



Documentos de trabajo

Economía y Finanzas

N° 19-05

2019

Uncertainty in Electricity Markets from a semi-nonparametric Approach

Alfredo Trespalacios, Lina M. Cortés,
Javier Perote

Uncertainty in Electricity Markets from a semi-nonparametric Approach[†]

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

June 4, 2019

Abstract

The spot price of electricity is highly skewed and heavy-tailed, as a result of the interaction of different variables that affect that market. Such characteristics impact the design of power plants with different technologies, fuel prices, and energy demand. This paper introduces the semi-nonparametric (SNP) approach to describe the uncertainty of different variables in an electricity market, reducing the limitations that normality and parametric density functions impose. The selection of probability density functions is achieved in terms of a finite Gram-Charlier expansion fitted by the maximum likelihood criterion. The study case is the Colombian electricity market, where the SNP distribution outperforms the normal distribution for spot price, national energy demand, the climate index ONI, and the series of hydrologic inflows of the system and some rivers. The results show that risk analysis in electricity markets requires the measurement of skewness, kurtosis, and high-order moments. The flexible methodology in our study has directly applications for implementing policies on electricity markets that improve the sustainability indicators of different systems. The particular characteristics of the series under analysis should be considered as a starting point for risk analysis and portfolio choice.

Keywords: Electricity market, SNP modeling, Risk management

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. University institution 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: lcorteds@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

The decentralization of electricity markets worldwide enables the participation of national and foreign private investors in a sector traditionally characterized by inefficiencies caused by the administration of state governments (Sioshansi, 2013). Such restructuring gave birth to markets that promote the interaction of agents with particular purposes, acting at peak supply and demand. Such supply is produced by electricity generators that transform primary forms of energy (solar, wind, and hydraulic power and natural gas) into electricity, while demand is represented by retailers (including industries, businesses, and households), who use electrical energy for production or well-being. Efficient resource management in this type of markets should consider the uncertainty of the volumes and the prices at which electrical energy is traded. Additionally, it requires portfolio managers to use practical risk management tools for asset planning, investment, commercialization, or management (Avci, Ketter, & van Heck, 2018).

According to Pérez Odeh, Watts, and Negrete-Pincetic (2018), the risks associated with investments and transactions in electricity markets affect the costs and profitability of projects. Such risks can be classified into four groups: technical, market, financial, and systemic. In the technical aspects are investment costs, maintenance, system failures during the operation, operation time, and construction time. Market risks include fuel prices, the opportunity cost intangible primary assets (such as the subjective price of water collected in a reservoir), energy price, renewable energy certificates, and energy demand. In turn, financial risks refer to the cost of capital, credit risk, and counterparty risk — finally, systemic risks due to transmission lines, macroclimatic conditions, environmental regulations, and social opposition. In line with the ideas above and in order to contribute to the improvement of risk-measuring tools in electricity markets, this work analyzes the variables that represent sources of technical, market, and systemic risks in Colombia: Spot Price, Energy Demand, Hydrologic Inflows, and conditions measured by the Oceanic Niño Index (ONI).

In different risk management applications, it is common to assume that changes of the variables under study can be fitted by a normal distribution. Although such an assumption has simplified uncertainty modeling in different fields of application, it becomes a limitation when market dynamics present skewness or the recurrence of extreme events result in heavy-tailed distributions. As a result of the interaction of the variables involved, the price of electrical energy exhibits stationarity characteristics, mean reversion, and heavy tails that have been empirically found by Benth, Kallsen, and Meyer-Brandis (2007); Geman and Roncoroni (2003); and Uribe and Trespalacios (2014), among other authors. Such characteristics provide reasons to suspect that the assumption of normality may be insufficient to represent the uncertainty of electricity price, demand, or supply. This work proposes the use of semi-nonparametric (SNP) techniques as a flexible alternative to adequately capture uncertainty.

The SNP approach for statistical modeling is based on the early work of Edgeworth (1907), but they were Sargan (1975) and Gallant and Nychka (1987) who introduced these analyses into econometrics. These authors derived Edgeworth and Gram-Charlier expansions to

approximate frequency functions and the asymptotic properties of related estimators.⁵ More recently, different authors have studied the properties of truncated SNP expansions to represent a valid probability density function (*pdf*), see e.g. Jondeau and Rockinger (2001), Ullah (2004) or León, Mencía and Sentana (2009).

The implementation of SNP methodologies for explaining economic series has been gradually adopted in different studies. Initially they were employed for option pricing (Jarrow and Rudd, 1982), but Brunner (1992) argued that the selection of such technique could prevent specification errors that underlie parametric modeling and allow the study of asymmetries in the real Gross National Product (GNP). In the same line, Gallant, Rossi, and Tauchen (1992) used SNP modeling to describe the comovements of prices and volumes of the stock market of the United States (US) in the period from 1928 to 1987. Mauleon and Perote (2000) proposed the use mixtures of SNP distribution to model the stock market of the US and the United Kingdom and Níguez and Perote (2012) implemented positive SNP transformations to evaluate the stock performance in US. Other works that adopt SNP approaches to the modeling of heavy-tailed model series are those by Cortés, Mora-Valencia and Perote (2016), who measure the productivity of researchers worldwide, and Cortés, Mora-Valencia, and Perote (2017), who estimate the size distribution of US firms in the period from 2004 to 2015.

This paper introduces the SNP approach to the modeling of energy markets, showing that the Gram–Charlier expansion outperforms the normal distribution when fitting different series of the Colombian electricity markets. Particularly, maximum likelihood (ML) criterion was used to fit the data from Electricity Demand, Spot Price, aggregated Hydrologic Inflows of the system, ONI, and the series of the Hydrologic Inflows of several rivers: Nare, Salvajina, Guavio, and San Carlos. This study opens a path for the formulation of stochastic models for the price, demand or determinants of supply in electricity markets, as well as the subsequent structuring of investment portfolios, without the limitations imposed by the assumption of normality.

The rest of the paper is structured as follows. Section 2 presents the variables under study and their significance in the treatment of uncertainty for decision making. Section 3 describes the model and the methodology, specifically the SNP distribution and the ML method, which is implemented in Section 4. Section 5 reports and discusses the results, and Section 6 summarizes the conclusions.

2. Electricity markets

This section explains some of the elements of the product, supply, and demand that impact the price and transactions in electricity markets. Their description provides insights on the variables studied in this work: Spot price, Demand, Hydrologic Inflows, and Climate Index.

⁵ There are many other SNP type of expansions (e.g. that of Cornish and Fisher, 1938, approximation) that are not considered in this paper.

The supply chain of electricity is composed of several activities: generation (transforming energy from primary sources into electricity), transmission (transporting the energy from production centers to consumption centers, hundreds of kilometers away), distribution (taking the energy to each customer), and retailing (responsible for billing, collection, and compensation of the supply chain). The final price of electricity is defined by the adequate compensation of operating costs, investment, and opportunity to each one of the agents that interact to provide said service. That requires agents to pay close attention to the availability of energy in the short, medium, and long term due to three conditions: (i) electricity should be produced at the same moments it is demanded; (ii) the availability of primary energy sources limits the production capacity in the short term, depending on external agents such as producers or transporters of fossil fuels or natural agents, e.g., hydrologic inflows; and (iii) investments in electrical infrastructure are capital intensive and present long capital recovery periods.

The limitations on electricity production and storage are key to price formation (Lucia & Schwartz, 2002). A different situation occurs in foreign exchange or the stock market and, in general, in financial markets, where the economic cycle is key and there are no physical limitations to interaction between agents (Pilipovic, 2007). Geman and Roncoroni (2003) claim that the supply curve in electricity markets presents an exponential growth caused by the aggregation of generation plants in charge of transforming different types of primary energy into electricity. Each plant represents production costs that may vary within a wide range of values and, in conditions of scarcity, may present important jumps.

Additionally, as described by Huisman and Huurman (2003), the marginal cost of meeting the electricity demand may be considered inelastic in the short term.⁶ Likewise, Atalla, Bigerna, and Bollino (2018) studied 117 countries that concentrate 95% of the world population. They found that, in the period from 1978 to 2012, electricity demand presented a price elasticity between -0.1 and -0.2 in most countries and, furthermore, it was determined by the behavior of the climate and the capital market. In the case of emerging economies, they found that their price elasticity is lower than that of developed nations, where the effects of energy efficiency are more pronounced. Similar results were found by Gutierrez-Pedrero, Tarancon, del Rio, and Alcantara (2018), who analyzed the relationship between electricity demand and Gross Value Added in the Member States of the European Union, and Campo and Sarmiento (2013), who calculated the elasticity of energy consumption with respect to the Gross Domestic Product of 10 countries in Latin America (Argentina, Bolivia, Brazil, Chile, Colombia, Ecuador, Paraguay, Perú, Uruguay, and Venezuela). Those results are also supported by Pasten, Saens, and Contreras Marín (2015), who analyzed the cointegration between energy consumption, capital, and workforce in 16 Latin American countries with data between 1971 and 2001, and Pinzón (2018), who found, applying Granger causality, that energy consumption causes economic growth in Ecuador.

⁶ This is because the movements of the electricity demand in the short term are not as fast and intense as the short-term movements of the supply curve of the market. This situation occurs as a result of conditions of sudden unavailability of generation plants, significant reduction of hydrologic inflows, or failures in the fuel supply or transmission networks, among others.

On the other hand, Benth, Kallsen, and Meyer-Brandis (2007) explain that the dynamics of electricity prices exhibit three important characteristics: seasonality over several time horizons (hourly, weekly, or monthly), mean reversion, jumps, and heavy tails. Such characteristics are also described in other studies (e.g., Lucia & Schwartz, 2002; Huisman & Huurman, 2003; Geman & Roncoroni, 2003; Pilipovic, 2007; Falbo, Fattore, & Stefani, 2010; Janczura, Trück, Weron, & Wolff, 2013; Trespalacios, Rendón, & Pantoja, 2012; Weron, 2014; Uribe & Trespalacios, 2014; Gonzalez, Moriarty, & Palczewski, 2017; Maradey, Pantoja, & Trespalacios, 2017; Dupuis, 2017; Mayer & Trück, 2018). Seasonality responds to economic and climate cycles. Because large-scale electricity storage is technically limited, the highest levels of demand occur when the national production is higher and the requirements of air conditioning, heating, or water pumping for agricultural processes increase. Mean reversion refers to a memory condition of the price regarding its long-term mean, that is, a significant and sudden increase of the series tends to be reverted in the long term. Heavy tails are, therefore, a consequence of the occurrence of sudden price jumps that lead the kurtosis of the series to present high values (over 3). The characteristics described above are key to suspect that the assumption of normality is not adequate to represent the uncertainty that underlies the interaction of the variables involved in electricity markets.

Electricity may be traded by agents using several mechanisms that include, without limitation, spot and long-term contract markets or forward contracts. The spot market is the place where energy can be traded instantaneously, and its price is formed by the immediate intersection of supply and demand (Trespalacios, Pantoja, & Fernández, 2017). Therefore, an agent that decides to perform all his or her transactions at spot price should assume possible short-term variations in the supply as well as the demand, which may generate volatility levels even 1000% higher than those of the price of natural gas (Pilipovic, 2007). Forward contracts are agreements to supply energy in the future with a price previously established. Such contracts seek to reduce agents' uncertainty levels about the future cash flow. In this case, the price of the contracts is formed based on spot price expectations and risk levels measured using a Forward Risk Premium (FRP). According to Pantoja (2011), the FRP in the Colombian electricity market is explained by variations in the climate conditions determined by the Oceanic Niño Index (ONI) because of its impact on hydrologic inflows to generation plants, in a country where hydroelectric power generation holds a share exceeding 50% of the entire system.

The ONI is a series of measurements used to describe El Niño Southern Oscillation (ENSO), which is a periodic irregular variation of the winds and temperature of the sea in the Pacific Ocean. 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 hydrologic conditions in Colombia and the northern area of South America, changing rainfall levels and the corresponding hydrologic inflow to hydraulic generation plants. Despite the chaotic nonlinear dynamics of ENSO, a probabilistic forecast can be produced based on an a priori classification of atmospheric conditions (Waylen & Poveda, 2002).

3. Model and Methodology

3.1. Normal distribution

If a variable y is normally distributed, then its *pdf* is described by Equation (1). Such *pdf* represents a symmetric density whose mean is captured by μ and full shape depends on one single parameter σ , e.g. σ^2 is the variance and $3\sigma^4$ the kurtosis. The standard normal distribution assigns values of zero and one to the location and dispersion parameters, respectively, according to Equation (2).

$$\phi(y|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(y-\mu)^2}{2\sigma^2}\right], \quad (1)$$

$$\phi(y) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{y^2}{2}\right]. \quad (2)$$

3.2. SNP distribution

A variable y is said to be SNP distributed if its *pdf* can be represented as in Equation (3), i.e. in terms of an (infinite) expansion of Hermite Polynomials (HPs) weighted by the standard normal function $\phi(y)$ – see Equation (2) – and a set of parameters defined as $\delta_s = \frac{1}{s!} \int_{-\infty}^{\infty} H_s(y) f(y) dy$. The HP of order s is obtained from the s -th order derivative of the standard normal distribution, as shown in Equation (4). Therefore the SNP distribution may be interpreted as an asymptotic expansion of the standard normal density that can be used to approximate any continuous and differentiable frequency function – see Ullah (2004) for further details on this expansion, which is usually referred as the Gram-Charlier Type A series.

$$f(y) = [\sum_{s=0}^{\infty} \delta_s H_s(y)] \phi(y), \quad (3)$$

$$H_s(y) = \frac{(-1)^s}{\phi(y)} \frac{d^s \phi(y)}{dy^s}, \quad (4)$$

Table 1 displays the coefficients of the first fourteen HPs. The zero-order Hermite Polynomial is assumed to be $H_0(y) = 1$, and the rest can be recursively computed from the relation in Equation (4), e.g. $H_1(y) = y$, $H_2(y) = y^2 - 1$ and $H_3(y) = y^3 - 3y$.

Table 1. Coefficients of the first fourteen Hermite Polynomials.

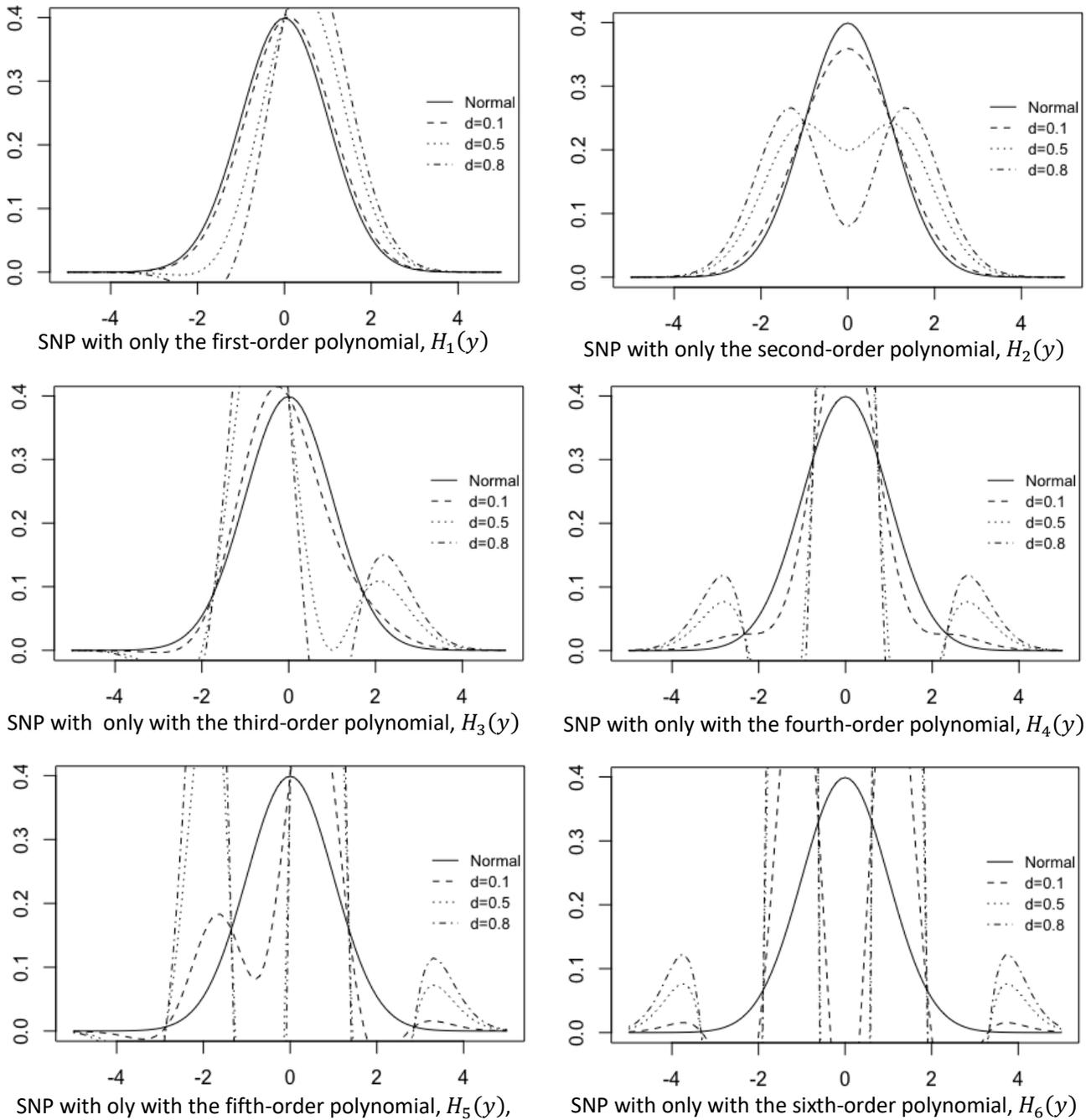
		Coefficients of Hermite Polynomials $H_s(y)$, from $s=0$ to $s=14$													
s	y^0	y^1	y^2	y^3	y^4	y^5	y^6	y^7	y^8	y^9	y^{10}	y^{11}	y^{12}	y^{13}	y^{14}
0	1														
1	0	1													
2	-1	0	1												
3	0	-3	0	1											
4	3	0	-6	0	1										
5	0	15	0	-10	0	1									
6	-15	0	45	0	-15	0	1								
7	0	-105	0	105	0	-21	0	1							
8	105	0	-420	0	210	0	-28	0	1						
9	0	945	0	-1,260	0	378	0	-36	0	1					
10	-945	0	4,725	0	-3,150	0	630	0	-45	0	1				
11	0	-10,395	0	17,325	0	-6,930	0	990	0	-55	0	1			
12	10,395	0	62,370	0	51,975	0	13,860	0	1,485	0	66	0	1		
13	0	135,135	0	270,270	0	135,135	0	25,740	0	2,145	0	78	0	1	
14	-135,135	0	945,945	0	945,945	0	315,315	0	45,045	0	3,003	0	91	0	1

The table shows the coefficients of the Hermite Polynomials. y^0 refers to the independent term of the polynomial, while y^4 denotes the coefficient of the independent variable of the polynomial to the fourth power. For example when $s=4$: $H_4(y) = y^4 - 6y^2 + 3$.

The marginal effect of the introduction of the first six HPs on the SNP expansion, compared to the constrained zero-term expansion, i.e. the standard normal, is illustrated in Figure 1. The plots show the the sensitivity of the ‘density’ to different values of the corresponding δ_s parameter. For example, the picture in the second row and first column in Figure 1 represents Equation (3) with $\delta_s = 0 \forall s \neq 3$; that is, the effect of incorporating only the third-order HP for different values of δ_3 (0.1, 0.5, and 0.8). When the normal distribution is compared to the SNP distribution, it can be seen that polynomials of even order s feature bimodality and high kurtosis, in both the right and left tails of the *pdf*. When the polynomial order s is odd the ‘density’ becomes skewed and thick-tailed in the right tail (provided that the parameter is positive). It can be also observed that the higher the order of the polynomial, the stronger the effect, for a given δ_s , e.g. the sixth-order HP has a higher impact on kurtosis than its second-order counterpart. Identifying such marginal effects of parameters δ_s enable the analyst to identify the type of components that should be involved in an SNP distribution when a distribution is fitted to a particular dataset.

Another issue that deserves attention is the fact that the introduction of different polynomials might yield the ‘density’ to negative values. This is a well-known shortcoming of Gram-Charlier expansions (see Barton & Dennis, 1952), which have been tackled by imposing positivity transformations (Gallant & Nychka, 1987), identifying the domains in which positivity holds (Jondeau & Rockinger, 2001) or implementing ML estimation procedures whose optimal solutions prevent the density from negative values (Cortés, Mora-Valencia & Perote, 2016).

Figure 1. Sensitivity of SNP expansion to different Hermite Polynomials



The figure compares the shape of the standard normal distribution, $N(0,1)$, with different expansions that incorporate a single HP polynomial in the SNP expansion in every picture. The marginal effect of each polynomial is plotted considering three different values of the corresponding δ_s parameter: 0.1, 0.5, and 0.8. The picture illustrates how the SNP expansion accounts for high kurtosis and modality and also for skewness when polynomials of odd s order are considered.

For empirical purposes the SNP expansion has to be truncated in a finite order n , which give rise to a distribution $g(y; \mathbf{d})$ that is parameterized in terms of vector $\mathbf{d} = (d_1, d_2, \dots, d_n) \in \mathbb{R}^n$. Equation (5) shows the finite SNP density with $\delta_0 = 1$ (see Proof 1 in the Appendix). It is noteworthy that truncated density integrates one as a consequence of the orthogonality of HPs ($\int_{-\infty}^{\infty} H_i(y)H_j(y)\phi(y)dy = 0$ for $i \neq j$).

$$g(y; \mathbf{d}) = [1 + \sum_{s=1}^n d_s H_s(y)]\phi(y), \quad (5)$$

The cumulative distribution function (*cdf*) a n -term SNP expansion may be obtained as follows (see Proof 2 in the Appendix),

$$G(y|\mathbf{d}) = \int_{-\infty}^x \phi(y)dy - \phi(y) \sum_{s=1}^n d_s H_{s-1}(y), \quad (6)$$

and the central moment of order s (μ_s) may be obtained from the first s even (odd) moments when s is even (odd), see the first eighth moments below (see Del Brio & Perote, 2012).

$$\begin{aligned} \mu_1 &= d_1, \\ \mu_2 &= 2d_2 + 1, \\ \mu_3 &= 6d_3 + 3d_1, \\ \mu_4 &= 24d_4 + 12d_2 + 3, \\ \mu_5 &= 120d_5 + 60d_3 + 15d_1, \\ \mu_6 &= 720d_6 + 360d_4 + 90d_2 + 15, \\ \mu_7 &= 5040d_7 + 2520d_5 + 630d_3 + 105d_1, \\ \mu_8 &= 40320d_8 + 20160d_6 + 5040d_4 + 840d_2 + 105. \end{aligned} \quad (7)$$

3.3. ML estimation

The relation between the parameters and the moments of the SNP density allows a direct implementation of the method of moments for estimation. Nevertheless this technique, although consistent, is not efficient and does not solve the positivity problem of truncated Gram-Charlier series. For this reasons we choose ML estimation for accurately fitting the density. The likelihood function is defined on the set of the parameters θ of a given *pdf* $h(\cdot)$, given a sample of realizations $\{x_i\}$ for $i = 1, \dots, N$. This function, $L(\theta)$, can be defined in terms of the product of the conditional pdfs as in equation (8).

$$L(\theta) = \prod_{i=1}^N h(x_i|\theta), \quad (8)$$

The ML estimate is the value $\hat{\theta}$ that maximizes the likelihood of the sample to be drawn from the chosen *pdf*. Usually, the log-likelihood function is optimized, which for the SNP distribution in Equation (5) results in the following function:

$$\begin{aligned} l_g &= \log[g(y_i|\mathbf{d})] = \sum_{i=1}^N \log[f(x_i|\theta)] = \\ &= -\frac{y_i^2}{2} - \log(\sqrt{2\pi}) + \log[1 + \sum_{s=1}^n d_s H_s(y_i)]. \end{aligned} \quad (9)$$

As the normal distribution is nested on the SNP specification, the performance of both densities can be naturally tested from the likelihood ratio (LR) test, which can be expressed in terms of the difference between the log-likelihood of the normal (l_ϕ) and SNP (l_g) and is distributed according to a χ_n^2 , n being the number of the parameters of the expansion – see Equation (10). Furthermore, the most accurate number of parameters in the SNP expansion may be also recursively chosen through LR tests.

$$LR = -2(l_\phi - l_g) \sim \chi_n^2, \quad (10)$$

4. Data description

The data used in this work are publicly available from two sources: the website of the Operator of the Colombian Electricity Market⁷ and the US National Oceanic and Atmospheric Administration (NOAA).⁸ All the variables are observed at monthly frequency from January 2000 until April 2018, except the ONI series, which contains information since 1950.

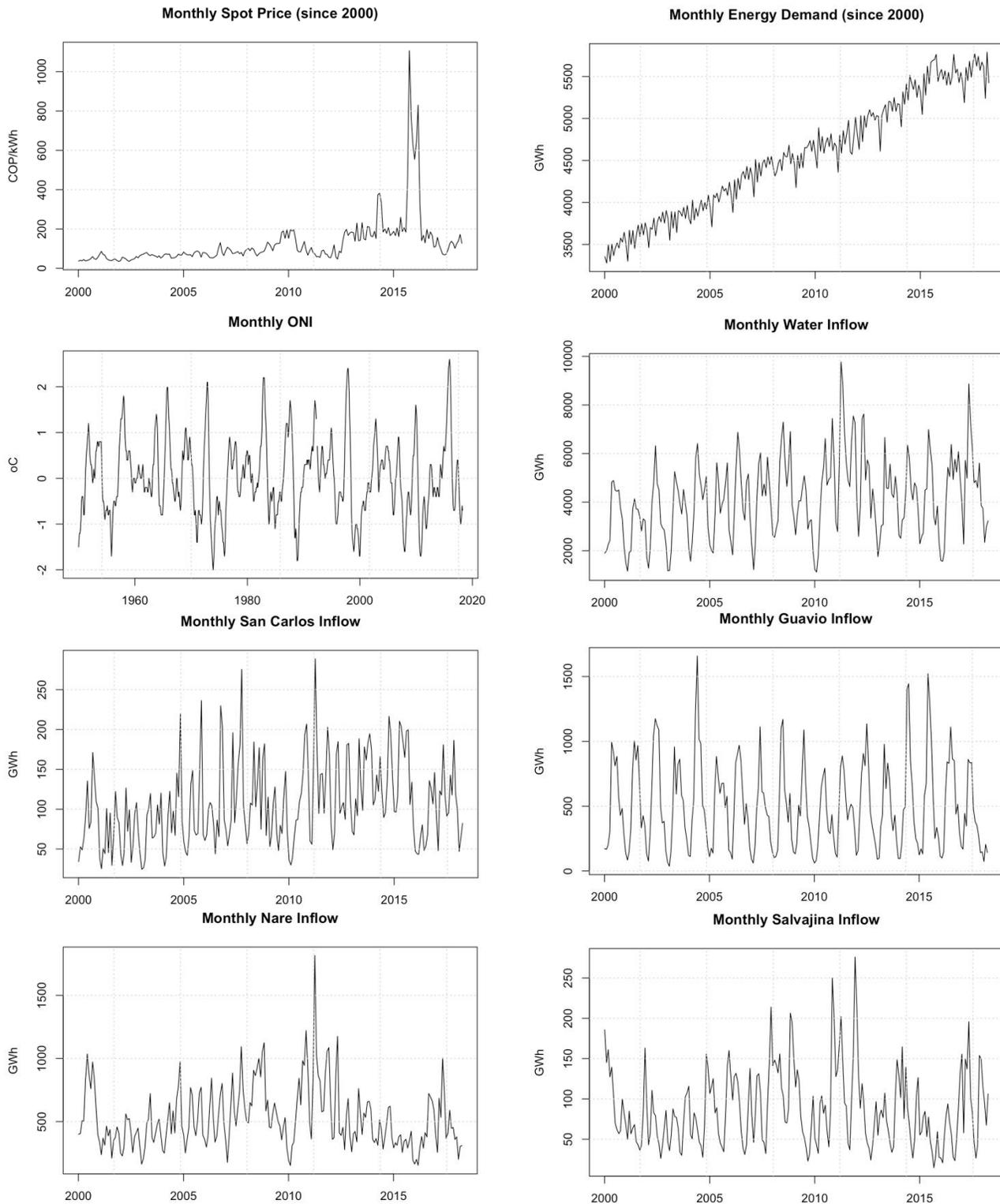
The Spot Price series corresponds to the average price of monthly energy traded in the energy market, measured in COP (Colombian pesos) per kWh (kilo Watt hour) (COP/kWh). The Demand series is the sum of all the energy consumed by the National Interconnected System (NIS) every month, measured in GWh⁹. The ONI series is the index reported by the NOAA every month, measured in °C. The Hydrologic Inflow series is the aggregate of all the amount of energy that the rivers in the system contribute to hydroelectric generation plants in a month, measured in GWh. The series of San Carlos, Guavio, Nare, and Salvajina Inflows measure the hydrologic inflows of each one of those rivers, which are necessary for generation (for the plants in San Carlos, Guavio, Guatapé, and Salvajina), measured in GWh.

⁷ www.xm.com.co

⁸ www.noaa.gov

⁹ 1 GWh equals 10⁶ kWh.

Figure 2. Time series for different energy variables.



The figures above are the time series of the variables analyzed in this work. The trends of Spot Price and Demand are noticeable.

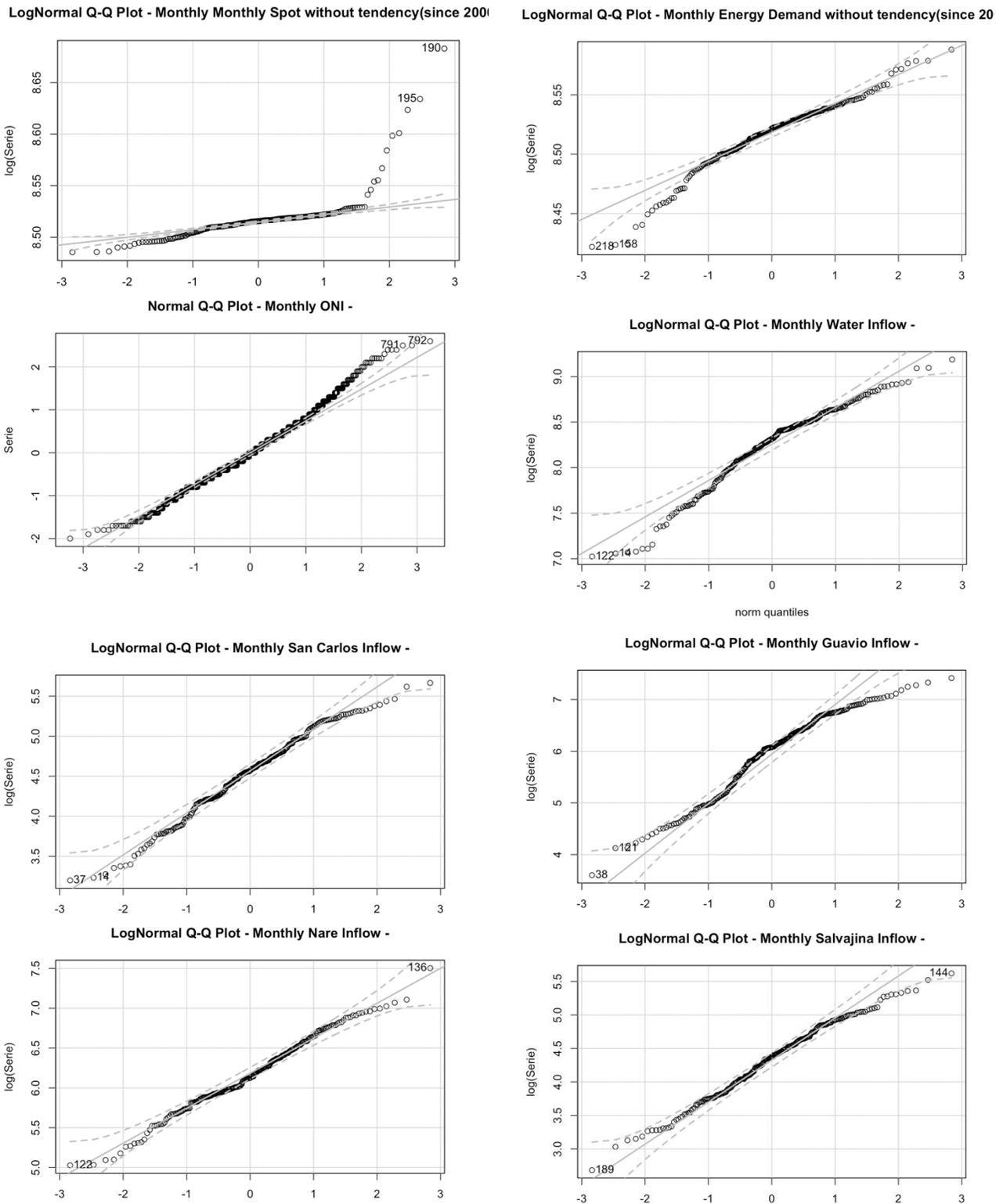
Figure 2 presents the distribution curves of the time series of each one of those variables. Remarkably, electricity Spot Price and Demand are the only series that present an evident trend, and Spot Price exhibits jumps, as described in previous literature about this and other markets. Given those circumstances, both Spot Price and Demand are analyzed in this work, eliminating the linear trend of the series and performing their natural log transformations. When to apply one or the other transformation to a particular series is explained below.

Figure 3 shows Q-Q plots comparing the percentiles of the series to the normal distribution. The right tail of the detrended Spot Price logarithm is far from being represented by a normal distribution. This also occurs with the detrended left tail of the series of the natural logarithm of demand, which cannot fit the normal distribution. Among all the series under analysis with their corresponding transformations, based on the Q-Q plot, only the logarithm of the Nare Inflow series can be modeled by a normal distribution (data remain within the bounds of the 95% confidence interval). No other series in Figure 3 maintains all its observations within the confidence interval, which provides evidence in favor of non-normality or log-normality to model the uncertainty of the electricity market.

This study considers different transformations of the time series. The type of transformation to apply to each series is selected following the nature of the data. Stationary series are modeled in two ways: the series and the natural logarithm of the series (ONI, NIS Inflow, San Carlos Inflow, Guavio Inflow, Nare Inflow, and Salvajina Inflow). In the case of nonstationary series (Spot Price and Demand), the time series is directly modeled after being detrended, both applying and not applying the natural logarithm. All the series under study reflect stationarity and clustering. Regarding Spot Price, Group 1 contains the months from January to August, and Group 2, from September to December. With respect to Demand, Group 1 contains January, February, April, and June; Group 2, the rest of the months. Group 1 of the ONI includes the months from January to May; its Group 2, June to December. Group 1 of NIS Inflow contains the months from January to March; its Group 2, the rest of the months.

The descriptive statistics of the series are found in Table 2. The series presents a noticeable leptokurtosis and positive skew. Although energy demand is the only series in this work that presents a kurtosis below 3, when detrended, such sample moment takes a value of 4.07. Detrended Spot Price is the series with the highest kurtosis (30.7), followed by Nare Inflow (5.9), and Detrended Demand in level. The series in a level whose kurtosis is closest to 3 is that of the ONI since 1950, followed by NIS Inflow.

Figure 3. Q-Q norm of the logarithm of the series and ONI.



The Q-Q plots of the logarithms of the series are compared with the normal distribution in a confidence interval of 95%. In the case of the ONI, the series is considered without calculating its natural logarithm. Spot Price and Energy Demand are detrended. The series comprises January 2000 to April 2018 period.

In practical applications, professionals commonly use the average as an indicator of expectations, in order to produce, for example, financial forecasts or analyze project feasibility. Table 2 reveals that the 50th percentile systematically differs from the mean of the samples. Professionals should consider this situation to interpret results and adequately guide decision makers.

Previous studies have emphasized asymmetrical conditions and heavy tails in electricity markets. This paper supports the fact that risk measurement in electricity markets should pay attention to the direction of the movements of the variables. That is to say; its objective is to improve the commonly-used measurement techniques applied to financial markets, which assume that positive movements are as probable and impactful as their negative counterparts. This idea is supported by the distances of extreme negative movements and extreme positive movements. For example, the distance between percentiles 50 and 1 of the Detrended Spot Price is 141 COP/kWh, while the distance between percentiles 99 and 50 is 546 COP/kWh. That shows that the extreme positive movement of the Detrended Spot Price is 3.87 times the extreme negative movement. Salvajina Inflow presents a similar condition, where the extreme positive movement (the difference between percentiles 50 and 1) is 2.41 times the extreme negative movement (the difference between percentiles 99 and 50).

Regarding Nare Inflow, the extreme positive movement is 2.41 times the extreme negative movement. Concerning Guavio Inflow, the extreme positive movement is 2.68 times its negative counterpart. In the case of Detrended Energy Demand, the extreme negative movements are higher: the extreme positive movement is 0.65 times the extreme negative movement. Although calculating the logarithm of the series smoothens their extreme movements, as shown in Table 2, after this transformation, the kurtosis is still above 3. As can be seen, the transformation achieves a kurtosis below three only for Guavio Inflow. This condition suggests that the log-normal distribution cannot capture the uncertainty either.

Table 2. Descriptive statistics of the variables in the Colombian electricity market.

Variable (1)	Treatment	Type (2)	Units	Mean	StdDev	Skewness	Kurtosis (3)	Percentile						
								1%	5%	25%	50%	75%	95%	99%
Spot Price	Series	Level	COP/kWh	125	128	4.34	26.6	37	40	64	85	150	263	744
		Logarithm		4.58	0.63	1.01	4.46	3.6	3.7	4.2	4.4	5.0	5.6	6.61
	Detrended (4)	Level	COP/kWh	0	111	4.54	30.7	-149	-106	-37	-8	12	64	538
		Logarithm		0.00	0.43	0.78	5.91	-1.02	-0.70	-0.22	0.00	0.19	0.61	1.44
Demand	Series	Level	GWh	4,552	690	0.06	1.88	3,311	3,502	3,936	4,547	5,121	5,626	5,763
		Logarithm		8.41	0.15	-0.14	1.93	8.10	8.16	8.28	8.42	8.54	8.64	8.66
	Detrended	Level	GWh	0	139	-0.64	4.07	-430	-274	-76	22	89	193	314
		Logarithm		0.00	0.03	-0.71	3.42	-0.086	-0.060	-0.016	0.001	0.022	0.046	0.054
ONI	Series	Level	°C	-0.05	0.81	0.63	3.91	-1.68	-1.40	-0.60	-0.15	0.40	1.31	2.48
	Series (5)	Level	°C	0.02	0.84	0.35	3.11	-1.70	-1.31	-0.50	0.00	0.50	1.50	2.20
NIS Inflow	Series	Level	GWh	4,137	1,631	0.42	3.13	1,175	1,713	2,949	4,096	5,059	6,908	8,630
		Logarithm		8.24	0.44	-0.65	3.08	7.07	7.45	7.99	8.32	8.53	8.84	9.06
San Carlos Inflow	Series	Level	GWh	106	52	0.77	3.20	27	38	68	97	137	200	235
		Logarithm		4.54	0.52	-0.32	2.60	3.29	3.64	4.21	4.57	4.92	5.30	5.46
Guavio Inflow	Series	Level	GWh	496	336	0.81	3.16	63	95	200	436	727	1,110	1,436
		Logarithm		5.94	0.80	-0.47	2.35	4.15	4.56	5.30	6.08	6.59	7.01	7.27
Nare Inflow	Series	Level	GWh	527	246	1.29	5.79	163	228	361	457	652	1,003	1,166
		Logarithm		6.17	0.45	-0.01	2.83	5.10	5.43	5.89	6.13	6.48	6.91	7.06
Salvajina Inflow	Series	Level	GWh	87	47	0.98	3.98	23	28	50	79	116	165	214
		Logarithm		4.32	0.56	-0.23	2.50	3.13	3.32	3.90	4.37	4.75	5.10	5.36

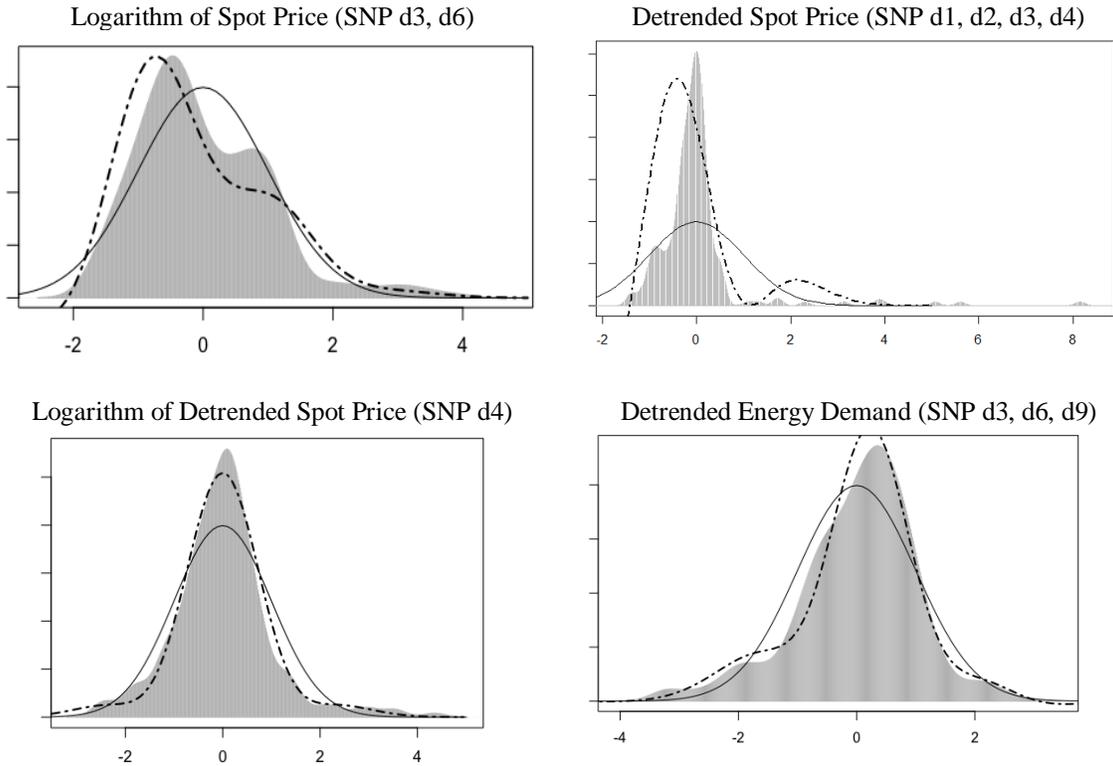
The table contains the descriptive statistics of the variables in the Colombian electricity market considered in this study. (1) Monthly series from January 2000 to April 2018. (2) Level refers to the series without transformations, while Logarithm refers to the natural logarithm of the series. (3) It refers to kurtosis. As a reference, the kurtosis of a normal distribution equals 3. (4) The detrended treatment is produced after removing a linear trend from the series. They are the residuals of a linear model where the series is the endogenous variable and a deterministic time trend is the exogenous variable. (5) This series contains information since January 1950.

5. Results and discussion

Figure 4 presents the SNP probability densities (dotted lines) estimated for Spot Price and Energy Demand, compared to the normal distribution (solid line) and the empirical *pdf* of the data (shaded area). The figure also shows the d_s parameters that better fit each one of the standardized series.

The pictures illustrate how the distribution captures the conditions of asymmetry, kurtosis, and bimodality, outperforming the normal distribution. The logarithm of Spot Price exhibits characteristics of bimodality, a heavy right tail, and a positive skew that are captured by the parameters d_3 and d_4 , which are statistically significant. Detrended Spot Price shows that the SNP distribution (d_1 , d_2 , d_3 , and d_4 being statistically significant) is also preferred than the normal to capture the height of the mode area of the distribution, as well as the right tail. In the case of the logarithm of the Detrended Spot Price, although the normal distribution achieves a better fit than in the other series in Figure 4, SNP captures kurtosis better because it is modeled with the parameter d_4 . Regarding Detrended Energy Demand, the empirical *pdf* presents a positive skew and the left tail of the distribution is heavier than that of the normal distribution. In this case, the SNP significantly improves data fit by including parameters d_3 , d_6 , and d_9 .

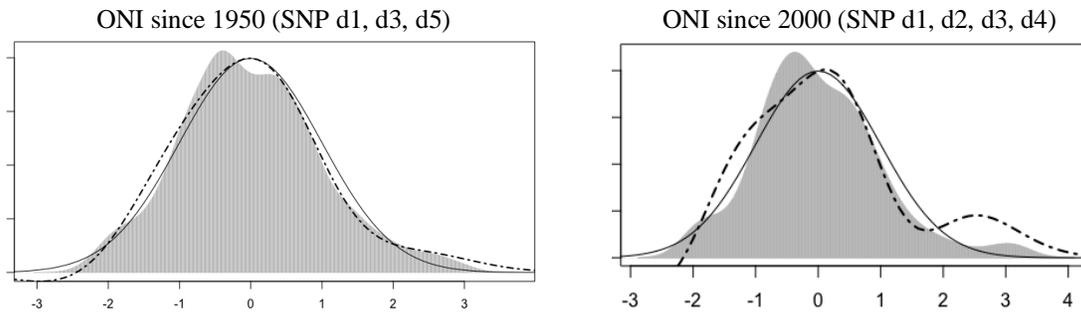
Figure 4. Probability density functions fitted to energy price and demand.



The shaded area represents the bar chart of the data. The standard normal function is marked with the solid line. The dotted line plots the semi-nonparametric distribution.

The ONI is the series that can be best modeled by the normal distribution. Figure 5 shows that the normal and SNP distributions fit the empirical *pdf* similarly. However, when considering ONI information since 2000, the normal distribution does not manage to capture the effect of a heavier right tail. Although the normal distribution may reflect the uncertainty of this series, it does so only for specific conditions of the sample. In turn, the SNP distribution can be adapted to the type of data or specific conditions of each sample.

Figure 5. Probability density functions fitted to the Oceanic Nino Index (ONI).



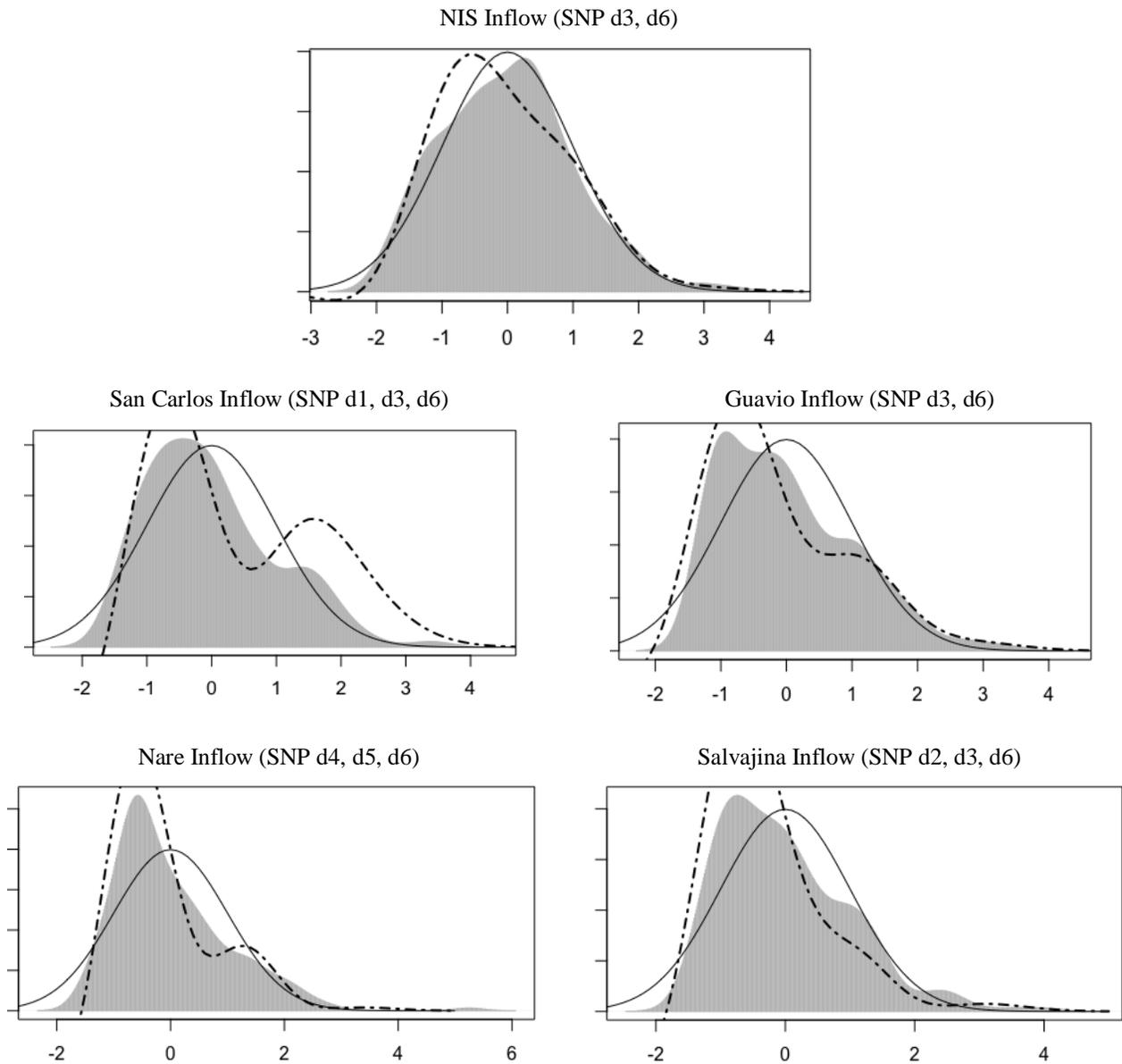
This figure shows the fit of the density functions for the ONI series with information since 1950 and since 2000. The shaded area represents the bar chart of the data. The standard normal function is marked with the solid line. The dotted line plots the semi-nonparametric distribution fitted to the data.

The density functions fitted to the hydrologic inflows NIS are shown in Figure 6. The inflows present a positive skew represented by parameters d_3 and d_5 in the fitted SNP distributions, which is accompanied by bimodality in San Carlos and Guavio, captured by parameter d_6 . The results obtained in these series help to understand the flexibility of an SNP distribution to incorporate the different features of each dataset.

The conditions of positive skew that can be seen in the series of hydrologic inflows are a signal that they commonly present values below average. For example, a hydroelectrical power plant designer who expects a plant to have inflows equal to the series average should consider that, after the plant is built, generation levels will be more commonly below average than above it. As a result, recovering the expected returns from the investment is harder.

Table 3 presents the results obtained with the estimation of the normal and SNP distributions for each variable. Panel A shows the parametric fit for each series, and Panel B does so for each standardized series. All the parameters of the SNP distribution present a confidence level equal to or greater than 95%, except for parameter d_6 of NIS Inflow, whose confidence level is equal or higher than 90%.

Figure 6. Probability density functions fitted to hydrologic inflow.



The figure above presents the fit to the series of hydrologic inflows of all the system and some rivers. The shaded area represents the bar chart of the data. The standard normal function is marked with the solid line. The dotted line marks the semi-nonparametric distribution fitted to the data.

The LR statistic shows that, compared to the Gaussian version, standardized series always present a better fit with the SNP distribution for a confidence level above 99%. Based on that, we conclude that the assumption of normality is indeed a limitation and that the *pdfs* of the SNP type are a more robust tool to represent the uncertainty in the variables of electricity markets.

Table 3. Fit of parametric and non-parametric probability distributions of variables in the Colombian electricity market.

Variable	Treatment	Panel A (Parametric)					Panel B (Seminonparametric SNP added over normalized series)						
		Distrib	Statistics	Mu	Sigma	LogLike	Statistics	Parameters	LogLike	LR			
Spot Price (COP/Wh) (1)	Logarithm of the series	Normal	Coef.	-2.327	0.626	303	d s	d 3	d 6	-292	39.2		
			Std Error	0.042	0.030		Coef	0.230	0.017				
			Pvalue	<0.1%	<0.1%		Pvalue	<0.1%	0.1%				
	Detrended	Normal	Coef	0.000	0.110	173	d s	d 1	d 2	d 3	d 4	-143	338
			Std Error	0.007	0.005		Coef	1.33	-1.57	1.27	-0.136		
			Pvalue	1	<0.1%		Pvalue	<0.1%	<0.1%	<0.1%	<0.1%		
Detrended	LogN (2)	Coef	2.30	0.011	178	d s	d 4				-295	32.7	
		Std Error	0.001	0.001		Coef	0.0918						
		Pvalue	<0.1%	<0.1%		Pvalue	<0.1%						
Demand (TWh) (3)	Detrended	Normal	Coef	0.000	0.139	122	d s (5)	d 3	d 6	d 9	-299	25.8	
			Std Error	0.009	0.007		Coef	-0.120	-0.013	0.0003			
			Pvalue	1.000	<0.1%		Pvalue	0.1%	0.2%	19.5%			
ONI (°C)	Series	Normal	Coef	-0.045	0.811	-266	d s	d 1	d 3	d 5	-1144	30.0	
			Std Error	0.055	0.039		Coef	0.094	0.210	0.035			
			Pvalue	0.406	<0.1%		Pvalue	0.7%	<0.1%	<0.1%			
Series (4)	Normal	Coef	0.020	0.835	-1016	d s	d 1	d 3	d 5	-297	37.9		
		Std Error	0.029	0.021		Coef	0.460	0.657	0.110				
		Pvalue	0.487	<0.1%		Pvalue	0.90%	<0.1%	0.16%				
NIS Inflow (TWh)	Series	Normal	Coef	4.14	1.628	-419	d s	d 3	d 6	-306	11.7		
			Std Error	0.110	0.078		Coef	0.121	0.010				
			Pvalue	<0.1%	<0.1%		Pvalue	0,5%	5,3%				
San Carlos Inflow (TWh)	Series	Normal	Coef	0.106	0.052	338	d s	d 1	d 3	d 6	-276	72.1	
			Std Error	0.004	0.002		Coef	0.932	0.686	0.015			
			Pvalue	<0.1%	<0.1%		Pvalue	<0.1%	<0.1%	0,23%			
Guavio Inflow (TWh)	Series	Normal	Coef	0.496	0.335	-72	d s	d 3	d 6	-287	49.0		
			Std Error	0.023	0.016		Coef	0.277	0.019				
			Pvalue	<0.1%	<0.1%		Pvalue	<0.1%	<0.1%				
Nare Inflow (TWh)	Series	Normal	Coef	0.527	0.245	-3	d s	d 4	d 5	d 6	-271	82	
			Std Error	0.017	0.012		Coef	0.238	-0.111	0.048			
			Pvalue	<0.1%	<0.1%		Pvalue	<0.1%	<0.1%	<0.1%			
Salvajina Inflow (TWh)	Series	Normal	Coef	0.087	0.047	122	d s	d 2	d 3	d 6	-274	74,508	
			Std Error	0.003	0.002		Coef	-0.350	0.423	0.025			
			Pvalue	<0.1%	<0.1%		Pvalue	<0.1%	<0.1%	<0.1%			

This table contains the results of the estimation of parametric (normality) models in Panel A and seminonparametric (SNP) models in Panel B for different series of electricity markets. In the SNP case, the terms of the polynomials are displayed when they are found to be statistically significant in the density specification. LogLike is the log-likelihood of the distribution selected in each case; and LR, the statistic of the Likelihood-Ratio test. (1) One kWh equals 1000 Wh. (2) For this estimation, the logarithm of the series is calculated and a regression is performed between the resulting series and a linear trend. The distributions of the residuals of such regression are fitted. (3) TWh equals 10³ GWh. (4) This series contains information since January 1950.

6. Conclusions and Policy Implications

This study proposes a novel and flexible SNP approach to approximate the distribution of Colombian electricity variables, finding that the Gram–Charlier expansion outperforms the normal distribution due to the multimodal, skewed and leptokurtic nature of the data. The variables analyzed in this work are Energy Demand, Spot Price, aggregated Hydrologic Inflows of the system, ONI, and the series of the Hydrologic Inflows of several rivers: Nare, Salvajina, Guavio, and San Carlos. These series exhibit leptokurtosis and positive skew, but are just an example of the behavior of most electricity markets, where the assumption of normality imposes severe limits to uncertainty modeling and imply misleading energy policies. Based on this result, studies on energy markets should assume that uncertainty is generated by a probability density function of the SNP type, whose normal distribution is a particular case.

The conditions of positive skew that can be seen in the series of Hydrologic Inflows are a signal that they commonly present values below the average, which cannot be captured with the normal distribution. For example, a hydroelectrical power plant designer who expects a plant to have inflows equal to the series average should consider the fact that, after it starts operations, generation will be more frequently below than above average. Such disbalance will hinder the recovering of the return on investments. The Spot Price presents the same conditions, where it is more common to find data below than above the average price. The levels of kurtosis identified in this study are a signal of the presence of extreme events, which may occur frequently and jeopardize the profitability of the investment in some term.

The sustainability of electricity markets depends on their adoption of adequate decision-making methods. Regulatory, supervisory, and control agencies should promote the implementation of the best tools in order to guarantee the sustainability of companies involved in electrical supply. This work provides tools that will complement the best practices of risk management. In this line, risk management in electricity markets should seek to improve analysis with the support of tools that are not limited to the assumption of normality. Such tools should be able to capture the effects of asymmetry, kurtosis, and even high-order moments in probability distributions. The variables of the electrical market present extreme events in an asymmetrical manner; for that reason, risk indicators that separate positive from negative variations in the series are necessary. Portfolio profitability analyses should focus on extreme positive or negative movements of the distributions; therefore, the risk measurements typically used in financial markets can be considered insufficient. Other techniques that have been developed in the field of SNP statistics should be further explored to model electricity markets.

All in all, professionals in this field should consider certain conditions for the treatment of deterministic and stochastic forecasts that may be captured if they are based on the assumption of an SNP uncertainty: (a) The average of the series is not equal to percentile 50, and bimodality conditions may exist. (b) The volatility and kurtosis in the random variables do not have a symmetrical shape. Therefore, the movements on the right side should be treated differently from those on the left. (c) Extreme events are not limited to some variables and can be more common than intuition suggests. (d) The SNP assumption does not compete

with normality; it is a compliment. Even a normal distributed variable should be described by an SNP distribution.

References

- Atalla, T., Bigerna, S., & Bollino, C. (2018). Energy demand elasticities and weather worldwide. *Economia Politica*, 35(1), 207-237.
- Avci, E., Ketter, W., & van Heck, E. (2018). Managing electricity price modeling risk via ensemble forecasting: The case of Turkey. *Energy policy*, 123, 390-403.
- Barton, D. E. & Dennis, K. E. R. (1952). The conditions under which Gram-Charlier and Edgeworth curves are positive definite and unimodal. *Biometrika*, 39(3/4), 425-427.
- Benth, F., Kallsen, J., & Meyer-Brandis, T. (2007). A non-Gaussian Ornstein–Uhlenbeck process for electricity spot price modeling and derivatives pricing. *Applied Mathematical Finance*, 14(2), 153-169.
- Brunner, A. D. (1992). Asymmetries in real GNP: A seminonparametric approach. *Journal of Business & Economic Statics*, 10(1), 65-72.
- Campo, J., & Sarmiento, V. (2013). The relationship between energy consumption and GDP: Evidence from a panel of 10 Latin American countries. *Latin America Journal of Economics*, 50(2), 233-255.
- Cornish, E., & Fisher, R. (1938). Moments and cumulants in the specification of distributions. *Review of International Statistics*, 5(4), 307-320.
- Cortés, L. M., 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: Statistical Mechanics and its Applications*, 485, 35-47.
- Del Brio, E., & Perote, J. (2012). Gram–Charlier densities: Maximum likelihood versus the method of moments. *Insurance: Mathematics and Economics*, 51(3), 531-537.
- 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.
- Edgeworth, F. Y. (1907). On the representation of statistical frequency by a series, *Journal of the Royal Statistical Society*, 70(1), 102–106.
- Falbo, P., Fattore, M., & Stefani, S. (2010). A new index for electricity spot markets. *Energy Policy*, 38(6), 2739–2750.
- Gallant, A. R., Rossi, P. E., & Tauchen, G. (1992). Stock prices and volume. *The Review of Financial Studies*, 5(2), 199-242.
- Gallant, A. R. & Nychka, D.W. (1987). Seminonparametric maximum likelihood estimation, *Econometrica*, 55(2), 363–390.
- 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.
- Gonzalez, J., Moriarty, J., & Palczewski. (2017). Bayesian calibration and number of jump components in electricity spot price models. *Energy Economics*, 65, 375-388.
- Gutierrez-Pedrero, M.J., Tarancon, M.A., del Rio, P., & Alcantara, V. (2018). Analysing the drivers of the intensity of electricity consumption of non-residential sectors in Europe. *Applied Energy*, 211, 743-754.
- Huisman, R., & Hurman, C. (2003). Fat tails in power prices. ERIM Report Series Reference No. ERS-2003-059-F&A. Available at SSRN: <https://ssrn.com/abstract=450992>

- Janczura, J., Trück, S., Weron, R., & Wolff, R. C. (2013). Identifying spikes and seasonal components in electricity spot price data: A guide to robust modeling. *Energy Economics*, 38, 96-110.
- Jarrow, R. & Rudd, A. (1982). Approximate option valuation for arbitrary stochastic processes, *Journal of Financial Economics*, 10(3), 347-369.
- Jondeau, E. & Rockinger, M. (2001). Gram-Charlier densities, *Journal of Economic Dynamics and Control*, 25(10), 1457–1483.
- León, A., Mencía, J., & Sentana, E. (2009). Parametric properties of semi-nonparametric distributions, with applications to option valuation, *Journal of Business and Economic Statistics*, 27(2), 176–192.
- Lucia, J. J., & Schwartz, E. S. (2002). Electricity prices and power derivatives: Evidence from the nordic power exchange. *Review of Derivatives Research*, 5(5), 5-50.
- Maradey, K., Pantoja, J., & Trespalacios, A. (2017). Evaluación de las garantías requeridas para cubrir los riesgos en los mercados de futuros de energía eléctrica. *Ecós de Economía*, 21(45), 67-105.
- 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(2), 225-239.
- Mayer, K., & Trück, S. (2018). Electricity markets around the world. *Journal of Commodity Markets*, 9, 77-100.
- Ñíguez, T.-M., & Perote, J. (2012). Forecasting heavy-tailed densities with positive Edgeworth and Gram-Charlier expansions. *Oxford Bulletin of Economics and Statistics*, 74(4), 600-627.
- Pantoja, J. (2012). Modelling risk for electric power markets. *Innovar*, 22(44), 51-66.
- Pasten, R., Saens, R., & Contreras Marín, R. (2015). Does energy use cause economic growth in Latin America? *Applied Economics Letters*, 22(17), 1399-1403.
- Pérez Odeh, R., Watts, D., & Negrete-Pincetic, M. (2018). Portfolio applications in electricity markets review: Private investor and manager perspective trends. *Renewable and Sustainable Energy Reviews*, 81(1), 192-294.
- Pilipovic, D. (2007). *Energy risk, valuing and managing energy derivatives*. Maidenhead: McGraw-Hill.
- Pinzón, K. (2018). Dynamics between energy consumption and economic growth in Ecuador: A granger causality analysis. *Economic Analysis and Policy*, 57, 88-101.
- Poveda, G., & Mesa, O. (1997). Feedbacks between hydrological processes in tropical South America and large-scale ocean-atmospheric phenomena. *Journal of Climate*, 10, 2690-2702.
- Sargan, J. D. (1975). Gram-Charlier approximations applied to t ratios of k-class estimators, *Econometrica*, 43(2), 327–346
- Sioshansi, F. P. (2013). *Evolution of Global Electricity Markets: New paradigms, new challenges, new approaches (1st ed.)*. Academic Press.
- 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*, 50, 148-157.

- Ullah, A. (2004). *Finite sample econometrics*. New York: Oxford University Press.
- Uribe, E., & Trespalacios, A. (2014). Contraste de modelos estocásticos para el precio de la energía en Colombia. *Revista de la Facultad de Ciencias*, 3(1), 41-55.
- Waylen, P., & Poveda, G. (2002). El Nino–Southern Oscillation and aspects of western South American hydro-climatology. *Hydrological Processes*, 16(6), 1247-1260.
- 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.

Appendix

Proof 1. The parameter of $H_0(y) = 1$ in the SNP expansion is $\delta_0 = 1$.

If parameter δ_s in the Gram-Charlier expansion satisfies $\delta_s = \frac{1}{s!} \int_{-\infty}^{\infty} H_s(y) f(y) dy$,

then $\delta_0 = \frac{1}{0!} \int_{\mathbb{R}} H_0(y) f(y) dy = \int_{-\infty}^{\infty} f(y) dy = 1$, since $f(\cdot)$ is a pdf.

Proof 2. Cdf of truncated SNP.

Let x be a random variable described by the pdf in Equation (5), then its cdf can be obtained as follows:

$$G(y|\mathbf{d}) = \int_{-\infty}^x [1 + \sum_{s=1}^n d_s H_s(y)] \phi(y) dy = \int_{-\infty}^x \phi(y) dy + \sum_{s=1}^n d_s \int_{-\infty}^x H_s(y) \phi(y) dy.$$

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

$$G(y|\mathbf{d}) = \int_{-\infty}^x \phi(y) dy + d_s \sum_{s=1}^n d_s \int_{-\infty}^x (-1)^s \frac{d^s \phi(y)}{dy^s} dy$$

$$\Rightarrow G(y|\mathbf{d}) = \int_{-\infty}^x \phi(y) dy + d_s \sum_{s=1}^n d_s (-1)^1 \phi(y) \frac{(-1)^{s-1} d^{s-1} \phi(y)}{\phi(y) dy^{s-1}} \Big|_{-\infty}^x$$

$$\Rightarrow G(y|\mathbf{d}) = \int_{-\infty}^x \phi(y) dy - \phi(x) \sum_{s=1}^n d_s H_{s-1}(x).$$