Electricity Price Forecasting in the Short Term Hybridising Fundamental and Econometric Modelling

Rodrigo A. de Marcos, Antonio Bello, Javier Reneses

***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.**

***Index Terms*—Econometric Models, Electricity Markets, Fundamental Models, Hybrid Models, Short-Term Forecasting**

I. 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.

All authors are with the Institute for Research in Technology (IIT) of the ICAI School of Engineering of Comillas Pontifical University, 28015, Madrid, Spain (respective e-mails: demarcos@comillas.edu, abello@comillas.edu and javierr@comillas.edu)

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 [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]–[6]), Artificial Intelligence (AI) techniques (e.g. Support Vector Machine, Radial Basis Function Networks [7]–[11]) and combinations of both types of models [12]–[18].

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 [19]. Furthermore, these models are capable of capturing the revealed behavioural aspects of market participants, such as strategic and speculative behaviour [20].

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 [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], [21], [22].

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 [20], [23]–[25]. 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 [26].

However, according to the work presented in [20], 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 [23]. 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 [24], [27]–[29] 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, [30]), ridge regression [31] and elastic nets (combination of LASSO and ridge regression, [32]). Some of these methods have been recently applied to electricity price forecasting in [33], 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 [20].

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 [23], 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 [23] 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 [23] 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.

This work's hybrid forecasting tool consists of a cost-production optimisation model which is linked to a neural network model. The proposed hybrid model has been tested on a real-size market with complex price dynamics: the Iberian (Portuguese and Spanish) electricity market. Furthermore, other benchmark models were tested on the same case studies so as to further validate the proposed forecasting model.

The remainder of this manuscript is organised as follows: section II describes this work's methodology. Section III presents the case studies in which the proposed forecasting method has been tested, as well as a comparison with other

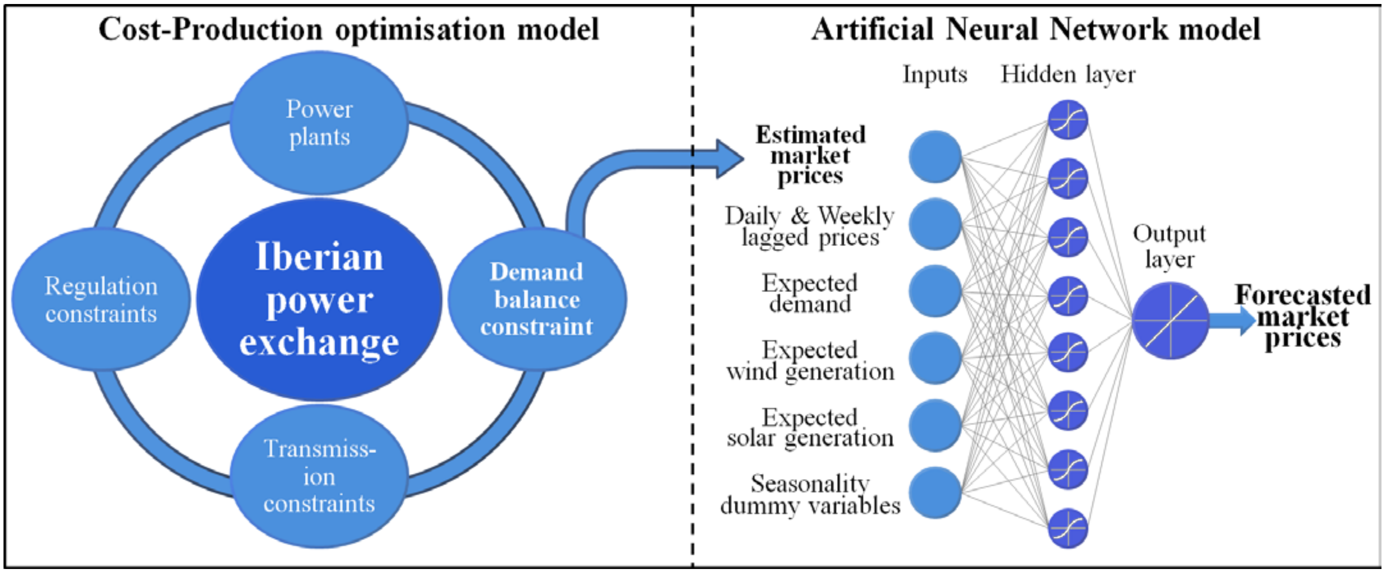

Fig. 1. Proposed hybrid forecasting model

forecasting procedures. Section IV contains the conclusions that were drawn in this work, including the suggestions for extensions and future developments of the proposed methodology.

II. 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 on 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 [20], 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.

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 [20] 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 on 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
- Expected system demand
- Expected system wind generation
- Expected system solar generation
- Dummy variable indicating if it is business day or not (non-business days include Sundays and holidays)
- Dummy variable indicating if it is Saturday or not

These variables are effortlessly obtainable from the Spanish ISO information website [34], 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.

The neural network model configuration is shown on Fig. 1, which is comprised of a hidden layer and an output layer. According to [35], 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

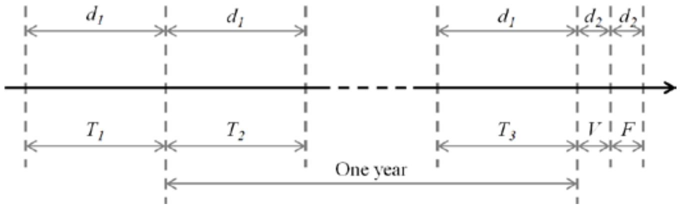

Fig. 2. Training, validation and test/forecast periods arrangement

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 [36]. 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):

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

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. This procedure was carried out five times for numerous days. As a result, the variables that have been mostly discarded are the following two:

- Two-day lagged electricity market prices
- Dummy variable indicating if it is Saturday or not

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 accuracy is measured in terms of the Mean Absolute Percentage Error (MAPE), which is computed as follows:

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

III. Case Studies, Results and Discussion

This section is composed of three 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 five benchmarks.

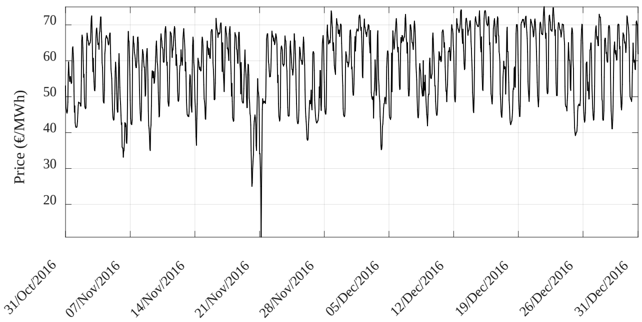

Fig. 3. Iberian electricity market prices for late 2016

A. 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 I), all of which pertain to the period depicted on 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 [34]. 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.

B. 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.

TABLE I
CASE STUDY TRAINING, VALIDATION AND FORECAST PERIODS

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

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-hour horizon and a rolling window of a 24-hour 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 found that the forecasts given in every simulation had a considerably higher volatility than those yielded by the neural network model with a 168-hour horizon, as well as a generally higher MAPE. Therefore, the rolling-window method was not utilised in these cases. Nevertheless, daily lagged prices were not included in the weekly forecasts, because, in reality, daily lagged price information becomes unknown if forecasting further than one day. Nevertheless, the estimated price from the fundamental model may contain, to some extent, the information of daily lagged prices.

The first case study (C_1) is displayed on 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 MAPE of 2.179%, which mostly corresponds to the early morning hours, whereas the late hours are considerably accurate.

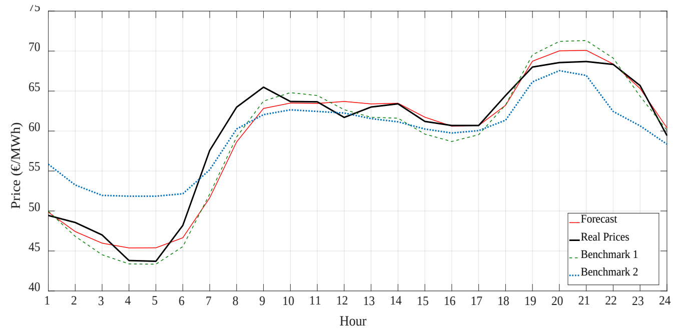

Fig. 4. Electricity price forecast for 15/Nov/2016 (C_1)

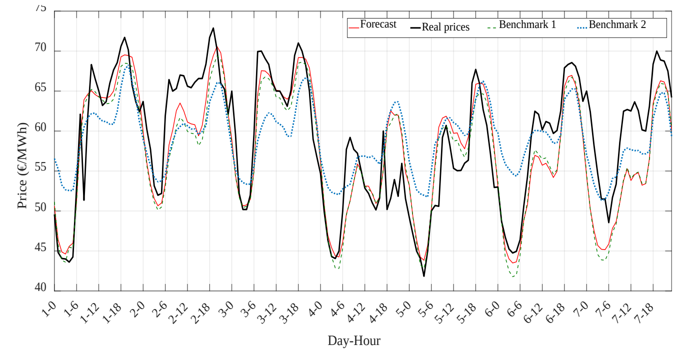

Fig. 5. Electricity price forecast from 05/Dec/2016 to 11/Dec/2016 (C_6)

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 to 11/Dec/2016 is displayed on Fig. 5. In this case study, the resulting MAPE is of 5.878%. Once again, it can be seen on 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.

C. 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 [37] and most recently applied to electricity price forecasting in [33]. This linear regression model can be represented as per the following equations in order to calculate the log-price $p_{d,h}$ at day d and hour h:

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

$$p_{d,h} = log\left(P_{d,h}\right) - \frac{1}{T}\sum_{t=1}^{T} log\left(P_{d,h}\right) \quad (4)$$

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

Furthermore, as mentioned before, a slight modification was carried out, which pertains to the logarithmic transform of Equation (4), where T refers to the training period. The mirror-log transform, recently applied to electricity price forecasting in [38], was applied due to the possibility of prices equal to zero in the Iberian power system, which is represented in the following equations:

$$n_{d,h} = \frac{\left(P_{d,h} - \mu_T\right)}{\sigma_T} \quad (5)$$

$$p_{d,h} = sgn\left(n_{d,h}\right)\left[log\left(n_{d,h} + \frac{1}{c}\right) + log\ (c)\right] \quad (6)$$

First of all, according to Equation (5), 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 c was set to 1/3 as done in [38].

TABLE II
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

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 [39]. 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 [40], [41]. Electricity prices were stabilised (variance) by means of the Box-Cox transformation [42]. 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: $\text{SARIMA}(1,0,0)_{168}(1,0,2)_{24}(1,0,0)_1$. 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:

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

The MAPE of these benchmarks along with those yielded by the proposed electricity price forecasting model on the seven case studies are displayed on TABLE II. 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 (almost 0.3% MAPE). 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.

IV. 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 [37], 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 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.

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.

REFERENCES

[1] A. Bello, J. Reneses, A. Muñoz, and A. Delgadillo, "Probabilistic forecasting of hourly electricity prices in the medium-term using spatial interpolation techniques," *Int. J. Forecast.*, 2015.

[2] R. Weron, "Electricity price forecasting: A review of the state-of-the-art with a look into the future," *Int. J. Forecast.*, vol. 30, no. 4, pp. 1030–1081, 2014.

[3] N. Bowden and J. E. Payne, "Short term forecasting of electricity prices for MISO hubs: Evidence from ARIMA-EGARCH models," *Energy Econ.*, vol. 30, no. 6, pp. 3186–3197, 2008.

[4] R. C. Garcia, J. Contreras, M. van Akkeren, and J. B. C. Garcia, "A

GARCH forecasting model to predict day-ahead electricity prices," *IEEE Trans. Power Syst.*, vol. 20, no. 2, pp. 867–874, 2005.
[5] C. W. Cheong, "Modeling and forecasting crude oil markets using ARCH-type models," *Energy Policy*, vol. 37, no. 6, pp. 2346–2355, 2009.
[6] Z. Tan, J. Zhang, J. Wang, and J. Xu, "Day-ahead electricity price forecasting using wavelet transform combined with ARIMA and GARCH models," *Appl. Energy*, vol. 87, no. 11, pp. 3606–3610, 2010.
[7] H. Chen, C. A. Canizares, and A. Singh, "ANN-based short-term load forecasting in electricity markets," in *2001 IEEE Power Engineering Society Winter Meeting. Conference Proceedings (Cat. No.01CH37194)*, 2001, vol. 2, pp. 411–415.
[8] J. P. S. Catalão, S. J. P. S. Mariano, V. M. F. Mendes, and L. A. F. M. Ferreira, "Short-term electricity prices forecasting in a competitive market: A neural network approach," *Electr. Power Syst. Res.*, vol. 77, no. 10, pp. 1297–1304, 2007.
[9] J. P. S. Catalão, H. M. I. Pousinho, and V. M. F. Mendes, "Hybrid wavelet-PSO-ANFIS approach for short-term electricity prices forecasting," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 137–144, 2011.
[10] I. P. Panapakidis and A. S. Dagoumas, "Day-ahead electricity price forecasting via the application of artificial neural network based models," *Appl. Energy*, vol. 172, pp. 132–151, 2016.
[11] D. Keles, J. Scelle, F. Paraschiv, and W. Fichtner, "Extended forecast methods for day-ahead electricity spot prices applying artificial neural networks," *Appl. Energy*, vol. 162, pp. 218–230, 2016.
[12] M. Shafie-Khah, M. P. Moghaddam, and M. K. Sheikh-El-Eslami, "Price forecasting of day-ahead electricity markets using a hybrid forecast method," *Energy Convers. Manag.*, vol. 52, no. 5, pp. 2165–2169, 2011.
[13] N. Amjady, "Day-Ahead Price Forecasting of Electricity Markets by a New Fuzzy Neural Network," *IEEE Trans. Power Syst.*, vol. 21, no. 2, pp. 887–896, 2006.
[14] P. Areekul, T. Senjyu, H. Toyama, and A. Yona, "A hybrid ARIMA and neural network model for short-term price forecasting in deregulated market," *IEEE Trans. Power Syst.*, vol. 25, no. 1, pp. 524–530, 2010.
[15] N. Chaâbane, "A hybrid ARFIMA and neural network model for electricity price prediction," *Int. J. Electr. Power Energy Syst.*, vol. 55, pp. 187–194, 2014.
[16] X. Yan and N. A. Chowdhury, "Mid-term electricity market clearing price forecasting: A hybrid LSSVM and ARMAX approach," *Int. J. Electr. Power Energy Syst.*, vol. 53, no. 1, pp. 20–26, 2013.
[17] Z. Yang, L. Ce, and L. Lian, "Electricity price forecasting by a hybrid model, combining wavelet transform, ARMA and kernel-based extreme learning machine methods," *Appl. Energy*, vol. 190, pp. 291–305, 2017.
[18] S. Talari, M. Shafie-khah, G. Osório, F. Wang, A. Heidari, and J. P. S. Catalão, "Price Forecasting of Electricity Markets in the Presence of a High Penetration of Wind Power Generators," *Sustainability*, vol. 9, no. 11, p. 2065, 2017.
[19] M. Cerjan, I. Krzelj, M. Vidak, and M. Delimar, "A literature review with statistical analysis of electricity price forecasting methods," in *Eurocon, 2013 IEEE*, 2013, no. July, pp. 756–763.
[20] A. Bello, D. W. Bunn, J. Reneses, and A. Munoz, "Medium-Term Probabilistic Forecasting of Electricity Prices: a Hybrid Approach," *IEEE Trans. Power Syst.*, 2016.
[21] A. Bello, "Medium-term forecasting of electricity prices: a hybrid methodology based on fundamental and technical analysis," PhD Thesis. Universidad Pontificia Comillas, Madrid, Spain, 2016.
[22] J. Nowotarski and R. Weron, "Recent advances in electricity price forecasting: A review of probabilistic forecasting," *Renew. Sustain. Energy Rev.*, 2017.
[23] V. González, J. Contreras, and D. W. Bunn, "Forecasting power prices using a hybrid fundamental-econometric model," *IEEE Trans. Power Syst.*, vol. 27, no. 1, pp. 363–372, 2012.
[24] N. V. Karakatsani and D. W. Bunn, "Forecasting electricity prices: The impact of fundamentals and time-varying coefficients," *Int. J. Forecast.*, vol. 24, no. 4, pp. 764–785, 2008.
[25] A. Bello, D. Bunn, J. Reneses, and A. Muñoz, "Parametric Density Recalibration of a Fundamental Market Model to Forecast Electricity Prices," *Energies*, vol. 9, no. 11, p. 959, 2016.
[26] R. Carmona and M. Coulon, "A Survey of Commodity Markets and Structural Models," in *Quantitative Energy Finance: Modeling, Pricing, and Hedging in Energy and Commodity Markets*, 2014, pp. 41–83.
[27] J. Bai and S. Ng, "Forecasting Economic Time Series Using Targeted Predictors," *J. Econom.*, vol. 146, no. 2, pp. 304–317, 2007.
[28] R. Weron and A. Misiorek, "Short-term electricity price forecasting with time series models: A review and evaluation," *Complex Electr. Mark.*, pp. 231–254, 2006.
[29] A. Gianfreda and L. Grossi, "Forecasting Italian electricity zonal prices with exogenous variables," *Energy Econ.*, vol. 34, no. 6, pp. 2228–2239, 2012.
[30] R. Tibshirani, "Regression Shrinkage and Selection via the Lasso," *J. R. Stat. Soc. Ser. B*, pp. 267–288, 1996.
[31] A. E. Hoerl and R. W. Kennard, "Ridge Regression: Biased Estimation for Nonorthogonal Problems," *Technometrics*, vol. 12, no. 1, p. 55, 1970.
[32] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," *J. R. Stat. Soc. Ser. B Stat. Methodol.*, vol. 67, no. 2, pp. 301–320, 2005.
[33] B. Uniejewski, J. Nowotarski, and R. Weron, "Automated variable selection and shrinkage for day-ahead electricity price forecasting," *Energies*, vol. 9, no. 8, 2016.
[34] Spanish System Operator Information System (e-sios). [Online]. Available: https://www.esios.ree.es/en.
[35] A. Bello, J. Reneses, and A. Muñoz, "Medium-term probabilistic forecasting of extremely low prices in electricity markets: Application to the Spanish case," *Energies*, vol. 9, no. 3, 2016.
[36] N. Amjady, "Short-Term Bus Load Forecasting of Power Systems by a New Hybrid Method," *IEEE Trans. Power Syst.*, vol. 22, no. 1, pp. 333–341, 2007.
[37] A. Misiorek, S. Trueck, and R. Weron, "Point and interval forecasting of spot electricity prices: Linear vs. non-linear time series models," *Stud. Nonlinear Dyn. Econom.*, vol. 10, no. 3, 2006.
[38] B. Uniejewski, R. Weron, and F. Ziel, "Variance Stabilizing Transformations for Electricity Spot Price Forecasting," *IEEE Trans. Power Syst.*, vol. PP, no. 99, 2017.
[39] J. Contreras, R. Espínola, F. J. Nogales, and A. J. Conejo, "ARIMA models to predict next-day electricity prices," *IEEE Trans. Power Syst.*, vol. 18, no. 3, pp. 1014–1020, 2003.
[40] A. Pankratz, "Building Dynamic Regression Models: Model Identification," in *Forecasting with dynamic regression models (Vol. 935)*, 2012, pp. 167–201.
[41] G. Box and G. Jenkins, *Time Series Analysis - Forecasting and Control*. San Francisco: Holden Day, 1970.
[42] G. E. P. Box and D. R. Cox, "An analysis of transformations," *J. R. Stat. Soc. Ser. B*, vol. 26, no. 2, pp. 211–252, 1964.

Rodrigo A. de Marcos received the B.Sc. degree of Industrial Engineer, and M.Sc. degree in Research in Engineering Systems Modelling from the Comillas Pontifical University, Madrid, Spain, in 2013 and 2016, respectively. Since 2014, he is a Ph.D. student at the Institute for Research in Technology (IIT) from the Comillas Pontifical University, Madrid, Spain. His areas of interest include commodity and electricity markets, price forecasting, simulation methods and risk management support.

Antonio Bello received the B.Sc. degree of Industrial Engineer, M.Sc. and Ph.D. degree in Power Systems from the Comillas Pontifical University, Madrid, Spain, in 2010 and 2012 and 2016, respectively. Currently, he is an Assistant Researcher with the Institute for Research in Technology (IIT) from the Comillas Pontifical University, Madrid, Spain. His areas of interest include operation, simulation models, planning of electricity and natural gas markets, energy forecasting and risk management support.

Javier Reneses received the Electrical Engineering degree and Ph.D. degree in Industrial Engineering from the Comillas Pontifical University, Madrid, Spain, in 1996 and 2004, respectively, and the degree in Mathematics from the Open University in Spain (UNED) in 2005. At present, he is an Associate Professor at the Institute for Research in Technology (IIT) from the Comillas Pontifical University, Madrid, Spain, and Visiting Scholar at the Lawrence Berkeley National Laboratory. His areas of interest include operation, simulation models and planning of electricity and natural gas markets, as well as regulation of energy systems.