UNIVERSIDAD EAFIT

Vigilada Mineducación

Documentos de trabajo

Economía y Finanzas

N° 19-14

2019

Modeling the electricity spot price with switching regime semi-nonparametric distributions

Alfredo Trespalacios, Lina M. Cortés, Javier Perote

cief
Centro de Investigaciones Económicas y Financieras
Universidad EAFIT

Modeling the electricity spot price with switching regime semi-nonparametric distributions[†]

Alfredo Trespalacios[‡]
Lina M. Cortés[§]
Javier Perote[**]

November 22, 2019

Abstract

Spot prices of electricity in liberalized markets feature seasonality, mean reversion, random short-term jumps, skewness and highly kurtosis, as a result from the interaction between the supply and demand and the physical restrictions for transportation and storage. To account for such stylized facts, we propose a stochastic process with a component of mean reversion and switching regime to represent the dynamics of the spot price of electricity and its logarithm. The short-term movements are represented by semi-nonparametric (SNP) distributions, in contrast to previous studies that traditionally assume Gaussian processes. The application is done for the Colombian electricity market, where El Niño phenomenon represents an additional source of risk that should be considered to guarantee long-term supply, sustainability of investments and efficiency of prices. We show that the switching regime model with SNP distributions for the random components outperforms traditional models leading to accurate estimates and simulations, and thus being a useful tool for risk management and policy making.

Keywords: Electricity markets, Gram-Charlier series, switching regime models, Ornstein–Uhlembeck process.

JEL Classification: C14, C22, C53, L94, L98, Q2

[†] Authors hold sole responsibility for the views expressed in this working paper, which may not necessarily reflect the opinion or policy of the Universidad EAFIT or the Center for Research in Economics and Finance (Cief).

[‡] Department of Finance, Faculty of Economic and Administrative Sciences. Institucion Universitaria Instituto Tecnológico Metropolitano (ITM). Address: Calle 54A No 30-01 Medellin, Colombia. E-mail: alfredotrespalacios@itm.edu.co

[§] Department of Finance, School of Economics and Finance. Universidad EAFIT. Address: Carrera 49 No 7 Sur-50 Medellin, Colombia. E-mail: lcortesd@eafit.edu.co. Corresponding author. Phone: +5742619500 ext. 9756. Fax: +5743120649.

[**] Department of Economics and IME, University of Salamanca. Address: Campus Miguel de Unamuno 37007 Salamanca, Spain. E-mail: perote@usal.es

1. Introduction

After nearly three decades of the wave of liberalization of electricity markets in the 1990s, electrical energy is one of the primary commodities involved in most industrial processes and daily human activities worldwide (Sioshansi, 2006; Peerbocus, 2007). The current commitments that governments, companies, and individuals are making in terms of sustainability are related, to a large extent, to the way primary, renewable, and non-renewable sources of energy are exploited for conversion or substitution (Burke and Stephens, 2018).

The way electricity is traded has been defined in multiple markets that mainly consider three salient characteristics of this commodity: (i) the limitation to store large amounts of electricity for long periods of time, (ii) the technical difficulties or environmental and social restrictions for long-distance transmission, and (iii) the intensive use of capital required for expanding systems at a large scale, which present long and uncertain payback periods. The most common products in an electricity market are defined for the transaction of expansion, energy, and auxiliary services. To buy and sell electricity, a set of mechanisms have been developed; they integrate the physical condition of a commodity that is produced at the same time it is transported and consumed, while ensuring the cash flows of the agents involved. The spot market, forward contracts, and future contracts are the most common mechanisms to define the quantities and prices of transactions between sellers and buyers (Trespalacios, Pantoja, and Fernández, 2017; Knaut and Paschmann, 2019).

The spot market defines, at every instant, the price of the underlying entity (electricity) according to the fit of the supply and demand offer in real time and to a first price sealed-bid auction. Both functions (demand and production) are modified at every time interval and, in some markets, they are different even at each node of the transmission network. In accordance with the customs and behaviors of the households, commerce, and the industry, at certain times of the day electricity consumption can be low or high. This exerts pressures on price reduction or increase. The pattern of the demand curve can also be found weekly, monthly, and annually. Likewise, the supply curve is affected by climate conditions, such as hydrology, solar radiation, and wind. In addition to the seasonal conditions, other unpredictable phenomena affect the price formation in the short or medium term: generator failures, modifications to the configuration of the transmission network, fuel price variations, or climatic phenomena. The interplay of these conditions results in a series of prices with certain dynamics different from those of other primary commodities or financial markets (Pilipovic, 2007; Weron, 2008).

According to Espen, Kallsen, and Meyer-Brandis (2005), the dynamics of the electricity price exhibit three significant characteristics: seasonality in several horizons (hourly, weekly, monthly, mean reversion, jumps, and heavy tails, which have been described by other authors (e.g. Lucia and Schwartz, 2002; Huisman and Huurman, 2003; Geman and Roncoroni, 2003; Cartea and Figueroa, 2005; Pilipovic, 2007; Falbo, Fattore, and Stefani, 2010, Escribano, Peña, and Villaplana, 2011; Weron, 2014; Trespalacios, Pantoja, and Fernández, , 2017; Dupuis, 2017; Mayer and Trück, 2018; Baum, Zerilli, and Chen, 2019; Hinderks and Wagner, 2019).

The seasonality of the demand is directly associated with industrial production, the consumption patterns of households, and the contributions of energy sources with storage limitations: water, wind, and solar radiation. Such seasonality is usually presented using categorical (dummy) variables or sinusoidal functions derived from the Fourier transform. Mean reversion, which has been widely applied in the study of interest rates, refers to the tendency of prices to stay around a central mean, thus avoiding long-term non-seasonality; for that reason, geometric Brownian motion is not suitable to represent the behavior of this asset. Once the electricity price moves away from such central mean, the probability that it will return to the starting point is higher than that of further distancing (Borovkova and Permana, 2006).

Electricity price jumps are due to the fact that the short-term variations of the price cannot be efficiently exploited by the agents in the market and the limitations to store massive amounts of energy for long periods of time. As a result of these factors, the efficient market hypothesis, which can be applied to financial markets, cannot be proven in the electricity market. These jumps produce significant increases in skewness and kurtosis, with heavy tails in the price series. Therefore, Poisson jump processes, regime switches, and other types of models are used to represent the stochastic nature of this phenomenon (Borovkova and Permana, 2006; Weron, 2008; Fanone, Gamba, and Prokopczuk, 2013; Knaut and Paschmann, 2019).

Geman and Roncoroni (2003) represent electricity spot price with a mean-reverted diffusion process with a periodic deterministic trend with seasonal effect, where the predictable short-term imbalances in the market are represented with a term of Gaussian white noise. These authors complemented the model proposed by Lucia and Schwartz (2002), also described by Pilipovic (2007), with a component of jumps that captures the asymmetry and kurtosis conditions. Although the results of the models previously mentioned are remarkable, they are limited because the short-term distortions of the price are governed by a normal distribution, which is a serious shortcoming to the validity of these models for representing highly leptokurtic phenomena. An alternative to the rigidity the assumption of normality is the application of semi-nonparametric (SNP) probability distributions, which relax the normality assumption by expanding the normal probability density function (pdf) distribution in a natural way that enables its modelling in terms of as much moments as necessary to capture salient features such as skewness, kurtosis, multimodality or extreme values.

Brunner (1992) showed that SNP statistical techniques are suitable to treat series with heavy tails or evidence of skewness, when normal distributions do not adequately represent the data under study; in addition, they avoid specification errors since Gram-Charlier (Type A) series has been proved to be the asymptotic distribution of any 'regular' pdf. Jondeau and Rockinger (2001) and Gallant, Rossi and Tauchen (1992) used SNP modeling to describe the behavior of the United States stock market. Mauleon and Perote (2000) also used the SNP distribution to model the stock market in the United States and the United Kingdom, while Ñíguez and Perote (2012) did so to evaluate the stock performance of the United States. In the last years this SNP approach has been applied to the modeling of many series – see e.g. Del Brio, Mora-Valencia and Perote (2014) and Cortés, Mora-Valencia and Perote (2016; 2017), but as far as we know and despite its clear advantages, it has only been applied for modeling variables related to electricity markets by Trespalacios, Cortés and Perote, (2019).

This paper proposes the construction of an electricity spot price model that considers key aspects reported in previous literature (seasonality, mean reversion, asymmetry, and high kurtosis) but incorporating a stochastic and mean-reverted SNP process to describe the predictable short-term imbalances instead of the traditionally used Brownian motion differential equation. Furthermore, the model specifies a regime switch as those described by Tsay (2010), where the probability of a regime change is a function of the time each regime is active. These modifications enable a more adequate representation of the spot price in the electricity market, capturing not only the variance but also the skewness, kurtosis, and higher-order moments as relevant features to describe price movements. The application of the model uses monthly series of the spot price of the Colombian electricity market since its start in 1995 until December 2018.

The results show that the SNP representation of the random components of the stochastic process outperforms the normal distribution, whether in typical conditions or those associated to substantial reductions of the hydrologic inflows. Therefore, an adequate measurement of risk in this type of market should be concentrated not only on calculating the standard deviation, skewness, or kurtosis but also on identifying higher-order moments. The activities related to electrical system planning should consider the effect of the occurrence of extreme events (in terms of probability, impact, and length) in order to identify the risk levels that could affect, among other elements, the payback period of the existing investments.

The following section describes the mathematical model proposed in this work, as well as the methodology that was applied. Section 3 describes the data used to calibrate the model, along with some basic concepts of electricity spot price formation in single node systems. In Section 4 the results are presented and analyzed. Finally, Section 5 summarizes the main conclusions and makes recommendations for future works, which could be applied by market professionals or researchers who study related fields.

2. Model and Methodology

2.1. Model

We propose a stochastic process for the spot price and the spot log-price, which is denoted as P_t and defined in Equation (1). Such process has three components: a deterministic one, $F(t)$; and two stochastic elements, X_t and J_t. X_t denotes a process with mean reversion, and J_t can be used to represent switching regimes and jumps:

$$P_t = F(t) + X_t + J_t. \tag{1}$$

The deterministic component, $F(t)$, enables the representation of the expected conditions for price formation, which can be fundamental market indicators such as demand levels, hydrologic inflows, fuel prices, availability of generation plants, and operation of transmission lines, among others. In particular, the deterministic element in the model incorporates three elements, as shown in expression (2): average price level (β_0), deterministic trend ($\beta_{trend} \cdot t$) and seasonal effect ($\sum_{i=1}^{m} \beta_i \cdot G_i$). Their representation has been adapted from the proposal by Lucia and Schwartz (2002) where a dummy variable (G_i) is assigned to each period $i = 1, \dots, m$ which may represent a group of months, as described below:

$$F(t) = \beta_0 + \beta_{trend} \cdot t + \sum_{i=1}^{m} \beta_i \cdot G_i. \qquad (2)$$

Parameter β_{trend} measures the average price growth from time $t - 1$ to time t and captures the average inflationary effect of the market. In turn, the constants β_i capture the average effect on the price (or log-price, depending on the case) that each group of months G_i produces. Variables G_i are defined as dummy variables that take value 1 if the month is in the selected group. For example, if Group 1 gathers the months from January to March, variable G_1 scores 1 in January, February, and March and 0 in the other months. In the case of the modeling of the log-price, parameters β_{trend} and β_i correspond to the elasticity of the price with respect to the time period and the occurrence of each group of months, respectively.

The component X_t , described in Equation (3), corresponds to a stochastic process of mean reversion that it is based on an Ornstein–Uhlenbeck process with two modifications in its random component. The first one is the fact that the magnitude of the random effect depends on the state of variable h_t, which is binary and can only take a values 1 and 0. The second modification affects the stochastic differential dz, which is not normal distributed, unlike the Brownian motion differential equation. This work considers that dz is described by a SNP pdf, see Section 2.2 for further details. Parameter κ in (3) is known as the mean reversion speed:

$$dX_t = -\kappa\, X_t\, dt + (1 - h_t)\, dz, \qquad (3)$$

$$dz \sim SNP(\mu_e, \sigma_e | \boldsymbol{d_e}).$$

Finally, component J_t denotes the effect of the regime switch h_t on the price. In cases where $h_t = 1$, the expected price increases the value of D, and the short-term random switches are governed by the random variable j_t. It is assumed that j_t is also described by an SNP probability density function:

$$J_t = h_t \cdot (D + j_t), \qquad (4)$$

$$j_t \sim SNP(\mu_j, \sigma_j | \boldsymbol{d_j}).$$

The regime switch is governed by the exogenous variable h_t. The specification of this term follows the Markov Switching Model, described by Tsay (2010) and the approximation in terms of prediction of electricity price jumps by Mount, Ning, and Cai (2006). The binary variable h_t has transition possibilities w_0 and w_1, which depend on the cumulative duration of each regime, as shown in equations (5) and (6):

$$w_0 = P\left(h_{t+1} = 0 \middle| \{h\}_{t-\tau_1}^{t} = 1\right) = \frac{1}{1-e^{-(\lambda_0^0+\lambda_1^0\cdot\tau_1)}}, \tag{5}$$

$$w_1 = P\left(h_{t+1} = 1 \middle| \{h\}_{t-\tau_0}^{t} = 1\right) = \frac{1}{1-e^{-(\lambda_0^1+\lambda_1^1\cdot\tau_0)}}, \tag{6}$$

where τ_0 and τ_1 are the number of periods where variable h_t remained unchanged in state 0 and 1, respectively. λ_0^0 λ_1^0 (λ_0^1 λ_1^1) are the parameters for the probability of transition from state 1 to 0 (0 to 1). This parametrization of the switching regime allows the modeling of events with both long and short durations. The parameters for w_0 and w_1 are estimated using the maximum likelihood method applied to a logit model.

2.2. SNP distribution

The most important contribution of this work is the incorporation of SNP functions into the components for price uncertainty and the two proposed regimes, either $h_t = 0$ or $h_t = 1$.

A standardized variable y exhibits a SNP distribution if its pdf is given by Equation (7), i.e. can be expressed in terms of the pdf of the standard normal, $\phi(y)$, and a weighted sum of its derivatives or the so-called Hermite Polynomials (HP), $H_s(y)$.

$$f(y) = \left[\sum_{s=0}^{\infty} \delta_s \cdot H_s(y)\right] \phi(y), \tag{7}$$

$$H_s(y) = \frac{(-1)^s}{\phi(y)} \cdot \frac{d^s\phi(y)}{dy^s},$$

$$\phi(y) = \frac{1}{2\pi} e^{-\frac{y^2}{2}}.$$

The HP form an orthonormal base that guarantee that the expansion in equation (7) is a pdf whose first *s* moments depend on the first δ_s parameters. These parameters may be expressed as $\delta_s = \frac{1}{s!}\int_{-\infty}^{\infty} H_s(y) f(y)\, dy$ and the HP can be recursively obtained, being $H_0(y) = 1$, $H_1(y) = y$, $H_2(y) = y^2 - 1$, $H_3(y) = y^3 - 3y$, $H_4(y) = x^4 - 6x^2 + 3$, $H_5(y) = x^5 - 10x^3 + 15x$, etc. For empirical purposes the expansion must be truncated at a finite degree *n* as in equation (8), and the vector $\boldsymbol{d} = (d_1, d_2, \dots, d_n) \in \mathbb{R}^n$ must ensure that $g(y; \boldsymbol{d}) \geq 0$.

$$g(y; \boldsymbol{d}) = [1 + \textstyle\sum_{s=1}^{n} d_s \cdot H_s(y)]\phi(y). \tag{8}$$

The SNP can be also characterized by the cummulative distribution function (cdf), which can be expressed in terms standard normal cdf, $\Phi(y)$ (see e.g. Cortés, et al., 2016):

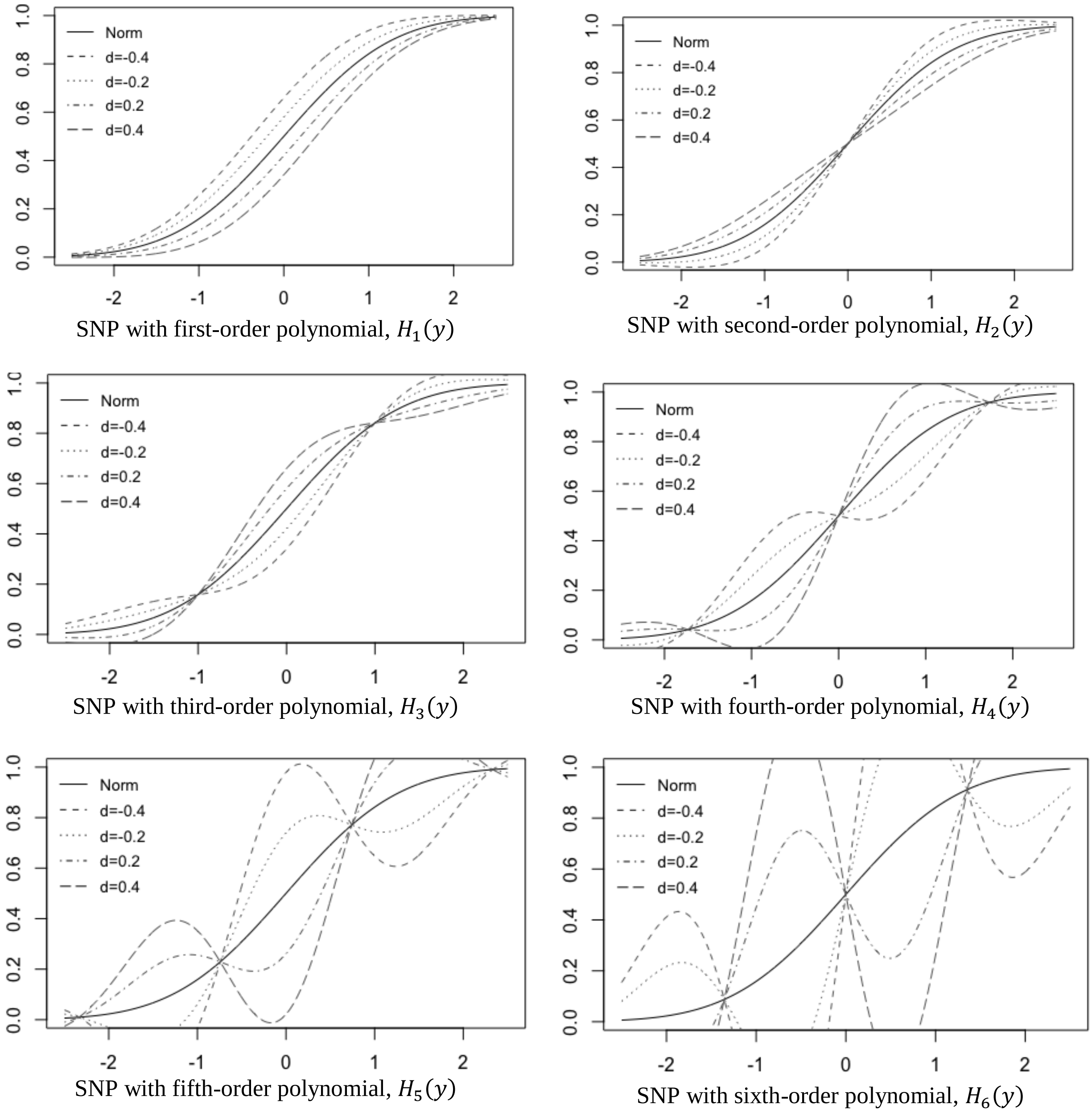
$$G(y|\boldsymbol{d}) = \Phi(y) - \phi(y) \sum_{s=1}^{n} d_s H_{s-1}(y). \quad (9)$$

The marginal effect of each one of the Hermite polynomials (from the first to the sixth order) on the cumulative standard normal distribution (CDF) is illustrated in Figure 1, which also shows the sensitivity to different values of δ_s. For example, the second plot in the first column in Figure 1 represents the function in (9) with $s = 3$; that is, the effect of incorporating only the third-order HP for different values of δ_3 (–0.4, –0.2, 0.2, and 0.4). When the normal distribution is compared to the SNP distribution, it can be seen that even order polynomials increase the slope of the cumulative distribution. The odd components modify the skew of the probability density function, shifting the cumulative function to the right (left) if the value of δ_s is positive (negative); this effect is more pronounced in the first-order polynomial.

We identified that the higher the order of the polynomial, the greater the magnitude of its effect for the same values of δ_s. The sixth-order HP has a higher impact on the shape of the cdf than its second-order counterpart. As the order of the polynomial increases, so does the marginal contribution of coefficient δ_s, e.g. when the value of δ_5 changes from 0.2 to 0.4 there is a more significant impact on the cdf than when δ_1 changes from 0.2 to 0.4. Those conditions on the marginal effects due to both the order of the polynomial *s* and the value of the its parameter δ_s lead the analyst to indentify the relevant components of the SNP distribution when fitting a particular dataset.

The condition of non-negativity of the Gram–Charlier series expansion has traditionally studied on the basis of the pdf representation. Figure 1 illustrates that the non-positivity issues also bring along conflicts in the cdf. For instance, depending on the polynomial order and the value of δs, some functions are not monotone increasing, and their values could be higher than 1 and even lower than 0; such conditions are not adequate for the characterization in terms of neither the pdf nor the cdf. This shortcoming of truncated Gram–Charlier series was first stated by Barton and Dennis (1952) and has been addressed in different ways in the literature. One possible solution lies in the use of maximum likelihood (ML) estimation and the selection of initial values that guarantee that convergence was not achieved in local optima.

Figure 1. Cumulative Probability Function for SNP and standard normal distributions.



SNP with first-order polynomial, $H_1(y)$

SNP with second-order polynomial, $H_2(y)$

SNP with third-order polynomial, $H_3(y)$

SNP with fourth-order polynomial, $H_4(y)$

SNP with fifth-order polynomial, $H_5(y)$

SNP with sixth-order polynomial, $H_6(y)$

Each figure compares the standard normal and SNP distributions (both with zero mean and unit variance) in terms of their respective cdfs. The normal cdf is depicted in a continuous line and the SNP cdfs with discontinous lines. Each plot depicts the marginal effect on the SNP cdf by considering only one HP and for four different values of its corresponding parameter δ_s={−0.4, −0.2, 0.2, 0.4}. The figures illustrate the sensitivity of the SNP cdf to the coefficients of the HPs.

2.3. Estimation

Equation (10) rewrites the model for spot price (P_t) or its logarithm (lnP_t) by grouping the different stochastic terms in a single component. It also extends the deterministic function $F(t)$, and considers h_t as a dummy or categorical variable, whose average effect is captured through constant D.

$$P_t = \beta_0 + \beta_{trend} t + \sum_{i=1}^{m} \beta_i G_i + D h_t + Y_t \tag{10}$$

$$Y_t = (1-\kappa) Y_{t-1} + (1-h_t)\epsilon_t + h_t j_t$$

The estimation is performed in several stages: Firstly, the parameters ($\boldsymbol{\beta}$) of the deterministic variables; secondly, the stochastic autoregressive component represented by Y_t. The residuals of this autoregressive process are governed by two different random variables, ϵ_t and j_t, which are independent and identically distributed as SNP random variables and have an effect on the price according to the value of h_t.

The parameters of the SNP density of ϵ_t and j_t (namely, $\boldsymbol{d}_\epsilon$ and $\boldsymbol{d}_j$) are estimated by implementing ML algorithms, recursively selecting the appropriate truncation order and the significant parameters according to linear restriction tests. As the normal is nested in the SNP, the best specification can be directly tested from the Likelihood Ratio (LR) in equation (11), which compares the loglikelihood under the normal (l_{normal}) and SNP (l_{SNP}). Under the null hypothesis of normality, the statistic follows a Chi-squared distribution with degrees of freedom equal to the number of parameters of the SNP expansion (*n*).

$$LR = -2(l_{normal} - l_{SNP}) \sim \chi^2_n, \tag{11}$$

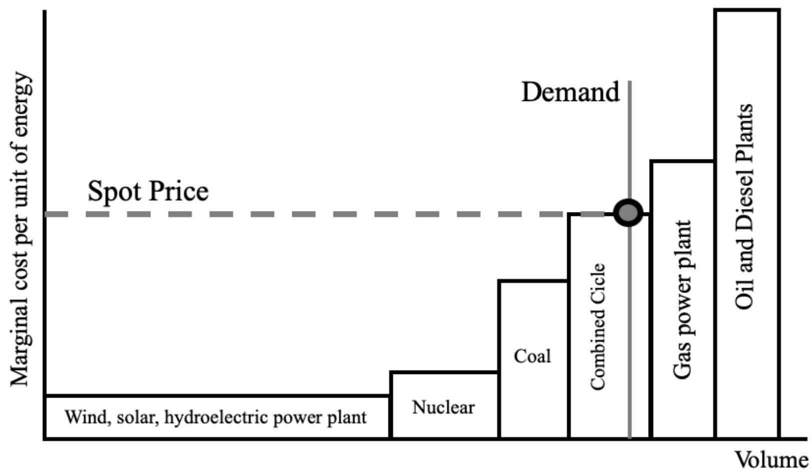
For the case of Colombian spot prices, the process h_t depends on the occurrence of El Niño phenomenon, which triggers the price increases. Consequently, its parameters are estimated with a Logit model and its fit is verified through the Hosmer–Lemeshow test. Finally, parameter *D* is the coefficient of linear regression that accompanies the occurrence of El Niño phenomenon, while components j_t and ϵ_t affect price formation depending on the occurrence or not of said phenomenon.

3. Electricity market and description of the data

The Colombian electricity market started operations in June 1995 after a series of legislative reforms that, among other aspects, separated the supply chain, enabled the participation of agents with private capital (in a sector that used to be dominated by state-owned companies), and created planning, regulation, supervision, and control mechanisms in order to maintain the safety and reliability of the electricity supply for the entire population. The activities in the sector are classified as generation, transmission, distribution, and commercialization.

Each one of them adopts a methodology to calculate the electricity rate that end users pay. The reference price is provided by the electricity market (or spot market), where the value of generation is calculated. This is the only activity, in Colombia, that takes place in an environment of free competition, along with commercialization in some market segments. In such market, different types of electricity generation technologies can participate with multiple production costs per unit of energy, as illustrated in Figure 2.

Figure 2. Spot price formation.



The figure shows the clearing spot price in a hypothetical market with several types of generation technology. Note that the most expensive resources can cause the spot price of electricity to present extreme events and a positive skewness. The figure is based on Bahar and Sauvage (2013).

Colombia has a single-node price system with only one reference price for the generation of electricity in the entire country. The situation is different in other markets, such as Chile or Mexico (also in Latin America), where there are multiple reference prices, one for each node in the system. The spot price in Colombia is defined with an hourly resolution, i.e a price associated with each hour of the day, and it is the result of a sealed-bid auction, where the bidders are the generators and the demand is passive. Every day, generation companies bid a price for the next day and declare the availability of each of their generation assets at each hour of the day. The spot price is defined as the price of the last generation resource that is necessary to meet the demand of the market, as illustrated in Figure 2. Thus, in electricity markets, the price is determined by the expectation of availability and the transaction price of the primary commodities required for generation: Hydrologic inflows, reservoir levels, winds, solar radiation, and the supply and transportation of coal, natural gas, and diesel fuel. Likewise, the behavior of the demand for electricity (boosted by the economic activity or extreme climate events such as El Niño phenomenon) determines the movements in the price series as well.

This study considers information of the price in the Colombian electricity market since its beginnings in 1995 until December 2018, with a monthly basis.[5] In Table 1, Panel A presents the descriptive statistics of the average monthly market price (spot price P_t) of electricity in Colombia and its transformations: first difference (ΔP_t), log-price (lnP_t) , and logarithmic return ($\Delta\ lnP_t$). Likewise, we considered different panels: Panel A analyzes the statistics for the complete series of prices and for the periods that either experienced El Niño phenomenon

[5] Information downloaded from Portal BI of the Colombian electricity operator, www.xm.com.co.

($h_t = 1$), or not ($h_t = 0$); Panel B displays the autocorrelation function for the complete series.

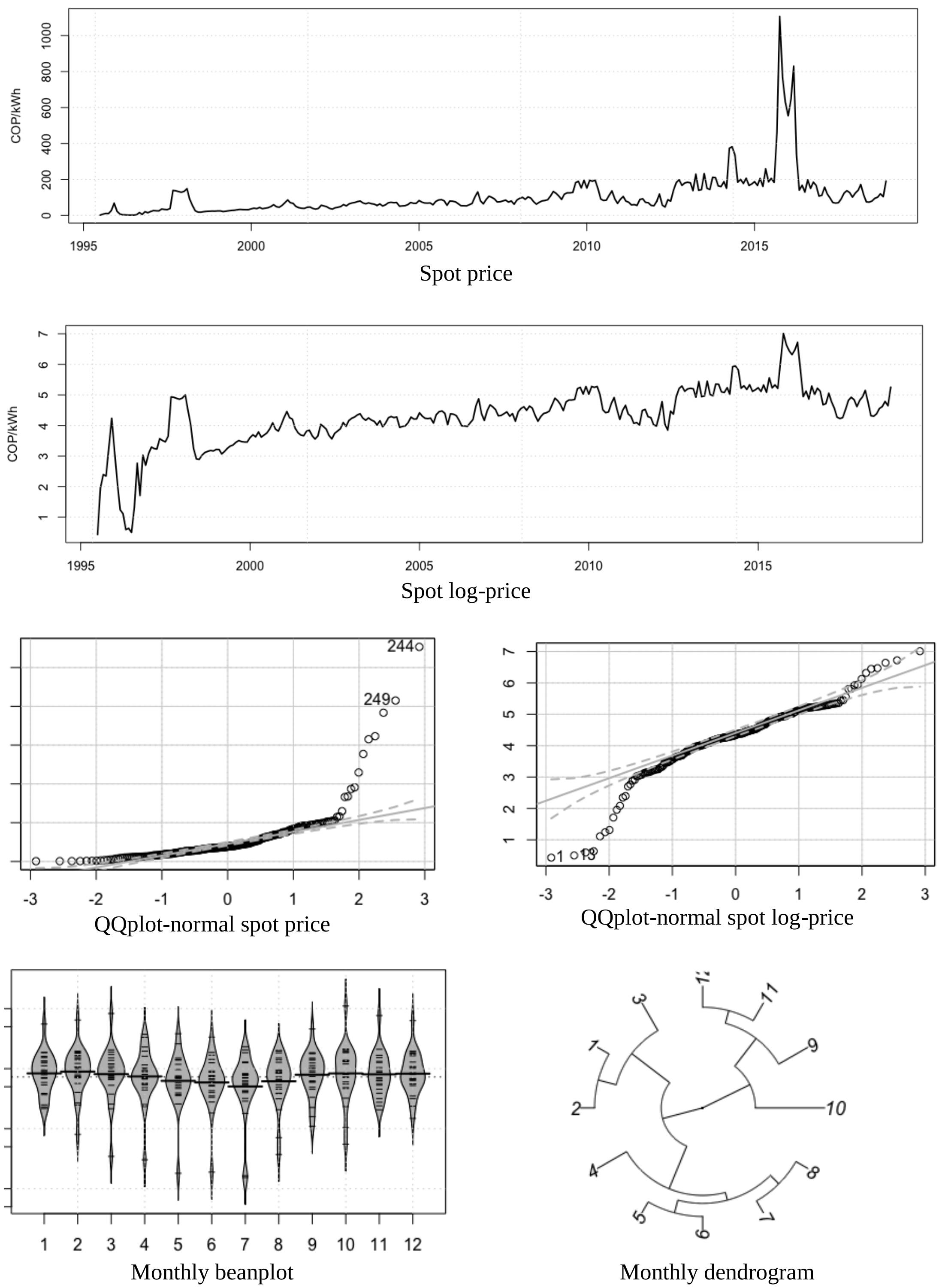
Table 1. Descriptive statistics of the electricity spot price in Colombia.

Panel A									
Series	**Mean**	**Std**	**Skewness**	**Kurtosis**	**P5%**	**P25%**	**P50%**	**P75%**	**P95%**
Period 1995-2018									
P_t	107.3	119.5	4.5	29.7	18.0	50.3	76.4	133.0	211.9
ΔP_t	0.3	94.3	-4.9	55.6	-82.7	-12.4	0.8	16.0	92.6
lnP_t	4.29	0.96	-1.06	6.43	2.89	3.92	4.34	4.89	5.36
$\Delta\ lnP_t$	0.02	0.34	0.71	7.63	-0.48	-0.13	0.02	0.15	0.48
Periods in which h_t=1									
P_t	247.1	265.8	1.8	5.2	65.0	76.8	139.0	195.5	792.2
ΔP_t	23.79	150.83	2.06	11.40	-103.38	-13.55	0.49	42.72	240.04
lnP_t	5.10	0.85	0.82	2.58	4.17	4.34	4.93	5.28	6.67
$\Delta\ lnP_t$	0.08	0.36	1.25	4.35	-0.32	-0.13	0.01	0.18	0.80
Periods in which h_t=0									
P_t	88.76	64.13	1.54	6.52	15.28	45.40	73.42	119.11	199.31
ΔP_t	0.76	34.02	0.41	16.59	-44.40	-8.68	0.92	10.31	40.57
lnP_t	4.18	0.92	-1.48	6.66	2.73	3.82	4.30	4.78	5.29
$\Delta\ lnP_t$	0.02	0.32	0.74	8.19	-0.42	-0.13	0.02	0.16	0.47
Panel B									
Autocorrelation Function (ACF) by lag order. Data from 1995 to 2018									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
P_t	0.846	0.702	0.631	0.596	0.561	0.421	0.341	0.313	0.303
ΔP_t	-0.025	-0.242	-0.120	0.005	0.348	-0.198	-0.175	-0.060	-0.037
lnP_t	0.907	0.836	0.773	0.715	0.673	0.650	0.625	0.601	0.581
$\Delta\ lnP_t$	0.030	-0.025	-0.038	-0.010	-0.032	-0.124	-0.137	-0.140	-0.016

Panel A presents descriptive statistics for the electricity spot price of the Colombian market and some transformations (differences, $\Delta P_t = P_t - P_{t-1}$, and natural logarithm, lnP_t). Statistics are shown for the complete series (1995-2018) and for the periods when the variable shifted regimes, depending on either the ocurrence El Niño phenomenon ($h_t = 1$) or not ($h_t = 0$). Panel B shows the autocorrelation values of the series under analysis.

The monthly series of electricity prices in Colombia (see Figure 3) presents an upward trend since the start of the market, with values ranging between 1.54 COP/kWh in July 1995 and 1106 COP/kWh in October 2015, and an average value of 107.3 COP/kWh. The spot log-price presents a smoother pattern, with an average of 4.29, and a range between 0.4312 and 7.0. The *qqplot* diagrams highlight the deviation of the sample with respect to the normal distribution, mainly at the tails, confirming the heavy tails featured by this type of commodities. This fact supports the modeling of the random components with SNP distributions. The beanplot in Figure 3 shows that the price is seasonal; between May and August, the average values are below the overall mean of the data. Such seasonality is also reflected in the dendrogram that groups in one segment the months from April to August; in another segment, January to March; and in the last group, from September to December. The effects of seasonality are not only visible in the average value of the series, they are also reflected in jumps in the shape of the distribution of probability. While June and July have a negative skewness more intense than that of other months of the year, October presents the highest extreme values.

Figure 3. Spot price and the spot log-price in Colombia.



The figure presents the series and qqplot diagrams for the spot price and the spot log-price in Colombia from 1995 to 2018, as well as the beanplots and dendrogram of the spot price.

An analysis of the first four moments of the spot price series and its transformations reveal features that are different from those produced by a normal distribution. All the transformations present asymmetry, with a skewness different from 0, and leptokurtosis, with a kurtosis above 3. In the case of the price, the right-side tail of the distribution is larger than its left-side counterpart: the 95th (5th) percentile is at a distance of 136 (58) COP/kWh from the 50th percentile. Regarding the periods when $h_t = 1$, the average of the series was found to be consistently higher than for the total of periods of time. During those periods of regime switching, the price series has a higher standard deviation with lower skewness and kurtosis. Therefore, the data of the prices during El Niño are closer to a normal distribution than those of the entire sample.

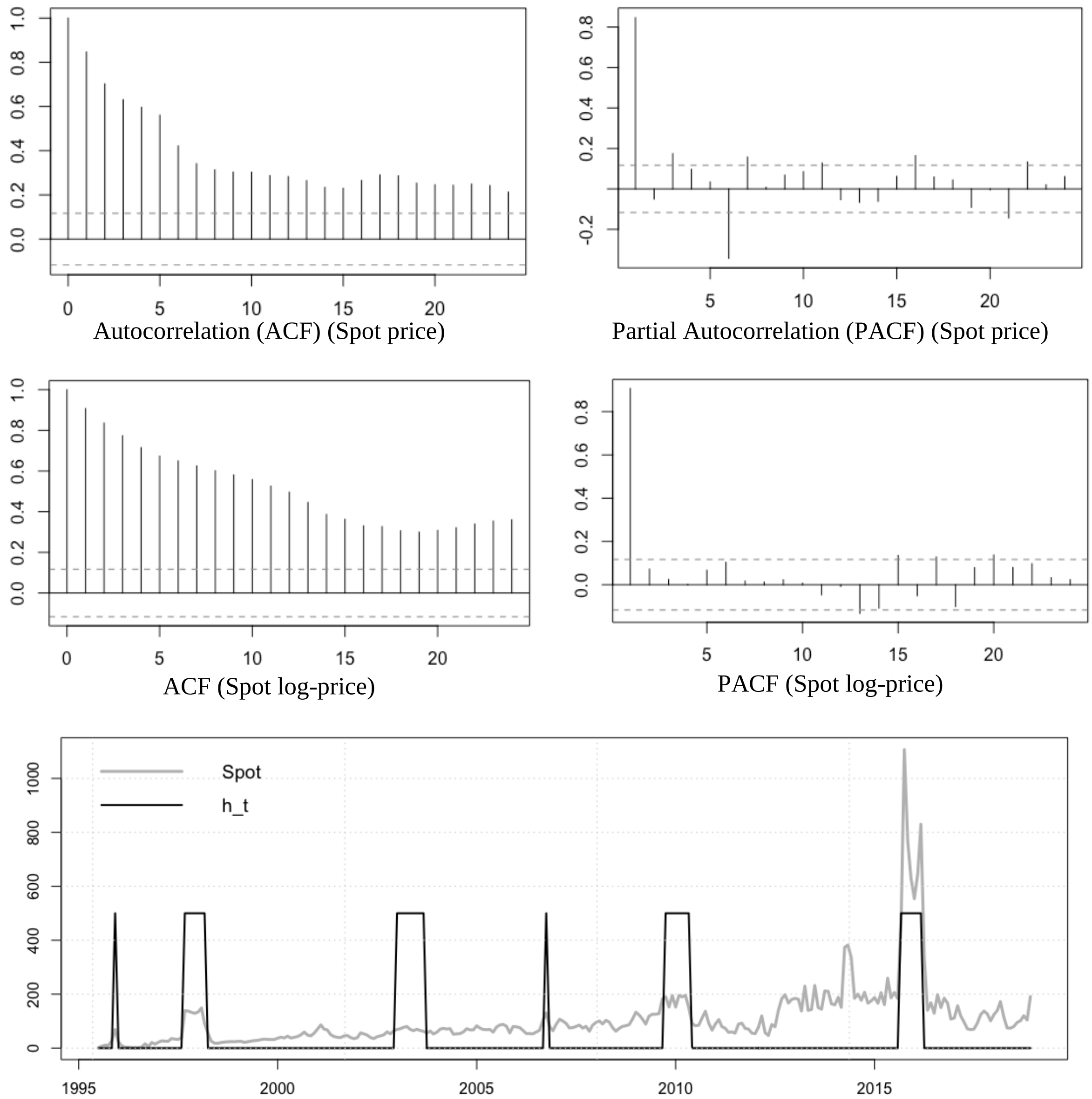
Panel B in Table 1 also lists the autocorrelation function for the proposed price transformations. Noticeably, the autocorrelation levels of the differentiated series are lower, suggesting that the series of price and log-price may contain autoregressive components. Similar observations can be made about the autocorrelation function (ACF) and the partial autocorrelation function (PACF) in Figure 4. The autocorrelation functions show a strong first-order relationship and quarterly, biannual, and annual seasonality. The incorporation of dummy variables for each one of the months enables to capture such a seasonal pattern, while the use of the Ornstein–Uhlembeck process that defines the process for X_t incorporates the effect of the first-order lag.

The presence of jumps in electricity prices is mainly due to the physical limitations of the storage of this commodity. For that reason, variations in demand, fuel availability, network stability, generation equipment capacity, hydrologic inflows, wind speed, or solar radiation, among others, are directly reflected in in the price of short-term transactions. If an agent buys a significant amount of electricity today in order to sell it in subsequent days when a positive jump of the price occurs, it could obtain a net profit in the operation and, at the same time, it could smooth the price series, increasing the price today and reducing the price in the future moment of the jump. Therefore, the adequate conditions to meet the assumptions of the efficient market hypothesis would be involved in the electricity market. Such hypothetical transaction can be completed as long as the agent can store the energy between today and the moment of the price jump.

The exogenous variable that affects the electricity price in Colombia the most (and some other countries in Latin America) is the variation in the hydrologic pattern due to El Niño Southern Oscillation (ENSO). ENSO is an irregular periodic variation of the winds and the sea temperature over the Pacific Ocean that mainly impacts the South American Pacific coast. It is associated with interannual precipitation variations and it presents long-term seasonality that lasts decades. The phase in which the sea temperature increases is known as El Niño, while the cooling phase, as La Niña.

Poveda and Mesa (1997) showed the way ENSO affects the climate and hydrological conditions in Colombia, changing rainfall levels and their corresponding contribution to hydraulic generation plants. As a result, the share of hydraulic generation can decrease from 80% in typical conditions to 50% during the warming phase. Moreover, the occurrence of El Niño is accompanied by an increase in the air temperature in Colombia, which also promotes the operation of cooling equipment and irrigation systems, which produce a subsequent increase in the demand for electricity. Thus, El Niño phenomenon reduces the supply and increases demand, pushing the electricity spot price upward. Despite the chaotic nonlinear dynamics of ENSO, a probabilistic forecast can be produced based on an a priori classification of the atmospheric conditions (Waylen and Poveda, 2002).

Figure 4. Autocorrelation of the electricity spot price in Colombia and regime switches (h_t).



Autocorrelation (ACF) (Spot price)

Partial Autocorrelation (PACF) (Spot price)

ACF (Spot log-price)

PACF (Spot log-price)

Regime switch moments, h_t

This figure presents the Autocorrelation Functions (ACF) and Partial Autocorrelation Functions (PACF) of series of the spot price and its logarithm. It also illustrates the moments at which the price has been affected by El Niño phenomenon and correspond to $h_t = 1$. Those periods do not exactly match to the moments when the phenomenon occurs, but to the periods when the market presented significant price movements; they were adjusted as a result of conversations with professionals in the sector.

The variable that governs the regime switch h_t, is constructed in accordance with the periods in which the electricity price in Colombia has been affected by the occurrence of El Niño phenomenon, as measured by the series Oceanic Nino Index (ONI), available on the website of the US National Oceanic and Atmospheric Administration (NOAA).[6] Among the moments when $h_t = 1$, the greatest impact in terms of mean and variance occurred in 2015. In that year, not only a strong El Niño occurred, but also this event coincided with a time when the thermal generation park was intensively substituting natural gas (with a relatively low variable production cost) with liquid fuels, oil or diesel (with a high variable production cost). As explained in the next section, the regime switch affects all features of the random component of the price: mean, variance, skewness, kurtosis, and higher-order moments.

4. Results and discussion

The estimation of the proposed model is implemented in three stages. First, we estimate the deterministic component; then, the conditional mean; and finally, the SNP distribution over the error terms in the two possible states of nature, defined by h_t. This section presents the estimation of the different processes and describes the results obtained in the simulations of price and log-price.

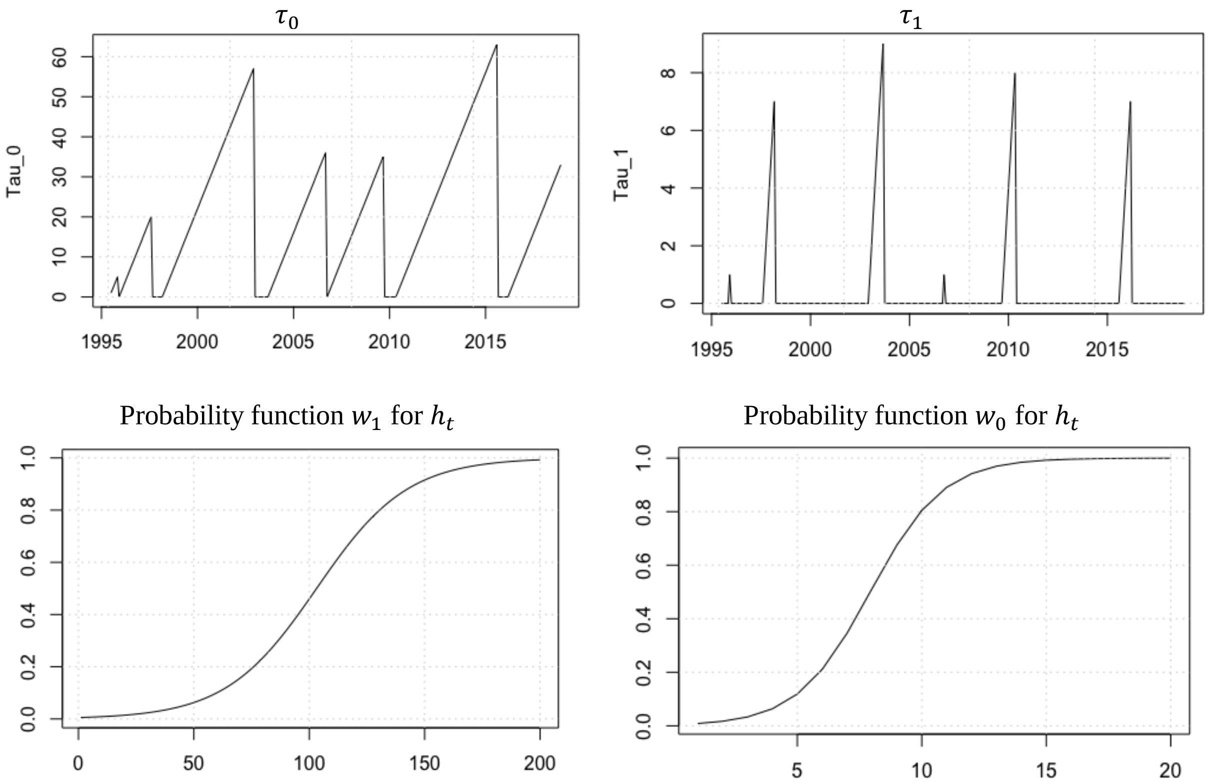
4.1. Regime switch

Variable h_t was modeled as a function of times τ_0 and τ_1. The historical values are shown in Figure 5, as well as the probabilities w_1 and w_0. It was found that, as time is accumulated in one of the states for h_t, the probability of a change of state increases. Probability w_0 is sensitive to the passage of time, e.g. there is a probability of 51% that the regime will change from $h_t = 1$ to $h_t = 0$ after 8 months in the state $h_t = 1$. On the other hand, w_1 captures the probability of transition from $h_t = 0$ to $h_t = 1$, and there is a 50% probability of this transition after 104 months in the state $h_t = 0$. Once the impact of El Niño phenomenon on the electricity spot price in Colombia has started, it is expected to be over in a matter of months. Regarding the maximum period that a regime could be maintained, we can observe that, after 13 months of variable h_t equaling 1, the probability that the variable will take a value of 0 is 97%. After 162 months of variable h_t in 0, the probability that the variable will change to 1 is 95%.

The probability functions in Figure 5 were constructed based on Equations (5) and (6) with the parameters in Table 2. The Hosmer–Lemeshow (HL) test does not enable us to reject the null hypothesis of correct specification; for that reason, the logit version proposed for these probability distributions is considered relevant to represent variable h_t.

[6] www.noaa.gov

Figure 5. Regime switch in Colombia h_t.



On the y-axis: probability, on the x-axis: τ_0 On the y-axis: probability, in the x-axis: τ_1

The figure shows the parameters of time used to calculate the probability of regime switch of h_t. The cumulative time that h_t equals 0 is τ_0, while the cumulative time that h_t equals 1 is τ_1. It also shows the function w_1 as a function of τ_1 and w_0 as a function of τ_1, which measure the probability that the value of h_t changes from 0 to 1 and from 1 to 0.

4.2. Deterministic function and mean reversion

The deterministic component, represented by $F(t)$, evidences the effects of the trend and seasonality of the spot price on the model of the spot price as well as of the spot log-price. The coefficients of trend and monthly grouping (β_{trend} and β_g, respectively) present confidence levels above 99%, except for parameter β_g of the model of spot price, whose confidence level exceeds 90%. According to these results, the group of months from May to August has an average price below the other months of the year.

The component X_t captures the effect of the mean reversion of the price in the model of the spot price as well as the spot log-price. The mean reversion speed κ is estimated using Equation (10), and $(1-\kappa)$ was found to have a confidence level above 99%. κ has an estimated value of 0.2783, which guarantees the seasonality of the process described above with an expected value in the long term equal to 0. Therefore, in the long term, the expected price is a linear combination weighed by the probability w_1, between $F(t)$ and D. This result demands additional reflection on the price conditions in the long term because it suggests that the long-term price expectation is adaptative. For that reason, the formation of the prices of derivatives (such as forward contracts, options, and swaps) should be, at each instant in time *t*, redefined as a function of the expectations associated with w_1.

The proposed model also enables the evaluation of the impact of the occurrence of macroclimatic phenomena on the price. In our application the occurrence of El Niño means an average increase in price of 163.52 COP/kWh. This value corresponds with the estimated value for parameter D and, it also corresponds with an increase in standard deviation in the random component, from 34.82 to 147.12 COP/kWh.

Table 2. Parameters of the price model. For spot price and spot log-price.

			Spot price			Spot log-price	
Component	Parameter		Coef	p-value		Coef	p-value
$F(t)$	β_0		-17.32	0.128		3.0002	< 0.001
	β_{trend}		0.7908	< 0.001		8.88E-03	< 0.001
	β_g	*May* *Jun* *Jul* *Aug*	-19.26	0.083	*May* *Jun* *Jul* *Aug*	-0.2443	< 0.001
X_t	$(1-\kappa)$		0.7217	< 0.001		0.8448	< 0.001
	κ		0.2783			0.1552	
	ϵ_t	μ_e	0.7056			0.02146	
		σ_e	34.82			0.3036	
		$d3$	0.0858	0.017	$d4$	0.1792	< 0.001
		$d4$	0.1322	< 0.001	$d6$	0.0297	0.0022
					$d8$	0.0026	0.0217
J_t	D		163.52	< 0.001		0.9750	< 0.001
	j_t	μ_j	-6.4532			-0.07155	
		σ_j	147.12			0.3235	
		$d3$	0.4061	< 0.001	$d4$	0.1425	< 0.001

Component	Parameter		Coef	p-value	HLTest*	
h_t	w_1	Lo	-5.2222	< 0.001	p-value:	0.7554
		l1	0.0502	0.033		
	w_0	Lo	-5.4283	< 0.001	p-value	0.9991
		l1	0.6846	< 0.001		

The table presents the parameters of the model proposed for the spot price of electricity in Colombia and the spot log-price. We included information with a monthly frequency from June 1995 to December 2018. To calculate the variable with regime switch h_t, we estimate two logit models: one for the moments at which h_t changes from 0 to 1, and one for the moments when it changes from 1 to 0. *The null hypothesis of the Hosmer–Lemeshow (HL) goodness-of-fit test is correct specification.

The variables that describe the uncertainty in the short term of the spot price (in both levels and logarithms) are ϵ_t and j_t with an effect on the price depending on the regime, $h_t = 0$ or $h_t = 1$, respectively. Table 3 presents the descriptive statistics, the fitted parameters of the SNP models and the LR test for the jointly significance of the models (i.e. normality *versus* SNP). Both series present heavy tails, with a kurtosis above 4, which is more significant for the model for spot price than for its logarithm. Nevertheless, the situation is different with respect to skewness: ϵ_t and j_t have a positive (negative) skew for the case of the spot price (spot log-price). In all cases the normal is clearly rejected for capturing short-term uncertainty.

The regime switch modifies the uncertainty conditions. The occurrence of El Niño phenomenon has a clear impact on the average, variance, skewness, and kurtosis of the price. For both the spot price and its logarithm, the absolute value of the variance and skewness increase with the presence of the phenomenon. However, the kurtosis exhibits a different pattern depending on the modeling of spot price or the spot log-price: the kurtosis of j_t is higher (lower) than that of ϵ_t in the former (latter).

All this evidence about the strong impact of the regime switch should be considered by the agents that participate in the market, so that the long-term energy supply, as well as the financial sustainability of the required investments, will be guaranteed. As a result, the investors will be able to make decisions about market expansion, structure, and performance by incorporating risky conditions (systemic or idiosyncratic) and their possible treatment in case of different conditions (or regimes).

4.3. Random components

The regime switch has a deep impact on the moment structure of the random component of price, measured by ϵ_t in normal conditions and j_t under the presence of El Niño. During the regime of El Niño standard deviation, skewness, and kurtosis increase along with the increase of uncertainty and risk. These facts support the need for capturing all these features with flexible representations as those provided by our model.

Table 3 displays the results for the estimation of both random components for the electricity spot price and its logarithm. Panel A lists the descriptive statistics for ϵ_t and j_t. Panel B presents the *ds* components fitted for the SNP distribution; and Panel C, percentiles with three types of measurements: empirical percentiles and fitted percentiles under both SNP and normal distributions.

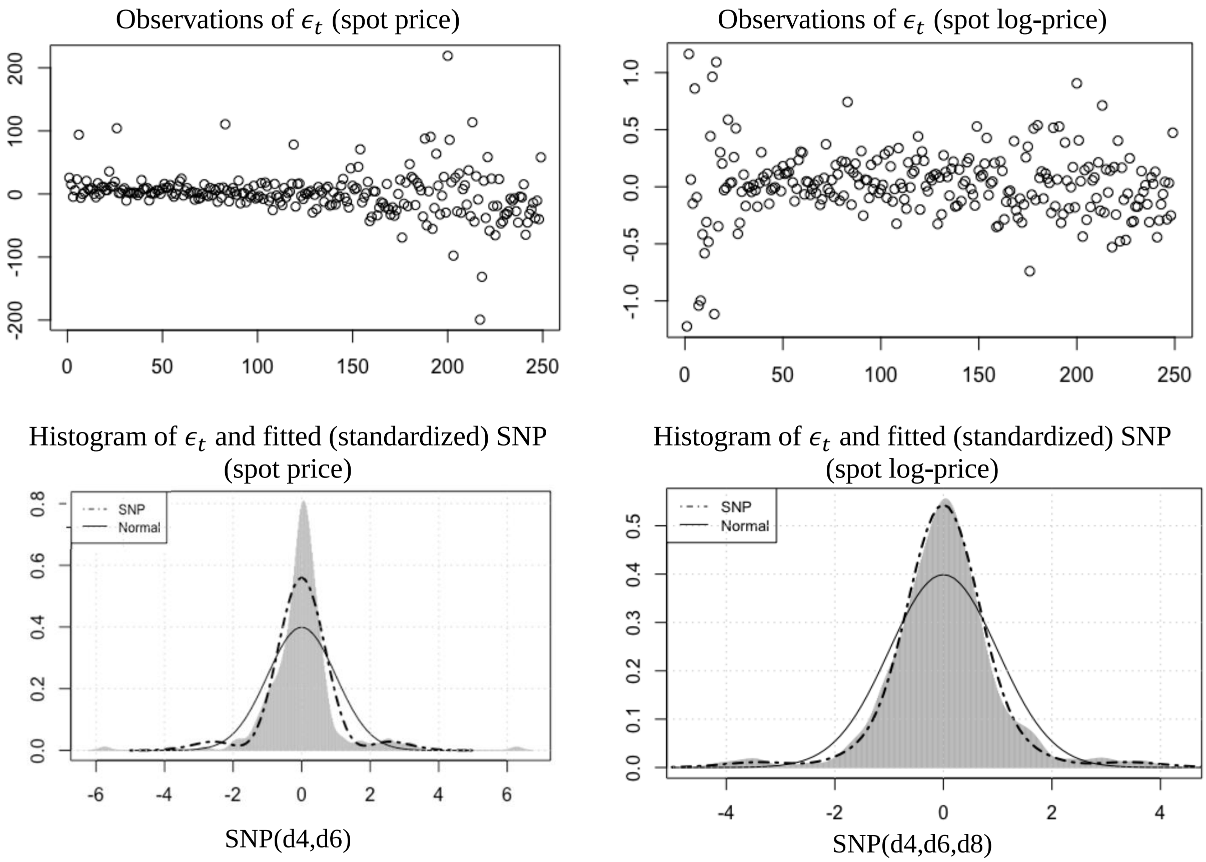
On the basis of the results in Table 3 (and Figures 6 and 7), we found that a SNP distribution can more appropriately describe the behavior of variables ϵ_t and j_t than a normal distribution. The model of the spot price ϵ_t presents a SNP distribution with parameters $(d_3, d_4) = (0{,}858\,, 0{,}1322)$, and j_t presents a SNP distribution with parameter $(d_4) = (0{,}4061)$, which are statistically significant for a confidence level of 99%. The likelihood-ratio test shows that, for a confidence level of 99%, the SNP outperforms the normal for modeling both ϵ_t and j_t. Figure 7 illustrates how the flexibility of the SNP specification describes the particularities of the right-side tail of the probability distribution of variable of j_t.

Table 3. Descriptive statistics and SNP fit for the random components of X_t and J_t.

		Panel A. Descriptive statistics				Panel B. Estimation						Panel C. Percentile					
						SNP (normalized series)											
		Mean	Std dev	Skewness	Kurtosis	di	d3	d4	d6	d8	LR		5%	25%	50%	75%	95%
Spot	ϵ_t	0.71	34.82	0.48	14.36	coef	0.0858	0.1322			629	Obs (1)	-44	-13	1	12	54
						p value	0.017	< 0.001			< 0.001	SNP (2)	-44	-18	-1	16	71
												Norm (3)	-57	-23	1	24	58
	j_t	-6.45	147.12	3.37	15.86	coef	0.4061				15.7	Obs	-136	-64	-34	-7	197
						p-value	< 0.001				< 0.001	SNP	-197	-125	-57	33	349
												Norm	-248	-106	-6	93	236
Log Spot	ϵ_t	0.02	0.30	-0.09	6.60	coef		0.1792	0.0297	0.0026	56.6	Obs	-0.41	-0.13	0.03	0.16	0.51
						p-value		< 0.001	0.0022	0.0217	< 0.001	SNP	-0.41	-0.13	0.02	0.17	0.45
												Norm	-0.48	-0.18	0.02	0.23	0.52
	j_t	-0.07	0.32	-0.18	5.00	coef		0.1425			9.30	Obs	-0.74	-0.23	-0.03	0.10	0.24
						p-value		< 0.001			0.002	SNP	-0.52	-0.23	-0.07	0.08	0.38
												Norm	-0.60	-0.29	-0.07	0.15	0.46

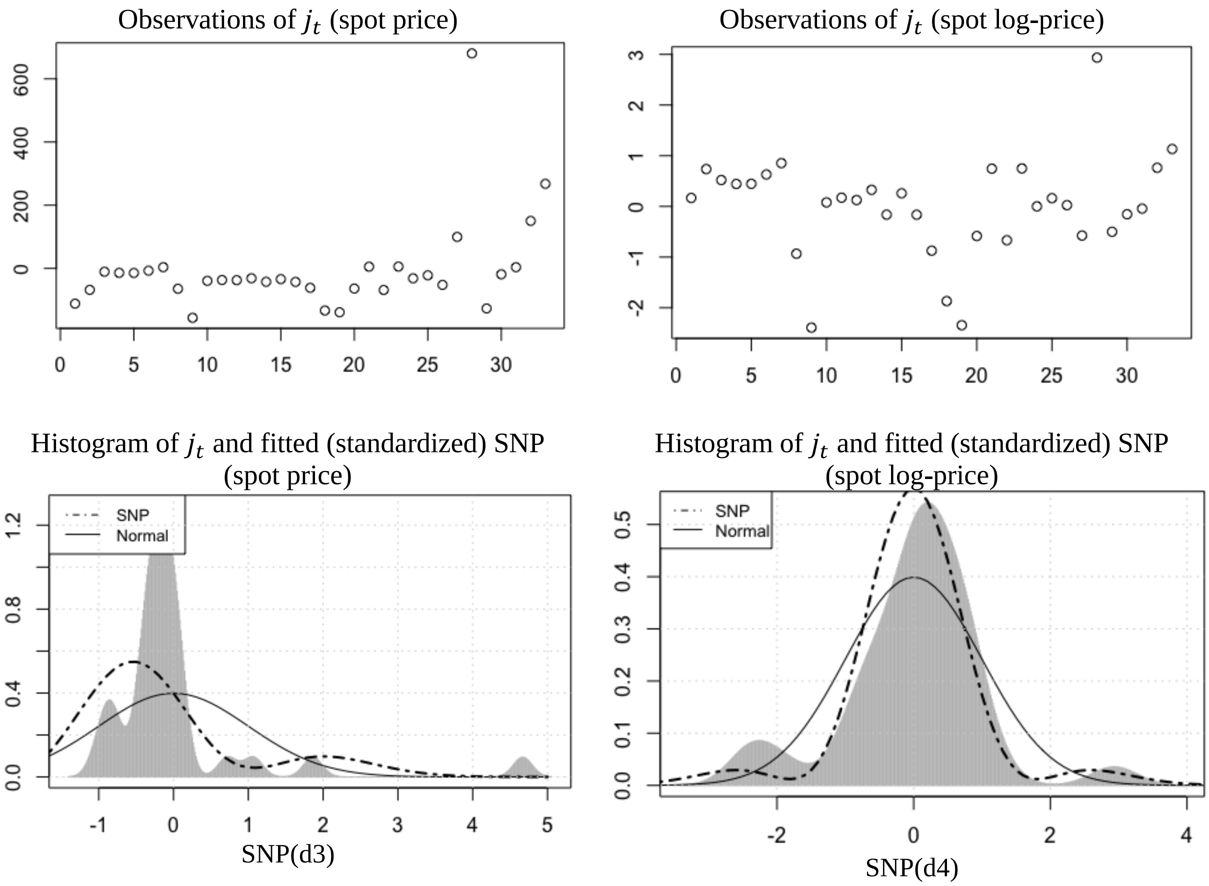
This table describes the random components of the spot price and its logarithm. Panel A presents the descriptive statistics for ϵ_t and j_t. Panel B presents the estimates for the d_s parameters of the SNP distributions. Panel C presents the empirical percentiles for ϵ_t and j_t compared to those ones of the fitted SNP and normal densities. The table only displays the SNP parameters that best fit the data, according to accuracy criteria and tests. For example, component ϵ_t for the spot log-price is explained in terms of d4, d6, and d8. The likelihood-ratio (LR) test for the joint significance of the models is also presented. P-values for this test, as well as for the all individual t-tests, are also provided.

Figure 6. Observations of the component ϵ_t when $h_t = 0$.



The figure shows the fit of the density functions for the stochastic component of ϵ_t, for both the spot price and its logarithm. The SNP distribution (dashed-dotted line) achieves a better description of the data density (shaded area) than the normal distribution (solid line).

Figure 7. Observations of the component j_t when $h_t = 1$.



The figure shows the fit of the density functions for the stochastic component of ϵ_t, for both the spot price and its logarithm. The SNP distribution (dashed-dotted line) achieves a better description of the data density (shaded area) than the normal distribution (solid line).

For the model of the spot log-price, we found that ϵ_t has a SNP distribution with parameters $(d_4, d_6, d_8) = (0{,}1792\ ,\ 0{,}0296\ , 0{,}00217)$, and j_t has a SNP distribution with parameter $(d_4) = (0{,}1425)$, which are statistically significant for a confidence level of 95%. The likelihood-ratio test shows that, for a confidence level of 99%, ϵ_t and j_t are better represented by the proposed SNP distributions than their normal counterpart. Again, from Figure 7 we can observe how the flexibility of the SNP specification allows a parsimonious description of both tails of the probability distribution of variable j_t.

Regarding the percentiles of the random components of price, we can observe that the SNP distribution also achieves a better performance than a normal distribution. This situation is relevant for estimation and simulation purposes because a better-defined cumulative probability distribution enables the generation of more coherent random numbers. Likewise, this condition on the percentiles shows how the measurement of risks in electricity markets should avoid the use of normality assumptions for the calculation of uncertainty indicators or expected loss.

4.4. Comparison of moments

In other to find more evidence on the performance of our model, this section simulates ten thousand possible trajectories of the spot price from 1995 to 2018, on the basis of the estimates in Table 3. For these trajectories we find that the calibrated models adequately capture the characteristics of seasonality, mean reversion, skewness, jumps, regime and switch, and, in general, they can describe the moments of the spot price of electricity. This situation can be observed in the price simulation figures in Appendix A and Table 4, which reports the first four sample moments and main sample percentiles compared to their estimates according to our model applied to the spot price and its logarithm.

In general terms, the model of the spot log-price accomplishes a better fit of the price percentiles, while the spot price model is better fitted to the mean and standard deviation.

Table 4. Model comparison by simulation.

Statistic		**Historical**	**Simulation**	
		Sample	**Model for spot price**	**Model for spot log-price**
Mean		107	105	122
Std. dev.		119	120	145
Skewness		4	2	4
Kurtosis		30	10	35
Percentile	5%	18	-52	14
	25%	50	28	38
	50%	76	96	78
	75%	133	162	154
	95%	212	285	367

This table presents a comparison of the electricity spot price moments in Colombia, obtained using three methods: (1) sampling, (2) simulating the spot price and (3) simulating the spot log-price. Simulations are based on the model proposed in Section 2.1. The model for the spot log-price exhibits best performance for the price percentiles, while the spot price model better fit the standard deviation.

The spot price model achieves a more accurate estimation of the two first moments. The model calculates a mean of 105 COP/kWh and a standard deviation of 120 COP/kWh, while the sample reports 107 COP/kWh and 119 COP/kWh, respectively. The spot price model exhibits a positive skewness, similar to the historical data, which is coherent with the description of the dynamics of the electricity price of some of the authors mentioned in the introduction. In the presence of extreme events, the model also describes the leptokurtosis of the series and can represent the presence of heavy tails. The greatest weakness of the spot price model is the generation of negative values for some of the simulated trajectories. In Table 4, the 5% percentile exhibits a negative value, which poses a challenge for the analyst, who should use the results for decision making. In practical applications, it may be convenient to incorporate a minimum value for the price in the simulation, systematically eliminate the trajectories with a negative value, or redefine the distributions for the stochastic components with some type of endogenous truncation algorithm. These findings are evidenced in Figure 8 in Appendix A.

The model of the spot log-price seems more suited for explaining skewness, kurtosis, and percentiles. In the case of skewness, it reports a value of 4, as in the sample. Regarding kurtosis, it presents a better performance than the spot price model, with a value of 35, compared to the value of the sample, 30. The nature of the proposed treatment on the series guarantees that all the simulated values are positive, which is consistent and does not make necessary to implement further algorithms to supervise the trajectories.

5. Conclusions

This work proposes a stochastic process with a component of mean reversion and regime switches to represent the dynamics of the spot price of electricity and its logarithmic form. For that purpose, three components were included: a deterministic one, another one of mean reversion, and a third one of regime switching. The short-term distortion of the mean reversion and regime switching components is represented by flexible semi-nonparametric (SNP) distributions with a probability density function defined by a finite Gram–Charlier expansion, in contrast to previous studies that assume the particular case of the Gaussian process. Therefore, the most important contribution of this paper is the modeling of the switching regime in terms not only of mean and standard deviation but also skewness, kurtosis, and higher-order moments.

Furthermore, we propose the incorporation of a binary regime switch component, where the probability of the transition is a function of the time passed in each regime. This representation enables the consideration of events that recurrently change the electricity spot price during one or several consecutive periods. Such is the case of El Niño phenomenon in the Colombian market, where its probability of occurrence is close to 50% after 104 months of its absence; and the probability ending is 50% if the phenomenon has taken place for 8 consecutive months. This switching regime model enables us to modify the spot price in the deterministic and random components, according to the behavior of the sample. Such generalization can be extrapolated in other markets (not only electricity), and its strength is complemented with the use of the SNP distribution, which captures the variations of asymmetry, kurtosis, and higher-order moments in case of regime switching.

For the Colombian electricity market, the modeling of spot prices and its logarithm shows that the SNP representation of the random components of the stochastic process significantly outperforms the normal distribution. However, when the spot price model is simulated, some trajectories present negative values, a situation that is not coherent with the type of variable being represented. This shortcoming does not seem to happen when we model log-prices. Simulations also exhibit more accurate performance for the first two moments when modeling spot price but the skewness, kurtosis and quantiles are outperformed by the use of logarithmic transformations. Therefore, depending on the objectives of the modeling and the series used the analyst should choose the best strategy. In general, we recommend taking the natural logarithm of the series and modeling random components with flexible SNP distributions.

As a matter of fact, the non-normality of random components of electricity prices cannot be overlooked, since it has serious implications for guaranteeing long-term supplies, sustainable investments and efficient market pricing. In this paper we propose a general and flexible model to overcome this problem. Planning, regulation, and control agencies, as well as agents who buy and sell electricity, can use these results to improve the way they measure risks, manage their portfolio, and define the expansion of the system. The activities related to electrical system planning should consider the effect of the occurrence of extreme events in terms of probability, impact, and length. If the regime switching (which, in Colombia, is produced by the occurrence of El Niño phenomenon in its different categories) further affects the levels of expectation and uncertainty, the agents involved in short- and long-term price formation should incorporate the assumed risks in the price that users pay. The expansion of systems should not only ensure the supply, long-term prices in accordance with income levels, or the needs of users but also the (systemic and idiosyncratic) risk levels that involve hidden or revealed costs that jeopardize the conditions of return of the invested capital or the objectives of development and social welfare. The model proposed in this study addresses the quantification of the possible impact on the cash flows of all the participants in the market in the short, medium, and long term.

Declaration of Competing Interest

None

References

Bahar, H., & Sauvage, J. (2013). Croos-Border Trade in Electricity and the Development of Renewables-Based Electric Power: Lessons from Europe. *OECD Trade and Environment Working Papers 2013/02*.

Barton, D.E. & Dennis, K.E.R. (1952). The Conditions under Which Gram-Charlier and Edgeworth Curves are Positive Definite and Unimodal. *Biometrika* 39, 425-427.

Baum, C., Zerilli, P., & Chen, L. (2019). Stochastic volatility, jumps and leverage in energy and stock markets: evidence from high frequency data. *Boston College, Economics Working-Papers* 67.

Borovkova S., Permana F.J. (2006). Modelling electricity prices by the potential jump-diffusion. In: Shiryaev A.N., Grossinho M.R., Oliveira P.E., Esquível M.L. (eds) Stochastic Finance. Springer, Boston, MA.

Brunner, A. D. (Jan de 1992). Asymmetries in real GNP: A seminonparametric approach. (L. Taylor & Francis, Ed.) *Journal of Business & Economic Statics*, 65-72.

Burke, M. J., & Stephens, J. C. (2018). Political power and renewable energy futures: A critical review, *Energy Research & Social Science*, 35, 78-93.

Cartea, Á., & Figueroa, M. (2005). A Mean Reverting Jump Diffusion Model with Seasonality. *Applied Mathematical Finance*, 313-335.

Cortés, L., Mora-Valencia, A., & Perote, J. (2016). The productivity of top researchers: a semi-noparametric approach. *Scientometrics*, *109*(2), 891-915.

Cortés, L. M., Mora-Valencia, A., & Perote, J. (2017). Measuring firm size distribution with semi-nonparametric densities. *Physica A*, 35-47.

Del Brio, E.B., Mora-Valencia, A., & Perote, J. 2014. VaR Performance during the Subprime and Sovereign Debt Crises: An Application to Emerging Markets. *Emerging Markets Review* 20, 23-41.

Dupuis, D. J. (2017). Electricity price dependence in New York state zones: a robust detrended correlation approach. *The Annals of Applied Statistics*, *11*(1), 248-273.

Escribano, A., Peña, I., & Villaplana, P. (2011). Modelling Electricity Prices: International Evidence. *Oxford Bulletin OF Economics AND Statistics*, 622-650.

Falbo, P., Fattore, M., & Stefani, S. (2010). A new index for electricity spot markets . *Energy Policy, 38*, 2739–2750.

Fanone, E., Gamba, A., & Prokopczuk, M. (2013). The case of negative day-ahead electricity prices. *Energy Economics*, 35, 22-34.

Gallant, A., Rossi, P., & Tauchen, G. (1992). Stock prices and volumens. *The Society for Financial Studies, 5*, 199-242.

Geman, H., & Roncoroni, A. (2003). A class of marked point processes for modelling electricity prices. *ESSEC Working Papers DR 03004, ESSEC Research Center, ESSEC Business School*, 2-70.

Hinderks, & Wagner. (2019). Factor models in the German electricity market: Stylized facts, seasonality, and calibration. *Energy Economics*, 1-13.

Huisman, R., & Huurman, C. (2003). Fat tails in power prices. *ERIM Report Series Research in Management*, 1-13.

Jondeau, E., & Rockinger, M. (2001). Gram-Charlier densities. *Journal of Economic Dynamics & Control, 25*, 1457-1483.

Knaut, A. & Paschmann, M. (2019). Price volatility in commodity markets with restricted participation. *Energy Economics*, 81, 37-51

Lucia, J., & Schwartz, E. (2002). Electricity prices and power derivatives: Evidence from the Nordic Power exchange. *Review of Derivatives Research*, 5-50.

Mauleon, I., & Perote, J. (2000). Testing Densities With Financial Data: An Empirical Comparison of the Edgeworth-Sargan density to the Student's t. *The European Journal of Finance*, *6*, 225-239.

Mayer, K., & Trück, S. (2018). Electricity markets around the world. *Journal of Commodity Markets, 9*, 77-100.

Mount, T., Ning, Y., & Cai, X. (2006). Predicting price spikes in electricity markets using a regime-switching model with time-varying parameters. *Energy Economics*, 62-80.

Ñíguez, T.-M., & Perote, J. (2012). Forecasting heavy-tailed densities with positive edgeworth and Gram-Charlier expansions. *Oxford Bulletin of Economics ans Statistics, 74*(4), 116-137.

Peerbocus, N. (2007): Assessment of Reforms in the Electricity Supply Industry: A Review of Some Recent Empirical Studies. *The Electricity Journal*, 20(2), 44-56.

Pilipovic, D. (2007). *Energy risk, valuing and managing energy derivatives.* New York: McGraw-Hill.

Poveda, G., & Mesa, O. (1997). Feedbacks between hydrological processes in tropical South America and large-scale ocean–atmospheric phenomen. *Journal of Climate, 10*, 2690-2702.

Sioshansi, F. P. (2006): Electricity Market Reform: What has the Experience taught us thus far? *Utilities Policy*, 14(2), 63-75.

Trespalacios, A., Cortés, L., & Perote, J. (2019). Uncertainty in electricity markets from a semi-nonparametric approach. *Energy Policy*, forthcoming. https://doi.org/10.1016/j.enpol.2019.111091

Trespalacios, A., Pantoja, J., & Fernández, Ó. (2017). *Análisis de mercados de electricidad.* Medellín, Colombia: Fondo Editorial Universidad EAFTI, Colección académica.

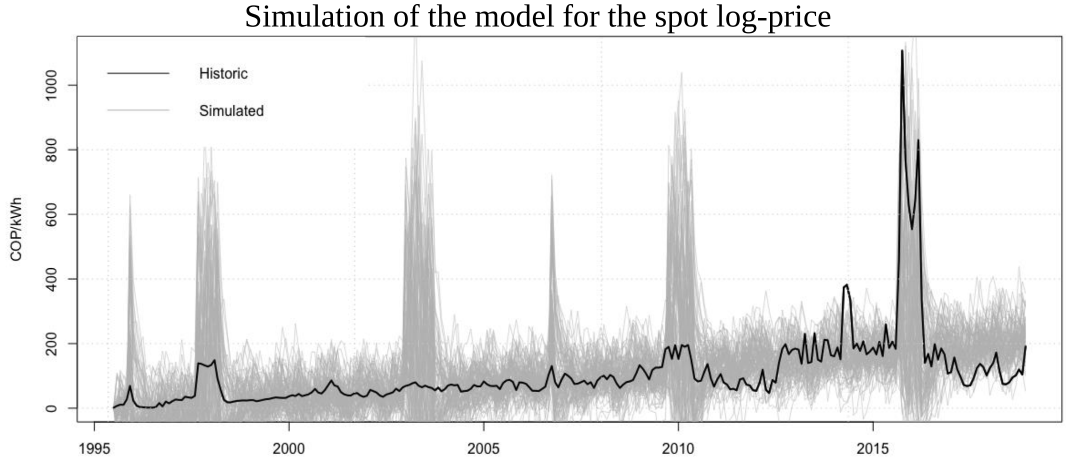
Trespalacios, A., Rendón, J. F., & Pantoja, J. (2012). Estrategia de cobertura a través de contratos a plazo en mercados eléctricos. *Academia, Revista Latinoamericana de Administración*, 148-157.
Tsay, R. S. (2010). *Analysis of Financial Time Series.* New Jersey: John Wiley & Son, Inc.
Waylen, P., & Poveda, G. (2002). El Nino–Southern Oscillation and aspects of western South American hydro-climatology. *Hydrological Processes*, *16*, 1247-1260.
Weron, R. (2008). Market price of risk implied by Asian-style electricity options and futures. *Energy Economics, 30*(3), 1098-1115
Weron, R. (2014). Electricity price forecasting: A review of the state-of-the-art with a look into the future . *International Journal of Forecasting*, 1030-1081.

Appendix A. *Historical simulation of the electricity spot price in Colombia*

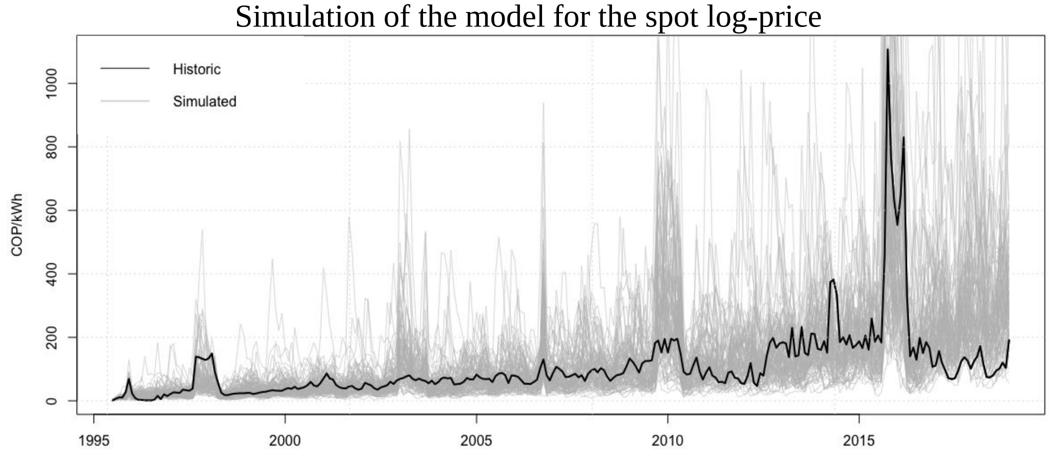
This appendix shows the simulation of the spot price and its accorfing to the model in Section 2.1 calibrated for the electricity price in Colombia from June 1995 to December 2018.

Figure 8. Historic and simulated data for spot price and its logarithm.

Simulation of the model for the spot log-price



Simulation of the model for the spot log-price



The figure above shows the monthly simulation of 100 stochastic trajectories of the spot price of electricity in Colombia from June 1995 to December 2018. Simulations are done for spot price and spot log-price on de basis of the model in Section 2.1. The black line shows the historical series of the price. It is noteworthy that, although the parameters used for the simulations have statistical in-sample significance to represent the electricity price, the simulation of the price can present negative values, a situation that does not occur when the log-price is modeled. However, both models can represent the structures of seasonality, mean reversion, skewness, kurtosis, and jumps.