IJEEP
INTERNATIONAL JOURNAL OF ENERGY ECONOMICS AND POLICY
EJ EconJournals
ISSN: 2146-4553

International Journal of Energy Economics and Policy

ISSN: 2146-4553

available at http: www.econjournals.com

International Journal of Energy Economics and Policy, 2018, 8(3), 97-106.

EJ EconJournals

Forecasting Electricity Price in Colombia: A Comparison Between Neural Network, ARMA Process and Hybrid Models

Jorge Barrientos Marín[1]*, Elkin Tabares Orozco[2], Esteban Velilla[3]

[1]Department of Economics, Faculty of Economic Science, University of Antioquia, Medellín, Colombia, [2]Department of Economics, Faculty of Economic Science, University of Antioquia, Medellín, Colombia, [3]Department of Electrical Engineering, Faculty of Engineering, University of Antioquia, Medellín, Colombia. *Email: jorge.barrientos@udea.edu.co

ABSTRACT

This study aims to predict electricity prices in the Colombian electricity market. To achieve this goal, conventional time series econometrics analysis and one alternative technique based on artificial intelligence algorithms have been implemented. We use autoregressive-moving-average models (ARMAX) and non-linear autoregressive neural networks (NARX). After estimating a hybrid model that combines ARMAX and ARNX models, including exogenous inputs, we forecasted an electricity price time series in a horizon of 12 months ahead (May, 2017). Results show that NARX model's performance is not significantly better than ARMAX's. After applying a Diebold-Mariano test for forecasting accuracy, the null hypothesis is not rejected. This suggests no significant difference in predictive accuracy between the competing methodologies.

Keywords: Stochastic Process, Autoregressive-moving-average, NARX, Random Walk, Predictive Accuracy, Electricity Spot Price
JEL Classifications: C01, C12, C22, C45, C53, L11, L94.

1. INTRODUCTION

Price formation of electricity in short-run markets, as the Colombian one, is a complex process that poses huge challenges for modelling and forecasting. Such complexity is given by the type of good dealt with in this market. This good's particular features include: (i) Electricity is a commodity that cannot be stored, (ii) it is traded in real time; (iii) electricity demand varies day by day; (iv) it has an enormous hydrological linked component; (v) the fixed costs are substantially greater with respect to the variables; (vi) it displays high volatility in spot prices; and (vii) lack of regulations compels producers to declare its real variable costs, which encourages agents to introduce information asymmetries and eventually abuse from market power. This situation generates speculative behaviors among agents of interest, which could have a strong incidence over electricity prices (Mustapha, 2012).

In this context, understanding and forecasting the electricity prices are fundamental processes for every agent involved in the production and supply chains, especially for producers and traders. In fact, analyzing the evolution of the electricity price is essential to design long-run contracts, also called forwards contracts. These contracts are bilateral agreements between agents, covering the risk induced by electricity price volatility in the trade pool. Over 80% of electricity in Colombia is traded using forwards instruments. Furthermore, understanding electricity price formation is crucial for investors, as prices provide the necessary information (market signaling) to expand the installed capacity of the system in the future. In general terms electricity price formation and its forecasting could help power generation firms, as well as consumers, in planning their strategies for maximizing profits and utilities based on mid and short-term perspectives (Murthy et al., 2014). In the long-run term aggregated demand is a very important variable for explaining electricity prices behavior (Smolen and Dudic, 2017). Then, electricity prices influence aggregated demand, which is also an important predictor of economic performance.

A review of international journals on electricity markets and forecasting methods shows a trend where modeling and forecasting electricity price is done using methodologies based on algorithms, as well as parametric and semi-parametric models. Harvey (1990), Nogales et al. (2002), Contreras et al. (2003), Zou and Yang (2004), Weron (2006), Karakatsani and Bunn (2008), Shafie-Khan et al. (2011), Andalib and Atry (2009) and Nan et al. (2014) have reported on the use non-linear autoregressive neural networks with exogenous variables (NARX) to forecast electricity prices, which relates to the main concern of this paper. However, modeling and forecasting electricity prices based on artificial intelligence is uncommon in the field (Weron, 2014). The NARX method seems to be attractive though, given its capacity to capture non-linear relationships while lacking from assumptions about the data generation process, even considering that NARX models are fully parametric.

Two methodologically interesting studies are those by Cadenas et al. (2016) and Santana (2006). In the first study, wind speed forecasting was carried out using ARIMA and NARX models. Then, predictive accuracy of the models was compared by computing the mean squared error. It was found that the NARX models had better performance in accuracy than the ARIMA models. In the second one, the usefulness of neural network methods were evaluated in predicting Colombian inflation, and compared the results yielded with forecasting as provided by seasonal autoregressive integrated moving average model (SARIMA). It was found that neural networks forecasts exhibit higher performance accuracy than SARIMA specification.

In recent years, there has been a growing interest in assessing predictive accuracy among competing forecasting methods under general loss functions and possibly non-normal errors, Diebold and Mariano (2002), Swanson and White (1995), White (2000), Corradi et al. (2001), Cuaresma et al. (2004). The proposed approaches can also provide useful groundwork for model comparison. Thus, in this paper we present a methodology for predictive accuracy closer to Diebold and Mariano (2002). In the Colombian case, the literature on modeling and forecasting electricity prices offers a similar panorama to the international one. A search for working and published papers from Colombian academic institutions and scientific journals offers a poor picture of the field, even though the Colombian electricity market is one of the most cited in worldwide research studies (e.g., economic regulation of electricity industry), particularly after the sector's deep reform in 1994.

For domestic contexts, there are few papers related to forecasting electricity prices using standard methods from the VARMA family (e.g., Lira et al. (2009), Sierra and Castaño (2010), Barrientos et al. (2012) and Barrientos and Martinez (2015)). Our research is aligned with one study by Agudelo (2015), who used NARX to forecast electricity spot prices. To the best of our knowledge, none of the studies reviewed seems to consider more than one methodology for modeling electricity pricing or carry out comparative procedures on predictive accuracy for forecasting.

One singular feature of the previously mentioned works is that the procedures were based on the assumption of non-stationarity of the data generating process or stationarity established by an Augmented Dicky-Fuller (ADF) test (Dickey, 1979). One issue found in this testing procedure is related to the time series level shifts. These shifts usually cheat the ADF test, so that the testing procedure does not reject the null hypothesis of the unit-root process, which is a strong conclusion for monthly electricity time series. Then, regime jumps or level shifts could be a common feature of the electricity prices, as it is in the Colombian case. One of the first papers in modeling and forecasting time series with this feature is that of Huisman (2003). However, in our case, to overcome this issue, we perform a stationary testing procedure based on Cavaliere and Giorgiev (2007), which tests a hypothesis for time series with multiple level shifts.

The development of our work can be summarized as follows: (i) We carried out an exhaustive analysis of the dynamic properties of the involved time series. Specifically, we show that regardless of the multiples level shifts exhibited by the monthly electricity price series, this series does not follow a unit root process. This conclusion results from the application of the Cavaliere test. (ii) We modeled and forecasted the electricity price by estimating NARX and ARMAX specifications with exogenous variables such as water-rivers supply, declared reserves, and demanded electricity. These variables can be considered as market fundamentals of electricity price formation in Colombia. (iii) In order to forecast the monthly electricity price series, not only do we use NARX and ARMA but we also implement a hybrid model combining these two methods, in the spirit of Shafie-Khan et al. (2011). In each case, we estimate the forecast's root mean-squared error. (iv) Finally, we use the Diebold and Mariano (2002) procedure in order to test the predictive accuracy of the competing econometric models.

This paper is structured as follows: Section 2 focuses on the empirical strategy. In Section 3, we describe the statistical models to estimate and test the procedures. Section 4 describes the empirical results. Finally, we present the main conclusions.

2. EMPIRICAL STRATEGY

2.1. Statistical Information

All variables are given in logarithmic form and the electricity price is given in monthly frequency from 01/2001 up to 05/2016. The set of variables used in this study are the following: (i) Energy prices, which correspond to the average monthly spot price, COL/kWh; (ii) hydro reserves (kWh), which is defined as the useful dam volume (water supply availability); (iii) water supplies (kWh), which indicate the physical conditions of river's water supply to dams; the demand (kWh) of the National Interconnected System (which stands in Spanish for Sistema Interconectado Nacional -); (iv) declared availability (kWh), which corresponds to the market supply, and it can be defined as the maximum net-power that one plant could supply to the NIS in a determined period of time and, finally, the ENSO, which takes positive (El Niño) and negative (La Niña) values. The data set used in this paper is provided by XM Company, which operates the Colombian electricity market. ENSO data is taken from the National Oceanic and Atmospheric Administration.

The Graphic 1 shows the evolution of energy price and the exogenous variables. It is clear that the electricity price presents level shifts as well as atypical values, especially for El Niño in the 2015–2016 period. Electricity demand and water-river supply series present seasonality. Then, we carried out a procedure to remove the seasonal component by estimating SARIMA models.

2.2. Stationary Testing

The stationary testing is based on the following empirical model:

$$\Delta y_t = \beta_0 + \beta_1 t + \gamma y_{t-1} + \sum_{i=1}^{p} \delta \Delta y_{t-i} + \varepsilon_t \quad (1)$$

In order to find the value for lag , we estimate all the models from p =0 (no lag) up to p = 12(T/100)$^{0.25}$ (up to the maximum value (equation)), where is the sample size. Next, we choose the best model using standard criteria such as Akaike (AIC) or Bayesian (BIC). Then, under null hypotheses of no-correlation and homoscedasticity, we perform the ADF test for energy price, which indicates that we cannot reject the null hypotheses of the unit-root process, Table 1.

Then, if we take into account the level shifts in electricity price, the ADF test is no longer useful. In order to overcome the level shifts issue, we carry out a stationary testing procedure following Cavaliere (2007). This procedure is based on the following empirical specification:

$$X_t = \varphi' Z_t + Y_t + \mu_t \quad (2)$$

$$Y_t = \alpha Y_{t-1} + u_t \quad (3)$$

$$u_t = \sum_{t-1}^{p} \gamma_i u_{t-1} + \mu_t \quad (4)$$

Where X_t is an observable variable, is a non-observable autoregressive process, Y_t conformable with Z_t and u_t is the non-observable level shifts component. The specification (2)-(4) assumes that every root of the underlying polynomial is >1 in absolute value, and the error term is a white noise. The expression for u_t is given by,

$$\mu_t = \sum_{s=1}^{t} \delta_s \theta_s \quad (5)$$

and we assume that δ_s is completely known (Table 2), then we estimates equation (5), so that we get:

$$\hat{\mu}_t = \sum_{s=1}^{t} \delta_s \Delta X_t \quad (6)$$

Table 1: Augmented Dickey-Fuller stationarity test for electricity price

n=169 interpolated Dickey-Fuller			
Test Statistic -2.2	1% critical value−4.02	5% critical value−3.45	10% critical value−3.2
Critical value (MacK innon) Z (t)=0.471			
Estimated model			
	Coefficient	**Standard error**	**t**
Trend	0.002*	0.001	2.505
Ll price	−0.191*	0.086	−2.238
Constant	0.696*	0.32	2.172

Regression Controls for 13 lags of the differenced ln p. *P<0.05 **P<0.01 ***P<0.001. Source: Author ‘s elaboration

Finally, the Equation (6) is used to obtain the de-jumped series, which we use for our testing procedure:

$$X_t^{\delta} = X_t - \hat{u}_t \quad (7)$$

Specifically, we perform ADF test for de-jumped series, assuming φ = 0 and p is known. De-jumped time series for electricity price is given in Figure 1.

Tables 3-7 show that time series involved in this paper are stationary with level shifts. It is worth noting that electricity demand and rivers’ water contribution, clearly show a seasonality behavior, however once we control for such seasonality we found that these series are stationary with level shifts. The rivers’ water contribution shows regular peaks and valleys, which are taken like extreme, but known, values.

3. METHODOLOGY

3.1. Autoregressive Moving Average Process with Exogenous Variables

The first model considered in this paper has been ARMA with exogenous variables, or ARMAX. In this model, the current value of the dependent variable y_t depends on p lags of y_t and k lags of the exogenous variable in the matrix x_t. A proper ARMAX estimation requires that the entire variables set be weakly stationary and ergodic. As we reported in section 2, the Cavaliere and Giorgiev testing procedure allows us to conclude that all variables are stationary with level shifts. Therefore, the model to be estimated is given by either of the following equations,

$$\phi_p(B) y_t = \theta_q(B) z_t + \sum_{i=1}^{k} \beta_i(B) x_{it} + \varepsilon_t \quad (8)$$

or

$$y_t = \frac{\theta_q(B)}{\phi_p(B)} z_t + \sum_{i=1}^{k} \frac{\beta_i(B)}{\phi_p(B)} x_{it} + \frac{\varepsilon_t}{\phi_p(B)}$$

Figure 1: De-jumped time series for electricity price

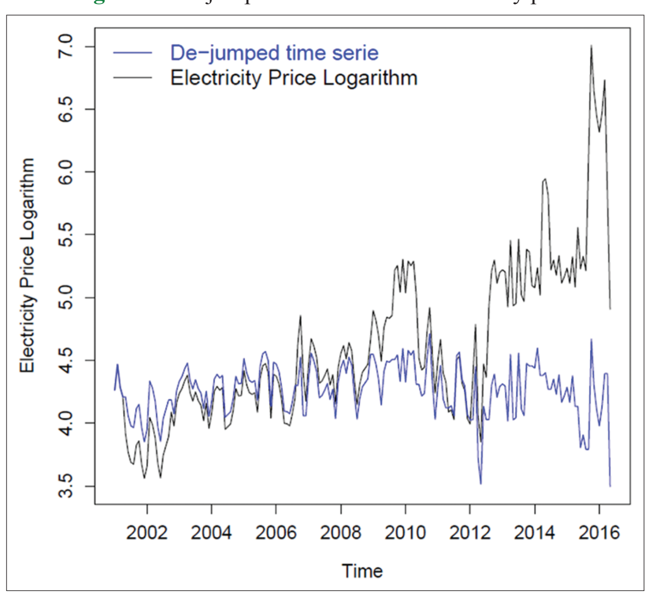


Table 2: Climatic events affecting the electricity energy prices

Date	Obs.	Event	Avg.
June-10	114–124	La Niña	4.52
May-11	125–133	A transition from La Niña to neutral ENSO conditions was made. However, atmospheric conditions continue to remain La Niña atmospheric conditions continue to remain La Niña	4.21
February-12	134–139	La niña weakens	4.31
August-12	140–155	Transition to El Niño conditions. Conditions exist at the boundary between ENSO neutral and weak El Niño.	5.17
December-13	156–159	neutral ENSO	5.11
Abril-14	160–162	There is a continuous evolution towards El Niño	5.9
July- 14	163–172	Decreases El Niño odds - neutral ENSO	5.21
May-15	173–176	El Niño	5.33

Source: Author's elaboration

Table 3: Augmented Dickey-Fuller test for de-jumped electricity price

n=184 interpolated Dickey-Fuller			
Test statistic	1% critical value	5% critical value	10% critical value
−7.64	−3.48	−2.8	−2.5
P-value for critical value (MacKinnon) Z (t)=0.000			
Estimated model			
	Coefficient	**Standard error**	**t**
L1 Dejumped	−0.52***	0.068	−7.6
Constant	2.24***	0.29	7.6

Regression controls for 13 lags of the differenced ln P. * *P*<0.05 **P<0.01 ***P<0.001.
Source: Author's elaboration

Table 4: Augmented Dickey-Fuller test for electricity demand

n=184 interpolated Dickey-Fuller			
Test statistic	1% critical value	5% critical value	10% critical value
−2.39	−3.2	−2.11	−1.98
P-value for critical value (MacKinnon) Z (t)=0.000			
Estimated model			
	Coefficient	**Standard error**	**t**
L1 Dejumped	−0.062***	0.025	−2.3
Constant	1.3***	0.56	2.4

Regression controls for 13 lags of the differenced ln P. * *P*<0.05** *P*<0.01 ***P<0.001.
Source: Author 's elaboration

Table 5: Augmented Dickey-Fuller test for declared availabity

n=184 interpolated Dickey-Fuller			
Test statistic	1% critical value	5% critical value	10% critical value
−2.99	−3.1	−2.71	−2.42
P-value for critical value (MacKinnon) Z (t)=0.061			
Estimated model			
	Coefficient	**Standard error**	**t**
L1 Dejumped	−0.10***	0.035	−2.9
Constant	2.3***	0.8	2.9

Regression controls for 13 lags of the differenced lnp. *P<0.05 **P<0.01 *** *P*<0.001.
Source: Author's elaboration

Where $\varepsilon_t \sim n(0,\sigma^2)$, $\beta i(B)=\beta_0+\beta_1 B+\beta_i B^i$, $\beta_i(B)=\beta_0+\beta_1 B+\beta_i Bi$, $\phi_P(B) = 1 - \phi_1 B - \ldots - \phi_p Bp$ and $\theta_q(B) = 1+\theta_1 B+\ldots+\theta_q B^q$.

The usual assumption on polynomials must be fulfilled, which means that all root should be inside the unit circle. Maximum Likelihood estimation (MLE) is carried out to estimate all parameters of interest. To forecast, we perform a standard recursive process. The first forecasted data obtained from the observable data set is used to forecast the value of the next period, and so on. Given that the model contains exogenous variables, the forecasting values of y_t should be conditional to the futures values of x_t. As suggested by Harvey (1990), the one-step ahead forecast of y_t can be expressed as:

$$\hat{y}_{T+l|T} = \sum_{i=1}^{p} \hat{\phi}_i \hat{y}_{T+l-i|T} + \sum_{i=1}^{k} \hat{\beta}_i \hat{x}_{T+l} + \sum_{i=1}^{q} \hat{\theta}_i \hat{z}_{T+l-i|T} \tag{9}$$

3.2. Artificial Neural Networks (ANN)

ANN are a computational tool that permits to find relationships or patterns between variables, interpolate data, predict and model no-linear behavior between variables (Kohonen et al. (1996) and Meireles (2003). One of the most common ANN architectures is the multilayer feedforward network, in which inputs and outputs are interconnected by weights, while transfer functions connect to each layer (hidden and output layers) formed by neurons. The first hidden layer receives the information from the outside and the output layer delivers the response of the network.

Another architecture that considers lags of the output variable is the nonlinear autoregressive networks with exogenous inputs (NARX). Due to the close loop in time-series models, this architecture permits to mix the advantages of autoregressive models (recurrent) with the capacity of ANN to predict the output variable (y_t) as a result of exogenous variables (x_t) and lags of y_t in a specific time t. When considering one hidden layer, and the output transfer function is a linear function, the relationship between the NARX model's variables can be shown in Equation (11), in which y_t is the predicted result; x_t is a set of exogenous variables as well as their respective lags; H is the number of neurons by layer; Alfa and Beta are the weight and bias of the ANN; G is the neural transfer function; e_t is the error, and sigma is the standard deviation of the error.

Table 6: Augmented Dickey-Fuller test for water contribution

n=184 interpolated Dickey-Fuller			
Test statistic −2.92	1% critical value −3.3	5% critical value −2.8	10% critical value −2.5
P-value for critical value (MacKinnon) Z (t)=0.056			
Estimated model			
	Coefficient	**Standard error**	**t**
L1 Dejumped	−0.066***	0.026	−2.5
Constant	1.47***	0.58	2.5

Regression controls for 13 lags of the differenced lnp. * *P*<0.05 **P<0.01 ***P<0.001. Source: Author's elaboration

Table 7: Augmented Dickey-Fuller test for water rerserves

n=184 interpolated Dickey-Fuller			
Test statistic −3.77	1% critical value −3.48	5% critical value −2.88	10% critical value −2.57
P-value for critical value (MacKinnon) Z (t)=0.0032			
Estimated model			
	Coefficient	**Standard error**	**t**
L1 Dejumped	−0.147***	0.039	−3.77
Constant	3.40***	0.9	3.77

Regression Controls for 13 lags of the differenced lnp. *P<0.05 **P<0.01 ***P<0.001. Source: Author's elaboration

$$y_t=\beta_*+\sum_{i=1}^{I}\varphi_i x_t^{(i)}+\sum_{h=1}^{H}\beta_h G\left(2\sigma_y\right)^{-1}\alpha_{*,h}+\sum_{i=1}^{I}\alpha_{i,h}x_t^{(i)}+e_t \quad (10)$$

The ANN model can be completed when the architecture, the layer number, neurons weights, and bias are defined (Haykin and Network, 2004). In this paper, the NARX structure with one hidden layer and the sigmoidal function (Equation 11), as transfer function, is selected. Other parameters such as weights and bias are obtained, thus maximizing the logarithm of likelihood error function, or by simply minimizing the average square error between the output of the model and the target when the number of neurons is defined. In this way, the last parameters can be obtained using numerical methods, for instance Levenberg-Marquardt back propagation, or the BFGS method by means of a Hessian Matrix (Broyden, 1970).

$$G\left(z_i\right)=\frac{1}{1+e^{-z}} \quad (11)$$

To avoid over-fitting of the ANN model, we use regularization to solve ill-conditioned inverse problems (Hinton, 1990). This methodology aims to spot trade-off between training data reliability and the model's output (y_t), thus minimizing total risk (R), as suggested by Velasquez et al. (2013). In the supervised learning process, risk can be linked to the model's performance, as measured by the squared error Equation 12, in which the term $\xi_c(W)$ is the penalty or regularization strategy connected to weights ($w_{p,h}$) of the layer p and neuron q (Equation 13), and z is λ is the factor that affects the training process, or regularization.

$$R\left(W\right)=\sum_{i=1}^{T}\left(\hat{y}_t-y_t\right)^2+\lambda\xi_c\left(W\right) \quad (12)$$

$$\xi_c\left(W\right)=\|w_{p,h}\|^2=\sum_{h=1}^{h}\sum_{p=1}^{p}w^2{}_{p,h} \quad (13)$$

3.3. Diebold and Mariano Test

Accuracy tests are usually run to forecast two same variables, y_t, from two different models. Let $\hat{y}^i_{(t+h|t)}$ for i = 1,2, two forecasting procedures from the same data-generating process. Then, forecasting errors are given by $\hat{\varepsilon}^i_{t+h|t}=y_{t+h}-\hat{y}^i_{t+h|t}$ for i = 1,2. At this point, it is important to keep in mind that forecasting electricity prices can be used to guide public policies on expanding electricity systems, investments decisions on electric plants, design bilateral agreements, among other economic decision-making purposes. Then, the expected loss associated with every single forecast error is induced by the decision-making problems faced by policy makers and agents in the market. Therefore, it is clear that the expected loss is a general function of $\hat{\varepsilon}^i_{t+h|t}$.

It is assumed that the forecasting procedure is performed for h periods ahead, in which = t0,…, T The results are two forecasting series of simple size T, two forecasting errors, and two loss functions denoted by $L\left(y_{t+h},\hat{y}^i_{t+h|t}\right)=L\left(\hat{\varepsilon}^i_{t+h|t}\right)$ for i=1,2. If the forecasting methods yield similar results, their loss function should be very close to each other. Forecasting accuracy can be evaluated by comparing the expected value of the differences of these two functions, as proposed by Diebold and Mariano (2002). This testing procedure is based on the following hypotheses:

$$H_0:E\left[L(\hat{\varepsilon}^1_{t+h|t})\right]=E\left[L(\hat{\varepsilon}^2_{t+h|t})\right] \quad (14)$$

versus

$$H_1:E\left[L(\hat{\varepsilon}^1_{t+h|t})\right]\neq E\left[L(\hat{\varepsilon}^2_{t+h|t})\right] \quad (15)$$

Or, in terms of a sample path of a loss-differential series $\{d_t\}_{t=1}^{T}$, these hypotheses are given by H_0: $E[d_t]=0$ against H_0: $E[d_t]\neq 0$, where $d_t=L\left(\hat{\varepsilon}^1_{t+h|t}\right)-L\left(\hat{\varepsilon}^2_{t+h|t}\right)$. Diebold and Mariano (2002) show that a consistent estimator of $E[d_t]$ is given by $\bar{d}=\frac{1}{T}\sum_{t=1}^{T}d_t$ and its limit distribution is given by $\sqrt{T}\left(\bar{d}-\mu\right)\rightarrow N\left(0,2\pi f_0(0)\right)$. As expected, under H_0 true, the statistical test $S=\bar{d}\Big/\sqrt{2\pi f_0(0)/T}$ converges to N (0,1) as $T\rightarrow\infty$. Then, high values of S are evidence against H_0.

4. EMPIRICAL RESULTS

4.1. Forecasting Exogenous Variables

In this section, we present the complete forecasting matrix, which is required for energy price forecasting. The seasonality unit-root test is performed for every subset of the exogenous variables. This testing hypothesis could be found in Osborn et al. (1988), and is called OCSB test. The results of this testing procedure are displayed in Figure 2. All estimated models have 80% and 95% bootstrap confidence-bands.

4.2. Estimation Models

The MLE of the ARMAX model's parameters are shown in Table 8. All the estimated parameters are statistically significant as, reported by the p-value. The Ljung-Box statistical test for 12 and 18 lags, with corresponding P-values 0.87 and 0.90, shows no correlation within the residuals of the ARMAX model. Moreover, McCleod-Li heteroscedasticity test could not reject the null hypothesis of constant variance of the residuals (Mcleod and Li, 1983).

The NARX model estimation was carried out including 8 lags of the electricity price, one hidden cover with ten neurons, while the exogenous entries are the same as those included in the ARMAX estimation. Every column in the matrix x_t was normalized by dividing each element by the max $\{x_t\}$, and the initial weights were randomly chosen in the set [−0.5, 0.5]. The activation function used is a sigmoidal specification for the staring cover and the linear function for the cover out. The Ljung-Box statistical test for 18 and 24 lags, with corresponding P-values 0.54 and 0.60, shows no correlation among the residuals of the NARX model.

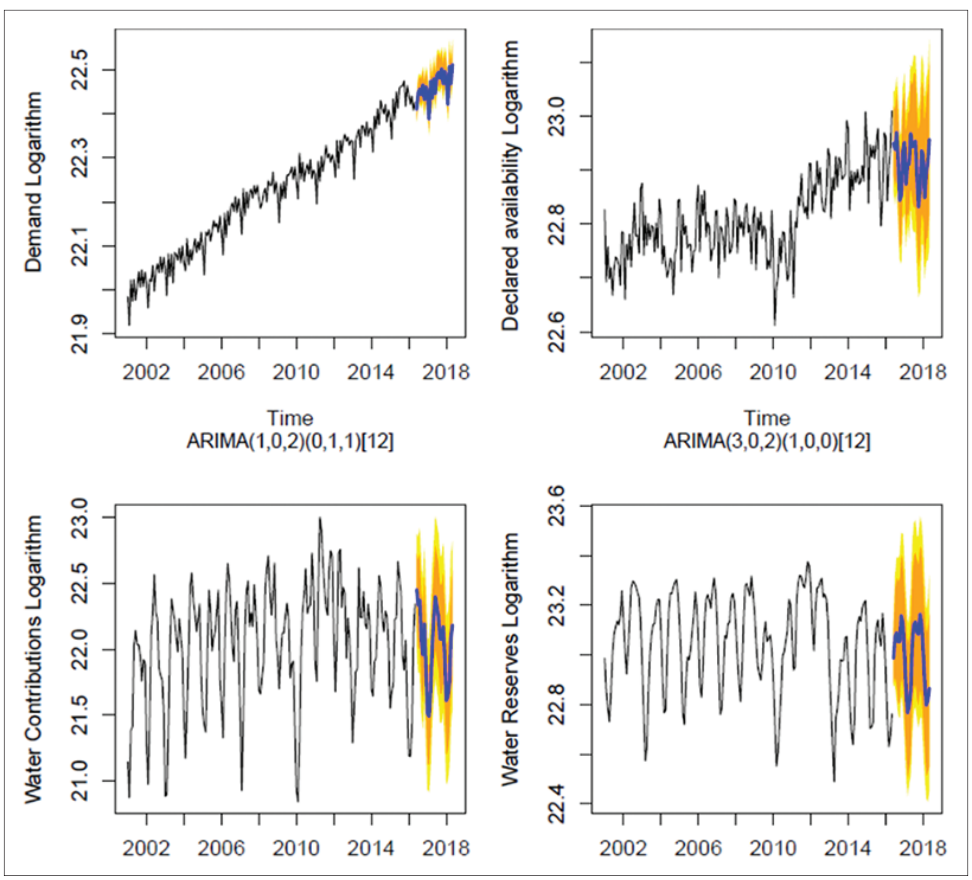
In order to get the final forecasting price, one hundred simulations of the NARX were made using an average on the entire set of values. Minimization of the loss function was made using the BFGS algorithm. Finally, the weight-decomposition parameter was fixed at 0.45. The hybrid model was constructed by means of weighing the results obtained with ARMAX and NARX using a 0.5 weight for each model.

4.3. Validation of the Models and Diebold and Mariano Accuracy Test

Selecting a sample to validate the forecasting was made by sorting the sample and splitting it into two subsets. The first subset was used for estimation (ARMAX model) or training (NARX), and the second observable data subset was compared with the forecasting made with the first data set. The validation has been made for two different sub-samples: The first period goes from 08/2014 and 06/2015, in which electricity prices stability can be observed; the second period goes from 08/2014 up to 05/2016, in which several atypical values are related to El Niño phenomenon, between 09/2015 and 05/2016.

Table 9 shows the estimated root mean squared error (RMSE) for the tree models. It is worth noting that this RMSE is a measurement of the differences between the forecasting yielded by the models and the data observed. In fact, the RMSE is a measure of accuracy that permits to compare forecasting errors of different models for a particular data set. Let us note that, in the first sub-sample, the ARMAX model exhibits the smallest RMSE. This means that such specification provides best forecasting values. However, when atypical values are considered in the second sub-sample, the NARX model displays better performance. Even though it seems difficult to capture and forecast level shifts and atypical values, the NARX model

Figure 2: Exogenous variables forecast



manages to capture much better the electricity prices evolution, providing more accurate forecasting values. Figures 3 and 4 compare observable data (black line) and forecasted data (red and blue lines), along with 95\% bootstrap confidence-bands (orange and yellow lines).

Table 10 shows the results related to the implementation of the Diebold and Mariano test for the two validation periods. In both cases, we cannot reject the null hypothesis where the difference between the loss functions be zero. Then, the forecasting of ARMAX or NARX models are statistically equal. This result is interesting because, in fact, the expected result is a scenario where NARX shows better performance than ARMAX (e.g., Cadenas et al. (2016) and Velasquez et al., 2013).

4.4. Forecasting Electricity Price

This section shows the performance of our forecasting up to 05/2017. We would like to highlight that we did not use monthly data from 06/2016 up to 05/2017. Figure 5 shows an increasingly persistent rising trend of the electricity price. Although it does not predict any atypical value in the near future, the results seem to predict some level shifts. According to Figure 6, it is clear that the implemented procedure provides a rather accurate pattern when the forecasting values are compared with the observed data, even though it does not reproduce the actual magnitude of the observed electricity prices.

Table 8: ARMAX model estimation

Dependent variable: Electricity price logarithm Method: Maximum likelihood observations: 185 Variable	Coefficient	Standard error	t	P-value
AR(l)	0.594	0.062	9.563	0.000***
AR (8)	−0.131	0.064	−2.042	0.041**
Constant	−36.429	5.381	−6.769	0.000***
Demand	3.388	0.301	11.233	0.000***
Demand (1)	1.202	0.215	5.579	0.000***
Declared availability	−1.623	0.266	−6.083	0.000***
Wlater contributions	−0.552	0.076	−7.176	0.000***
Wlater reserves	0.653	0.275	2.372	0.017**
Water reserves (1)	−1.163	0.247	−4.699	0.000***
Child phenomenon (1)	0.148	0.027	5.491	0.000***
D98	0.272	0.123	2.206	0.027**
Dl lO	−0.370	0.123	−3.010	0.002***
D152	0.401	0.122	3.290	0.001***
D137	−0.448	0.137	−3.253	0.001***
D138	−0.491	0.136	−3.612	0.000***
D161	0.516	0.137	3.766	0.000***
D162	0.223	0.131	1.699	0.090*
D179	0.545	0.143	3.811	0.000***
D180	0.381	0.141	2.691	0.007***
D184	0.399	0.123	3.237	0.001***
D186	−0.546	0.145	−3.747	0.000***
Sigma	0.0199			
Log-likelihood	98.07			
AIC	−154.15			

Source: Author's elaboration

Finally, Table 11 shows that ARMAX, NARX, and Hybrid models behave similarly as shown by the RMSE, which does not constitute a proper tool for model selection. At the beginning we could think that NARX is not well trained. However, when its performance about forecasting is evaluated by computing the mean absolute error (MAE), which operates in L1-norm, we find that NARX has a smaller forecasting error than the ARIMAX and Hibryd models. Since any loss function computed in L2-norm squares the error, then usually NARX model will show a much larger forecasting error than the loss function in L1-norm, so the NARX will be much more sensitive than the ARIMAX.

5. CONCLUSIONS

The results stemming from empirical analyses show that the monthly electricity price series displays no stochastic and non-predictable behaviors. In other words, the price series is stationary with level shifts. The Diebold and Mariano test suggests that both procedures yield rather similar forecasting values among competitive models. However, NARX performs slightly better than ARMAX, particularly when atypical values are take into account. Based on these facts, the estimation methodologies reported in this paper, training and forecasting, performed adequately for our sample, while the performance of the model implemented for the 06/2016 period yielded less precise forecasting of electricity price. However, it is worth noting the fact that although our procedure seems not to accurately forecast electricity prices level, it can reproduce future price patterns with high precision.

Finally, concerning forecasting values, the procedure implemented in this paper suggests that electricity prices will display an increasing trend. In fact, our forecasting procedure shows that

Table 9: Error forecasting in validation

Period	Model	RMSE
August 2014-June 2015	ARMAX	0.121
	NARX	0.143
	Hybr id	0.130
August 2014- May 2016	ARMAX	0.419
	NARX	0.404
	Hybr id	0.410
Source: Author ‘s elaboration		

RMSE: Root mean squared error

Table 10: Diebold-Mariano test between the ARMAX model and the artificial neuronal network

Period	Forecast horizon	Statistical	P-value
August 2014-June 2015	11	0.875	0.808
August 2014-May 2016	21	0.747	0.772

Source: Author's elaboration

Table 11: Forecasting error: L_2-norm versus L_1-norm

Period	Model	RMSE	MAE
June 2016–May 2017	ARMAX	0.595	0.482
	NARX	0.594	0.456
	Hybrid	0.60	0.489

Source: author's elaboration. MAE: Mean absolute error

Figure 3: Validation of the models for sample (August 2014–June 2015)

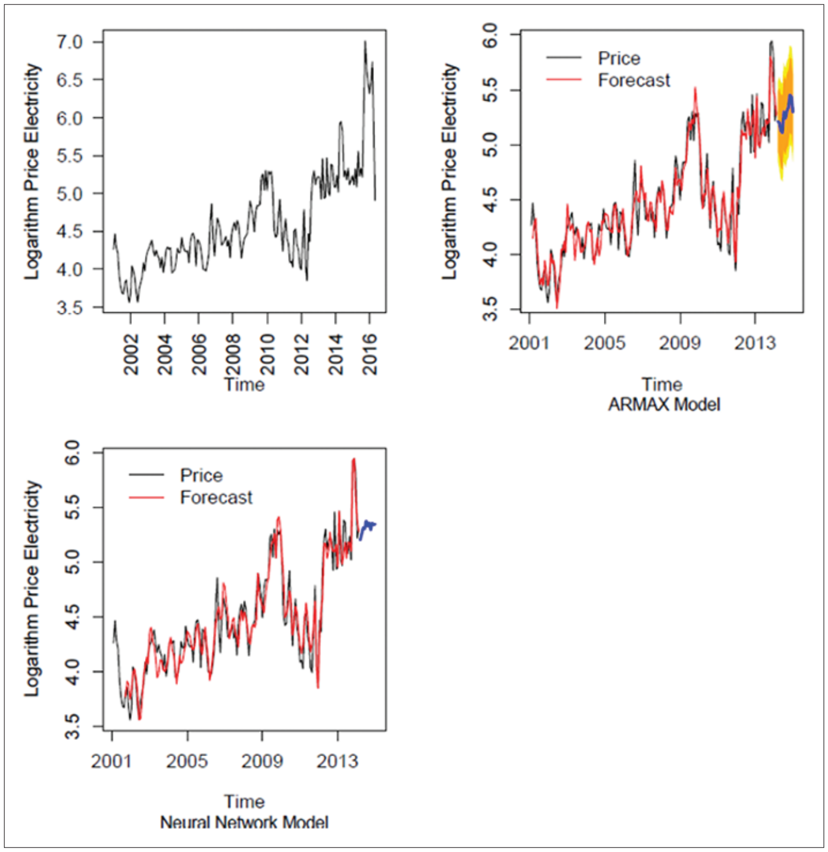


Figure 4: Validation of the models for Sample (August 2014–May 2016)

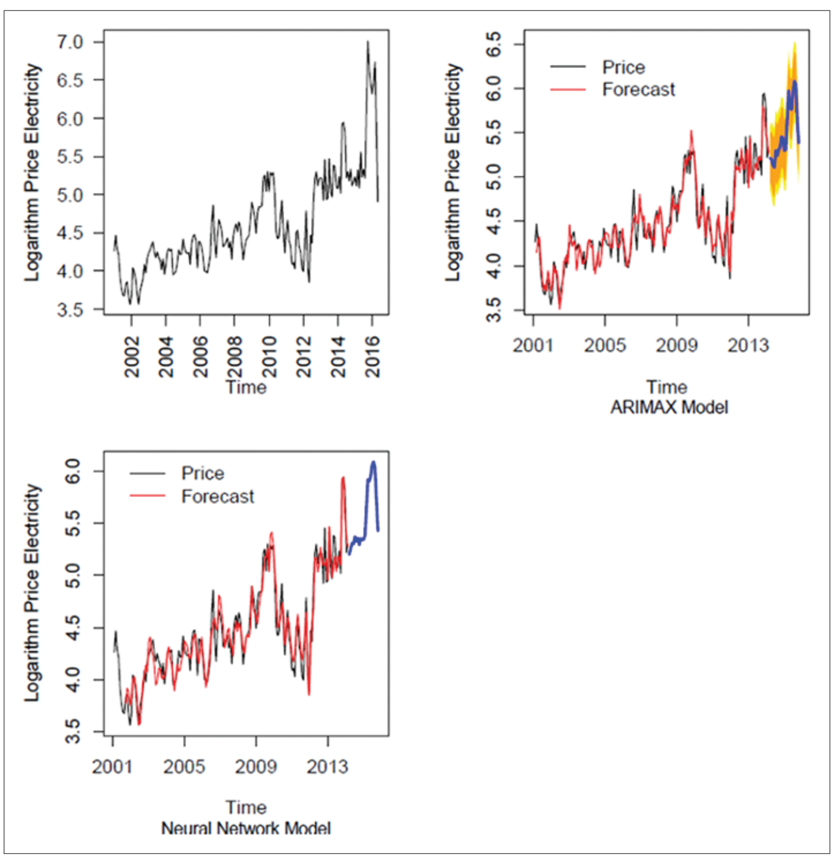


Figure 5: Price forecast through May 2017

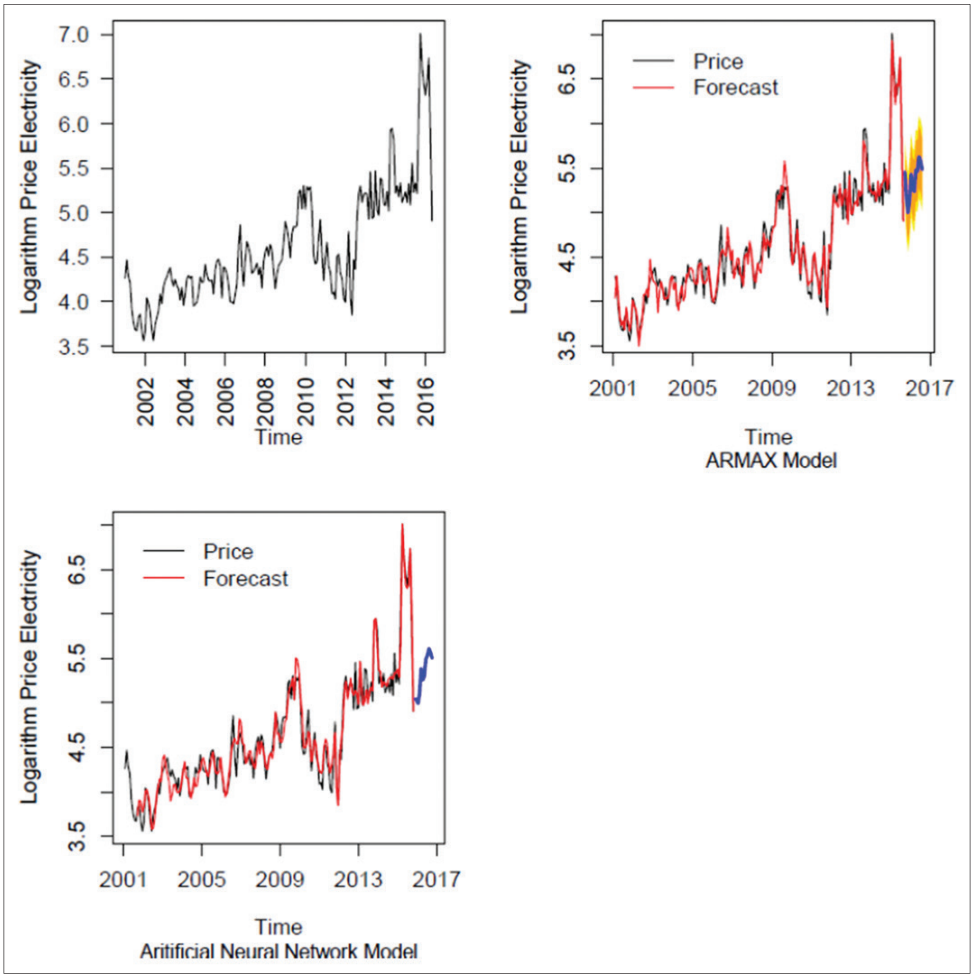
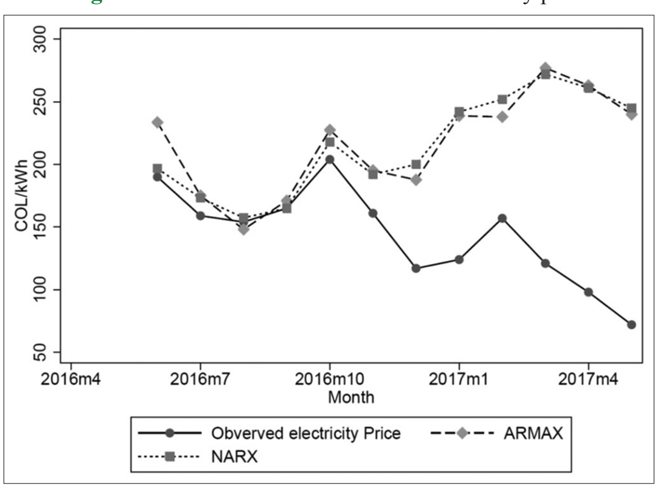


Figure 6: Observed versus forecasted electricity price



monthly electricity prices, in the short-run, will display level shifts and some form of volatility.

REFERENCES

Agudelo, A.P., López-Lezama, J.M., Velilla, E. (2015), Predicción de la electricidad en bolsa mediante un modelo neuronal no-lineal autoregresivo con entradas exógenas. Información Tecnológica, 26(6), 99-108.

Andalib, A., Atry, F. (2009), Multi-step ahead forecasts for electricity prices using NARX: A new approach, a critical analysis of one-step ahead forecasts. Energy Conversion and Management, 50(3), 739-747.

Barrientos, J., Martinez., M. (2015), La hidrología a como predictor del comportamiento del precio de energía en bolsa. Perfil de Coyuntura Económica, 25, 90-123.

Barrientos, J., Rodas, E., Velilla, E., Lopera, M., Villada, F. (2012), Modelo para el pronóstico del precio de la energía eléctrica en Colombia. Lecturas de Economía, 77(22), 91-127.

Broyden, C.G. (1970), The convergence of a class of double-rank minimization algorithms: 2. The new algorithm. IMA Journal of Applied Mathematics, 6(3), 222-231.

Cadenas, E., Rivera, W., Campos-Ameszcua, R., Heard, C. (2016), Wind speed prediction using a univariate ARIMA model and a multivariate NARX model. Energies, 9(2), 109.

Cavaliere, G., Geogiev, I. (2007), Testing for unit roots in autoregressions with multiple level shifts. Econometric Theory, 23(6), 1162-1215.

Contreras, J., Espinola, R., Nogales, F.J., Conejo, A.J. (2003), ARIMA models to predict next-day electricity prices. IEEE Transactions on Power Systems, 18(3), 1014-1020.

Corradi, V., Swanson, N.R., Olivetti, C. (2001), Predictive ability with cointegrated variables. Journal of Econometrics, 104(2), 315-358.

Cuaresma, J.C., Hlouskova, J., Kossmeier, S., Obersteiner, M. (2001), Forecasting electricity spot-prices using linear univariate time-series models. Applied Energy, 77(1), 87-106.

Dickey, D.A., Fuller, W.A. (1979), Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association, 74(366a), 427-431.

Diebold, F.X., Mariano, R.S. (2002), Comparing predictive accuracy. Journal of Business and Economic Statistics, 20(1), 134-144.

Harvey, A.C. (1990), Forecasting, Structural Time Series Models and the Kalman Filter. UK: Cambridge University Press.

Haykin, S., Network, N. (2004), A comprehensive foundation. Neural Networks, 2, 41.

Hinton, G.E. (1990), Connectionist learning procedures. Artificial Intelligence, 40(1-3), 185-234, 1989. Reprinted. In: Carbonell, J., editor. Machine Learning: Paradigms and Methods. Cambridge, MA: MIT Press.

Huisman, R., Mahieu, R. (2003), Regime jumps in electricity prices. Energy Economics, 25(5), 425-434.

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

Kohonen, T., Oja, E., Simula, O., Visa, A., Kangas, J. (1996), Engineering applications of the self-organizing map. Proceedings of the IEEE, 84(10), 1358-1384.

Lira, F., Muñoz, C., Nunez. F., Cipriano, A. (2009), Short-term forecasting of electricity prices in the Colombian electricity market. IET Generation, Transmission and Distribution, 3(11), 980-986.

McLeod, A.I., Li, W.K. (1983), Diagnostic checking ARMA time series models using squared-residual autocorrelations. Journal of Time Series Analysis, 4(4), 269-273.

Meireles, M.R., Almeida, P.E., Simoes, M.G. (2003), A comprehensive review for industrial applicability of artificial neural networks. IEEE Transactions on Industrial Electronics, 50(3), 585-601.

Murthy, G.G.P., Sedidi, V., Panda, A.K., Rath, B.N. (2014), Forecasting electricity prices in deregulated wholesale spot electricity market-a review. International Journal of Energy Economics and Policy, 4(1), 32.

Mustapha, U.M. (2012), The role of speculation in the determination of energy prices. International Journal of Energy Economics and Policy, 2(4), 279.

Nan, F., Bordignon, S., Bunn, D.W., Lisi, F. (2014), The forecasting accuracy of electricity price formation models. International Journal of Energy and Statistics, 2(1), 1-26.

Nogales, F.J., Contreras, J., Conejo, A.J., Espínola, R. (2002), Forecasting next-day electricity prices by time series models. IEEE Transactions on Power Systems, 17(2), 342-348.

Osborn, D.R., Chui, A.P., Smith, J.P., Birchenhall, C.R. (1988), Seasonality and the order of integration for consumption. Oxford Bulletin of Economics and Statistics, 50(4), 361-377.

Santana, J.C. (2006), Predicción de series temporales con redes neuronales: Una aplicación a la inflación colombiana. Revista Colombiana de Estadística, 29(1), 77-92.

Shafie-Khan, M., Moghaddam, M.P., Sheikh-El-Eslami, M. (2011), Price forecasting of day-ahead electricity markets using a hybrid forecast method. Energy Conversion and Management, 52(5), 2165-2169.

Sierra, J., Castaño, E. (2010), Pronóstico del precios spot del mercado eléctrico colombiano con modelos de parámetros variantes en el tiempo y variables fundamentales. Estadística Aplicada: Didáctica de la Estadística y Métodos Estadísticos en Problemas Socioeconómicos. Colombia: Universidad Nacional de Colombia.

Smolen, J., Dudic, B. (2017), The role of residual demand in electricity price analysis and forecasting: Case of czech electricity market. International Journal of Energy Economics and Policy, 7(5), 152-158.

Swanson, N.R., White, H. (1995), A model-selection approach to assessing the information in the term structure using linear models and artificial neural networks. Journal of Business and Economic Statistics, 13(3), 265-275.

Velasquez, J.D., Fonnegra, Y., Villa, F.A. (2013), Pronóstico de series de tiempo con redes neuronales regularizadas y validación cruzada. Revista Vínculos, 10(1), 267-279.

Weron, R. (2006), Short-term Electricity Price Forecasting with Time Series Models: A Review and Evaluation. Poland: Hugo Steinhaus Center, Wroclaw University of Technology.

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

White, H. (2000), A reality check for data snooping. Econometrica, 68(5), 1097-1126.

Zou, H., Yang, Y. (2004), Combining time series models for forecasting. International Journal of Forecasting, 20(1), 69-84.