method may result in a highly challenging task due to the high level of complexity of the neural network training algorithms.

Appendix A

The fundamental model that has been used in this work's proposed methodology consists of a traditional market equilibrium model in which the costs of each market agent are minimised. In this model, perfect competition has been assumed, and thus total system costs are simultaneously minimised in its objective function. Taking into account that the main decision variables are the production quantities of each generation unit q_i, the following equation represents the model's objective function:

$$\min_{q_i} \sum_i c_i \tag{11}$$

The term c_i is the cost function of generation unit i, which mainly consists of the following costs: fuel, start-up, CO_2 emission and maintenance costs. Apart from the corresponding technical and physical constraints (e.g. maximum/minimum power, start-up and shut-down periods, etc.) of the system's generation units, the most important element of this optimisation problem is the demand vs. generation balance equation (for every hour t):

$$\textstyle\sum_i q_{i,t} = D_t \quad : \lambda_t \ \forall \ t \tag{12}$$

Each unit's production quantities, i.e. $q_{i,t}$, are mainly limited by this constraint, and therefore its dual variable can be considered as the system's marginal price λ_t, which represents the market clearing price that is utilised for the econometric component of this work's proposed hybrid model. Furthermore, it is worth noting that interconnections with adjacent markets, such as France, have been taken into account. Given this model's purpose, the lack of non-linear terms (e.g. the conjectured-price response quadratic term of Ref. [23]) and the unit coupling conditions (i.e. start-up and shut-down bounds), the nature of the optimisation model is a relaxed mixed-integer problem (RMIP), which has been run with the CPLEX solver (version 12.5.1.0).

The parameters, i.e. input data, of this optimisation model are listed below:

- $\underline{P_i}$: minimum power output of unit i
- $\overline{P_i}$: maximum power output of unit i
- $U_{i,0}$: initial unit commitment state of (thermal) unit i (active or inactive)
- $FP_{i,t}$: fuel cost per unit of volume produced for (thermal) unit i and hour t
- S_i: cost per start-up operation for (thermal) unit i
- A_i: tonnes of CO_2 emitted per unit of volume produced for (thermal) unit i
- B: penalty per tonne of CO_2 emitted
- M_i: maintenance costs per unit of volume produced for (thermal) unit i
- $I_{i,t}$: expected hydro inflow for (hydro) unit i and hour t
- $R_{i,0}$: initial hydro reservoir level of (hydro) unit i
- $W_{i,t}$: expected wind generation for (wind power) unit i at hour t
- D_t: expected system demand at hour t

All generation units considered may belong to one of these sets:

- T: thermal power units (nuclear, CCGT, OCGT, and coal)
- H: hydro power units
- W: wind power units

Solar generation units are not physically considered and their expected production is a known parameter that is added to the term on the left-hand side of Eq. (12). Furthermore, as mentioned on the beginning of this manuscript's Section 2, the generation units that share similar technical features were aggregated into larger units, therefore reducing the number of them. Although the operation and behaviour of the market is thus not very accurately modelled, the main objective of this model is to provide an output (i.e. market clearing prices) that reflects changes regarding the market fundamentals. Moreover, the main computational statistics of this simplified optimisation model are also mentioned at Section 2.

The main variables of the optimisation model, apart from the system's marginal price, are listed below:

- $q_{i,t}$: production of generation unit i at hour t
- c_i: total costs of unit i
- cf_i: fuel costs of (thermal) unit i
- cs_i: start-up costs of (thermal) unit i
- cc_i: CO_2 emission costs of unit i
- cm_i: maintenance costs of (thermal or wind power) unit i
- $u_{i,t}$: integer variable that indicates if generation unit i at hour t is active or not
- $y_{i,t}$: binary variable that indicates if generation unit i at hour t is starting up
- $z_{i,t}$: binary variable that indicates if generation unit i at hour t is shutting down
- $r_{i,t}$: hydro reservoir level of (hydro) unit i at hour t
- $s_{i,t}$: energy spilled amount of (hydro or wind power) unit i at hour t
- $p_{i,t}$: pumped amount of (hydro) unit i at hour t

In addition to Eq. (12), the following set of equations represents this model's main constraints for every hour :

$$\underline{P_i} \leq q_{i,t} \leq \overline{P_i} \quad \forall \ i, t \tag{13}$$

$$y_{i,t} - z_{i,t} = u_{i,t} - u_{i,t-1} \quad \forall \ i, t > 1 \tag{14}$$

$$y_{i,1} - z_{i,1} = u_{i,1} - U_{i,0} \quad \forall\ i \tag{15}$$

$$c_i = cf_i + cs_i + cc_i + cm_i \quad \forall\ i \tag{16}$$

$$cf_i = \sum_t q_{i,t} \cdot FP_{i,t} \quad \forall\ i \in T \tag{17}$$

$$cs_i = \sum_t y_{i,t} \cdot S_i \quad \forall\ i \in T \tag{18}$$

$$cc_i = \sum_t q_{i,t} \cdot A_i \cdot B \quad \forall\ i \in T \tag{19}$$

$$cm_i = \sum_t q_{i,t} \cdot M_i \quad \forall\ i \in \{T, W\} \tag{20}$$

$$r_{i,t} = r_{i,t-1} + I_{i,t} - q_{i,t} - s_{i,t} + p_{i,t} \quad \forall\ i \in H, t > 1 \tag{21}$$

$$r_{i,1} = R_{i,0} + I_{i,1} - q_{i,1} - s_{i,1} + p_{i,1} \quad \forall\ i \in H \tag{22}$$

$$q_{i,t} = W_{i,t} - s_{i,t} \quad \forall\ i \in W, t \tag{23}$$

References

[1] A. Bello, J. Reneses, A. Muñoz, A. Delgadillo, Probabilistic forecasting of hourly electricity prices in the medium-term using spatial interpolation techniques, Int. J. Forecasting 32 (2016) 966–980, https://doi.org/10.1016/j.ijforecast.2015.06.002.

[2] R. Weron, Electricity price forecasting: a review of the state-of-the-art with a look into the future, Int. J. Forecasting 30 (2014) 1030–1081, https://doi.org/10.1016/j.ijforecast.2014.08.008.

[3] N. Bowden, J.E. Payne, Short term forecasting of electricity prices for MISO hubs: evidence from ARIMA-EGARCH models, Energy Econ. 30 (2008) 3186–3197, https://doi.org/10.1016/j.eneco.2008.06.003.

[4] R.C. Garcia, J. Contreras, M. van Akkeren, J.B.C. Garcia, A GARCH forecasting model to predict day-ahead electricity prices, IEEE Trans. Power Syst. 20 (2005) 867–874, https://doi.org/10.1109/TPWRS.2005.846044.

[5] C.W. Cheong, Modeling and forecasting crude oil markets using ARCH-type models, Energy Policy 37 (2009) 2346–2355, https://doi.org/10.1016/j.enpol.2009.02.026.

[6] Z. Tan, J. Zhang, J. Wang, J. Xu, Day-ahead electricity price forecasting using wavelet transform combined with ARIMA and GARCH models, Appl. Energy 87 (2010) 3606–3610, https://doi.org/10.1016/j.apenergy.2010.05.012.

[7] A. Cruz, A. Muñoz, J.L. Zamora, R. Espínola, The effect of wind generation and weekday on Spanish electricity spot price forecasting, Electr. Power Syst. Res. 81 (2011) 1924–1935, https://doi.org/10.1016/j.epsr.2011.06.002.

[8] J.P.S. Catalão, S.J.P.S. Mariano, V.M.F. Mendes, L.A.F.M. Ferreira, Short-term electricity prices forecasting in a competitive market: a neural network approach, Electr. Power Syst. Res. 77 (2007) 1297–1304, https://doi.org/10.1016/j.epsr.2006.09.022.

[9] N. Amjady, F. Keynia, Electricity market price spike analysis by a hybrid data model and feature selection technique, Electr. Power Syst. Res. 80 (2010) 318–327, https://doi.org/10.1016/j.epsr.2009.09.015.

[10] J.P.S. Catalão, H.M.I. Pousinho, V.M.F. Mendes, Hybrid wavelet-PSO-ANFIS approach for short-term electricity prices forecasting, IEEE Trans. Power Syst. 26 (2011) 137–144, https://doi.org/10.1109/TPWRS.2010.2049385.

[11] I.P. Panapakidis, A.S. Dagoumas, Day-ahead electricity price forecasting via the application of artificial neural network based models, Appl. Energy 172 (2016) 132–151, https://doi.org/10.1016/j.apenergy.2016.03.089.

[12] D. Keles, J. Scelle, F. Paraschiv, W. Fichtner, Extended forecast methods for day-ahead electricity spot prices applying artificial neural networks, Appl. Energy 162 (2016) 218–230, https://doi.org/10.1016/j.apenergy.2015.09.087.

[13] H.S. Sandhu, L. Fang, L. Guan, Forecasting day-ahead price spikes for the Ontario electricity market, Electr. Power Syst. Res. 141 (2016) 450–459, https://doi.org/10.1016/j.epsr.2016.08.005.

[14] N. Amjady, Day-ahead price forecasting of electricity markets by a new fuzzy neural network, IEEE Trans. Power Syst. 21 (2006) 887–896, https://doi.org/10.1109/TPWRS.2006.873409.

[15] P. Areekul, T. Senjyu, H. Toyama, A. Yona, A hybrid ARIMA and neural network model for short-term price forecasting in deregulated market, IEEE Trans. Power Syst. 25 (2010) 524–530, https://doi.org/10.1109/TPWRS.2009.2036488.

[16] L.D.S. Coelho, A.A.P. Santos, A RBF neural network model with GARCH errors: application to electricity price forecasting, Electr. Power Syst. Res. 81 (2011) 74–83, https://doi.org/10.1016/j.epsr.2010.07.015.

[17] M. Shafie-Khah, M.P. Moghaddam, M.K. Sheikh-El-Eslami, Price forecasting of day-ahead electricity markets using a hybrid forecast method, Energy Convers. Manag. 52 (2011) 2165–2169, https://doi.org/10.1016/j.enconman.2010.10.047.

[18] M. Ortiz, O. Ukar, F. Azevedo, A. Múgica, Price forecasting and validation in the Spanish electricity market using forecasts as input data, Int. J. Electr. Power Energy Syst. 77 (2016) 123–127, https://doi.org/10.1016/j.ijepes.2015.11.004.

[19] X. Yan, N.A. Chowdhury, Mid-term electricity market clearing price forecasting: a hybrid LSSVM and ARMAX approach, Int. J. Electr. Power Energy Syst. 53 (2013) 20–26, https://doi.org/10.1016/j.ijepes.2013.04.006.

[20] Z. Yang, L. Ce, L. Lian, Electricity price forecasting by a hybrid model, combining wavelet transform ARMA and kernel-based extreme learning machine methods, Appl. Energy 190 (2017) 291–305, https://doi.org/10.1016/j.apenergy.2016.12.130.

[21] S. Talari, M. Shafie-khah, G. Osório, F. Wang, A. Heidari, J.P.S. Catalão, Price forecasting of electricity markets in the presence of a high penetration of wind power generators, Sustainability 9 (2017) 2065, https://doi.org/10.3390/su9112065.

[22] M. Cerjan, I. Krzelj, M. Vidak, M. Delimar, A literature review with statistical analysis of electricity price forecasting methods, Eurocon, 2013 IEEE (2013) 756–763, https://doi.org/10.1109/EUROCON.2013.6625068.

[23] A. Bello, D.W. Bunn, J. Reneses, A. Munoz, Medium-term probabilistic forecasting of electricity prices: a hybrid approach, IEEE Trans. Power Syst. 32 (2016) 334–343, https://doi.org/10.1109/TPWRS.2016.2552983.

[24] A. Bello, D. Bunn, J. Reneses, A. Muñoz, Parametric density recalibration of a fundamental market model to forecast electricity prices, Energies 9 (2016) 959, https://doi.org/10.3390/en9110959.

[25] J. Nowotarski, R. Weron, Recent advances in electricity price forecasting: a review of probabilistic forecasting, Renew. Sustain. Energy Rev. 81 (2018) 1548–1568, https://doi.org/10.1016/j.rser.2017.05.234.

[26] V. González, J. Contreras, D.W. Bunn, Forecasting power prices using a hybrid fundamental-econometric model, IEEE Trans. Power Syst. 27 (2012) 363–372, https://doi.org/10.1109/TPWRS.2011.2167689.

[27] N.V. Karakatsani, D.W. Bunn, Forecasting electricity prices: the impact of fundamentals and time-varying coefficients, Int. J. Forecasting 24 (2008) 764–785, https://doi.org/10.1016/j.ijforecast.2008.09.008.

[28] R. Carmona, M. Coulon, A survey of commodity markets and structural models, Quantitative Energy Finance. Modeling, Pricing, and Hedging in Energy and Commodity Markets, (2014), pp. 41–83, https://doi.org/10.1007/978-1-4614-7248-3.

[29] J. Bai, S. Ng, Forecasting economic time series using targeted predictors, J. Econom. 146 (2007) 304–317, https://doi.org/10.1016/j.jeconom.2008.08.010.

[30] R. Weron, A. Misiorek, Short-term electricity price forecasting with time series models: a review and evaluation, Complex Electr. Mark (2006) 231–254.

[31] A. Gianfreda, L. Grossi, Forecasting Italian electricity zonal prices with exogenous variables, Energy Econ. 34 (2012) 2228–2239, https://doi.org/10.1016/j.eneco.2012.06.024.

[32] R. Tibshirani, Regression shrinkage and selection via the Lasso, J. R. Stat. Soc. Ser. B (1996) 267–288, https://doi.org/10.1111/j.1467-9868.2011.00771.x.

[33] A.E. Hoerl, R.W. Kennard, Ridge regression: biased estimation for nonorthogonal problems, Technometrics 12 (1970) 55, https://doi.org/10.2307/1267351.

[34] H. Zou, T. Hastie, Regularization and variable selection via the elastic net, J. R. Stat. Soc. Ser. B Stat. Methodol. 67 (2005) 301–320, https://doi.org/10.1111/j.1467-9868.2005.00503.x.

[35] B. Uniejewski, J. Nowotarski, R. Weron, Automated variable selection and shrinkage for day-ahead electricity price forecasting, Energies 9 (2016), https://doi.org/10.3390/en9080621.

[36] Spanish System Operator Information System (e-sios). https://www.esios.ree.es/en (last accessed on 28th May 2018).

[37] A. Bello, J. Reneses, A. Muñoz, Medium-term probabilistic forecasting of extremely low prices in electricity markets: application to the Spanish case, Energies 9 (2016), https://doi.org/10.3390/en9030193.

[38] N. Amjady, Short-term bus load forecasting of power systems by a new hybrid method, IEEE Trans. Power Syst. 22 (2007) 333–341, https://doi.org/10.1109/TPWRS.2006.889130.

[39] B. Uniejewski, R. Weron, F. Ziel, Variance stabilizing transformations for electricity spot price forecasting, Working Paper, 2017. https://ideas.repec.org/p/wuu/wpaper/hsc1701.html.
[40] F.X. Diebold, R.S. Mariano, Comparing predictive accuracy, J. Bus. Econ. Stat. 13 (1995) 253–263, https://doi.org/10.1080/07350015.1995.10524599.
[41] B. Liu, J. Nowotarski, T. Hong, R. Weron, Probabilistic load forecasting via quantile regression averaging on sister forecasts, IEEE Trans. Smart Grid 8 (2017) 730–737, https://doi.org/10.1109/TSG.2015.2437877.
[42] A. Misiorek, S. Trueck, R. Weron, Point and interval forecasting of spot electricity prices: linear vs. non-linear time series models, Stud. Nonlinear Dyn. Econ. 10 (2006), https://doi.org/10.2202/1558-3708.1362.
[43] B. Uniejewski, R. Weron, F. Ziel, Variance stabilizing transformations for electricity spot price forecasting, IEEE Trans. Power Syst. 33 (2018) 2219–2229, https://doi.org/10.1109/TPWRS.2017.2734563.
[44] J. Contreras, R. Espínola, F.J. Nogales, A.J. Conejo, ARIMA models to predict next-day electricity prices, IEEE Trans. Power Syst. 18 (2003) 1014–1020, https://doi.org/10.1109/TPWRS.2002.804943.
[45] A. Pankratz, Building dynamic regression models: model identification, Forecasting with Dynamic Regression Models vol. 935, (2012) 167–201, https://doi.org/10.1002/9781118150528.ch5.
[46] G. Box, G. Jenkins, Time Series Analysis—Forecasting and Control, Holden Day, San Francisco, 1970, https://doi.org/10.1111/j.1467-9892.2009.00643.x.
[47] G.E.P. Box, D.R. Cox, An analysis of transformations, J. R. Stat. Soc. Ser. B 26 (1964) 211–252.