(2023), **07-04**, 1–32

MASTER THESIS DOCUMENT

Predictive Modeling for Day-Ahead Pricing in Electricity Markets: A Methodological Approach for Colombia

Sam Minors* and Santiago Manotas-Arroyave*

Barcelona School of Economics, MSc. Data Science Candidate.
*Corresponding author. Email: sam.minors@bse.eu; santiago.manotasarroyave@bse.eu

Abstract

This article provides a comprehensive overview of the use of machine learning and deep learning methodologies in predicting spot prices in liberalized electricity markets, with a focus on the Colombian wholesale energy market. The study presents a methodological framework for detecting potential imperfect competition behavior within the market, which is divided into two primary phases. The literature review highlights the limitations of traditional inferential statistical methods and emphasizes the necessity for more sophisticated techniques to capture the complexities and non-linearities inherent in electricity markets. The study's findings underscore the need for ongoing research in this domain and offer a comprehensive understanding of the field, constructing a robust foundation for a more profound comprehension of the multifaceted dynamics of energy markets and allowing for the envisioning of potential market trajectories in the face of emerging technologies and methodologies. Overall, this study provides valuable insights into the potential for predictive modeling to improve decision-making in the Colombian energy market and beyond; as well as providing exhaustive metrics for out-of-sample architecture comparison.

Keywords: Predictive Modeling, Machine Learning, Energy Markets.

Supervisors[0]:

- Hannes Müller, Ph.D. - IAE-CSIC and BSE.
- Christian Brownlees, Ph.D. - UPF and BSE.
- Elliot Motte, Ph.D. Candidate - UPF.

A thesis presented as part of the fulfillment of the Barcelona School of Economics Master's in Data Science for Decision Making.

[0]Institut d'Anàlisi Econòmica - Universitat Autònoma de Barcelona, Universitat Pompeu Fabra.

Contents

List of Figures

List of Tables

1. Introduction

1.1 Motivation

In likewise manner to that of Riascos, Chitiva, and Salazar (2022) and Poggi, Di Persio, and Ehrhardt (2023), we make use of information from the Colombian energy market, liberalized during the 90s and with its latest update in regulations and configuration in 2010. These liberalizations during such period responded to a need to promote competition and efficiency. Nevertheless; intrinsic features such as inelastic demand, market concentration, and high storage costs pose a challenge to achieving this Wolak (2003), Newbery (2008).

Here, the possibility to make spot pricing predictions becomes imperative for policy makers and corporate stakeholders regardless of individual motivation, as this poses an opportunity to reap benefits or counteract against market anomalies. This capability offers an opportunity to capitalize on favorable market conditions or counteract any anomalies that may arise. Therefore, our research objective from this study is to establish a tool that can effectively reflect and raise alerts to central planners when the electricity market exhibits variations that require price forecasting. Such alerts can prompt further investigation into potential instances of market power exertion by certain agents. It is important to note, however, that our methodology provides grounds for further assessment rather than offering open-and-shut evidence in itself. In this, we highlight the leverage of extant artificial intelligence (**AI**) to assist forecasting with Machine Learning (**ML**) archetypes.

In hand with this, the controlling body in Colombia[1] introduced a tool during 2017 to monitor the electricity market and foster competition among electricity generators. This tool aims to encourage generators to adopt more efficient and transparent business practices, monitoring their behavior in three scenarios: bilateral contracts, energy exchange, and positive reconciliations. Additionally, it is expected that the contracting mechanisms being developed by various market agents will contribute to improving market efficiency and transparency.

1.2 Characterization of the Colombian Energy Market

Upon closer examination of the Colombian energy market, a crucial point of departure is understanding the composition of the market by technology. As of December 31, 2022, the energy landscape in Colombia is quite concentrated, featuring a mix of renewable and non-renewable sources. The market's total capacity, 18.7 GW, is distributed across five primary energy technologies: Biomass, Thermal (Coal/Gas), Wind, Hydro, and Solar. Figure 1 reveals these preliminary numbers.

Given Colombia's access to an abundant river network, Colombia generates 66.8% of total capacity from hydro power. Yet, the country's rainfall is impacted by the El Niño climate system and can affect the availability of this largest generating source.

The second largest generating source is Thermal power (Coal/Gas) technology representing approximately 30.5% of the total capacity. Predominantly non-renewable, these thermal power plants provide a substantial part of the energy mix. Their role has been especially significant in periods of low rainfall when hydroelectric power production may diminish.

To aid in this delineation of the inner workings of said market, Figure 2 displays the plants that on average had the biggest market share across 2022 for the generation of power at the country-grid level. As consistent with our first showing, these plants corresponding almost exclusively to hydro generation. As we move forward, we will evaluate whether this outcome of the first-stage power market is congruent with the goals of controlling bodies. Forbye to this overview, the Appendix 1 includes an exhaustive explanation of the intricacies of this environment as set by law.

[1] *Superintendencia de Industria y Comercio* – Superintendency of Industry and Commerce (SIC), by its initials.

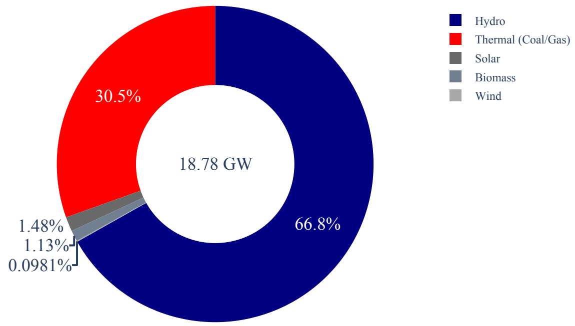

Figure 1. Installed capacity by generation technology (Dec 2022). Source: XM. Authors' elaboration.

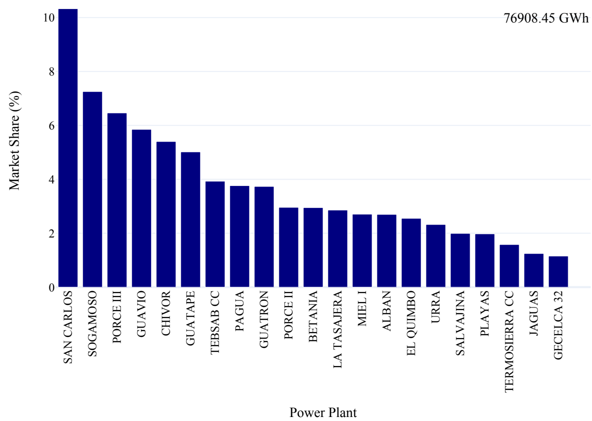

Figure 2. Top plants by capacity for the overall market (2022). Source: XM. Authors' elaboration.

2. Literature Review

This literature review is fundamentally structured around three main axies. Firstly, it explores the liberalization of electricity markets as a strategy to enhance efficiency and quality within generation markets; secondly, it scrutinizes the strategic behavior of agents participating in such markets; and finally, it delves into the utilization of machine learning techniques to detect potential behaviors that might impact competitive levels within the market.

Despite the fact that reforms geared towards the liberalization of energy markets initiated around 1990 with the British market model, there is still a recognized need for the continuous adjustment and improvement of institutional designs to increase competition within these markets. Given their characteristics, the introduction of the spot market was a significant contribution from the British model towards nurturing competition in electricity markets in the 1990s; countries such as Colombia have adopted said process of liberalizing the spot energy market since 1995.

According to Haselip (2007) and Weinmann (2007), within the developing world, Latin America has been at the forefront of implementing far-reaching reforms in the electricity sector. Beginning with the gradual privatization of utility companies in Chile in the 1980s, the main wave of restructuring, disaggregation, and creation of competitive wholesale markets expanded in the first half of the 1990s from Chile to its neighboring countries Argentina, Bolivia, and Peru; it later extended to Colombia, Brazil, and several states in Central America and the Caribbean.

Although the fundamental hypothesis of electricity market liberalization models is the increment in competition amongst agents, it is evident that the technical characteristics of these markets demand continuous monitoring of potential strategic behaviors by the participating agents (suppliers):

- Inelastic demand.
- Market concentration.
- High energy storage costs.

This monitoring aims to prevent agents from maximizing their individual profitability at the expense of social welfare.

Several authors have been conducting studies to understand the behavior of energy prices in countries that have formalized processes of liberalizing their energy sector. Nagayama (2009) formulated an empirical model to observe the impact of electricity prices on the selection of an electricity sector liberalization model. His analysis included data from 78 countries in the period from 1985 to 2003. As a result of his work, it is concluded that the development of liberalization models in the electricity sector does not necessarily reduce electricity prices. In fact, contrary to expectations, there was an upward trend in price in all modeled markets.

Regarding the specific context of the Colombian electricity market, authors such as Correa-Giraldo, Garcia-Rendon, and Perez (2021) and Suarez (2022) have highlighted the importance of analyzing the strategic behavior of market participants. Their results suggest that, despite reforms aimed at increasing competition in the electricity sector, high levels of concentration among some market participants result in certain market power. This power is associated with the strategic behavior of generation resources under certain conditions and uncertainty.

Moreover, the need to promote the integration of non-conventional renewable energies into Colombia's generation matrix has led some authors, such as Mastropietro et al. (2020), to propose a regulatory framework to introduce a multiple settlement system. This system comprises a binding daily market, followed by intraday sessions and a balancing market. The main debate focuses on how to address the complexities arising from the introduction of a multiple settlement system in a context where sessions are settled on the basis of uniform prices.

On the other hand, despite the original intention to promote competition in energy markets through structural reforms aiming at the liberalization of these markets, a particular interest has arisen in the literature to analyze the market power that agents can exert. Studies in this regard are broad and diverse.

Prabhakar Karthikeyan, Jacob Raglend, and Kothari (2013) present a comprehensive review on market power, incorporating several indices used in the analysis of market power, and chronicling the evolution of research and development in the field of market power within energy markets that embarked on liberalization reforms aimed at fostering competition.

Fabra and Toro (2005) provided an example of the early method to forecast price and potential anomalies. To analyse time-series data, they employed a time-varying transition-probability Markov-switching model. Although theoretically challenging, the model uses conventional econometric techniques and provides an easy way to spot distinctive price patterns and possibly collusive behaviours. The model's straightforward interpretation of the data is made possible by its simplicity, but it lacks predictive power and may oversimplify complicated market behaviours.

Lundin and Tangerås (2020) posit that the horizontal shifts observed in supply curves in wholesale electricity markets are consistent with Cournot-type models. These authors scrutinized the

Scandinavian market, Nord Pool, during the period of 2011-2013. They rejected the null hypothesis of perfect competition in all specifications. Their findings suggest that the average cost-price margin throughout the sampling period was approximately 4%.

In the same context, Fridolfsson and Tangerås (2009) found no evidence of systematic exploitation of market power at the system level in Nord Pool. According to the authors, local market power stemming from transmission limitations appears to be more problematic in certain pricing zones of the Nordic countries.

In the literature, there are other studies that analyze the market power of agents in the electric systems of different regions and countries. Ciarreta, Nasirov, and Silva (2016) conduct a comprehensive analysis of the market power problem in the Spanish electricity generation sector, and they examine how and to what extent the market has evolved in terms of market power issues following market liberalization reforms. The methodology applied in this study includes the typical structural and ex-post behavioral measures used to estimate the potential for market power, namely: concentration ratios (CR) (for the largest and the three largest providers), the Herfindahl-Hirschman Index (HHI), Entropy, the Pivotal Supply Index (PSI), the Residual Supply Index (RSI), and the Residual Demand Elasticity (RDE). The results are presented for the two largest Spanish generating companies (Endesa and Iberdrola) that operate in the Iberian Electricity Market (MIBEL), and in the Spanish daily electricity market. The results show evidence that these companies have behaved much more competitively in recent periods than at the beginning of market liberalization. Additionally, the article discusses significant structural and regulatory changes through market liberalization processes in the Spanish daily electricity market.

Mulder and Schoonbeek (2013) evaluated the influence of a series of events on the degree of competition in the Dutch wholesale electricity market during the period 2006-2011 using a decomposition method based on the Residual Supply Index.

Lise, Hobbs, and Hers (2008) use a static computational game theory approach for a fully open European electricity market and can take into account the strategic interaction among electricity producing companies: Firstly, in the scenario of perfect competition, prices differ due to the presence of various generation technologies and the limited capacity to exchange electricity between countries. In addition, when large companies exercise their market power, model executions indicate that prices are higher in countries where the number of companies is low. Secondly, dry weather would increase prices in the Nordic countries, which are rich in water resources, followed by the Alpine countries. The price response would be 20% higher with market power. Thirdly, greater transmission capacity would reduce prices in countries with high prices and would also reduce the impact of market power. Therefore, greater transmission capacity can improve market competitiveness.

The third and final strand of the literature review, delves into the burgeoning field of machine learning and Deep Learning (**DL**) techniques utilized for the forecasting of spot pricing in electricity generation markets; a realm of study that has seen rapid advancement in the recent years. This burgeoning area of research harnesses the power of advanced algorithms and computational models to accurately predict fluctuations in spot pricing, and to detect anomalies within these markets. The effectiveness and efficiency of these techniques are continually being refined, making them a promising tool for future market analysis. Throughout this section, various influential and pioneering studies that have shaped this sub-field are critically examined and discussed; with the aim of highlighting their implications amongst the state of the art and the future of electricity market price forecasting.

Foremost, a transition from simple statistical methods to cutting-edge machine learning and AI techniques is evident; as complexity, prowess, and capabilities grow exponentially.

The article by Razmi, Buygi, and Esmalifalak (2021) discusses a machine learning approach for detecting collusion in electricity markets based on Nash equilibria assumptions. The methodology involves using Support Vector Machines (**SVM**) and Classification and Regression Trees (**CART**) algorithms to identify colluding firms using data related to market equilibrium for machine training and validation. The results show that SVM outperforms other methods due to its capability in providing a good out-of-sample generalization with proper parameter tuning (specifically for C and γ, defined variables). The study concludes that machine learning approaches can be useful in detecting collusive behaviors in electricity markets; an otherwise complex process, where it is particularly hard to find reasons and rules for justifying safe versus fraudulent behaviors.

Garcia Pires and Skjeret (2023) present a comprehensive analysis of screening methods for detecting partial cartels in retail electricity markets. The study focuses on the Norwegian retail electricity market, utilizing data from the NordPool wholesale price. The uncoverings of the study reveal several key insights.

Firstly, the article highlights that most available screen methods for cartel detection are based on measures of centrality and dispersion, such as mean and standard deviation of prices. In spite of this, these measures are not derived from economic theory and were originally developed for full cartels. Ergo, the study proposes the use of third and fourth moments to control for patterns related to seasonality and production costs.

Furthermore, the research identifies two periods in the dataset where the pricing of electricity contracts changed abruptly. These changes are statistically significant and cannot be explained by factors on the supply side. The observed price decreases during the late summer of 2013 resemble what a coalition breakdown would look like, according to the theory of coalition formation. Correspondingly, the theory suggests that different subsets of firms choose different pricing strategies; with insiders opting for higher prices and outsiders choosing lower prices. The findings from the analysis of the Norwegian retail electricity market differ from the theoretical framework, providing valuable viewpoints for future research and policy-making in the field of electricity market regulation.

The work by Esmaeili Aliabadi and Chan (2022) undertakes an extensive exploration of Deep Q-Network (**DQN**) algorithms and their role in the prediction of spot pricing within electricity markets. It pays particular attention to the bidding strategies employed by Power Generating Companies (**GenCos**) in deregulated electricity markets, further scrutinizing the potential influence of learning algorithms on competition and overall market outcomes. Employing a multifaceted case study; replete with numerous Nash and collusive equilibria, the authors evaluate the capability of DQN models to perpetuate collusive behaviors within these markets. The revelations from this study augment the existing body of knowledge by showcasing the efficacy of DQN algorithms in predicting spot pricing, thence underscoring their potential to upset competition within deregulated electricity markets. The study also pinpoints practical impediments that could possibly obstruct the real-world application of these techniques. The results of this investigation elucidate the potential of learning algorithms to shape market outcomes and emphasizing the importance of understanding the repercussions of emerging technologies on fair competition.

Vega-Márquez et al. (2021) provide a thorough investigation into the utilization of deep learning architectures for forecasting day-ahead electricity prices within the Spanish electricity market. The central aim of the research is to navigate the intricate challenges inherent in electricity price prediction; such as the multitude of influential factors, non-linearity, lack of seasonality, and the price volatility over time. The researchers juxtapose various renowned models, optimizing their parameters with data from three distinct time spans: a normal period, a fraud period, and a quarantine period during the COVID-19 pandemic.

The results from this study underscore the efficacy of deep learning models in forecasting electricity prices, demonstrating a superior performance compared to other established models in the literature. This finding emphasizes their potential to reliably encapsulate the complexities of the electricity market, while highlighting the importance of temporal considerations and the specific characteristics of the dataset when developing forecasting models. The study also underscores the necessity to incorporate exogenous variables, such as demand and renewable energy capacity, to amplify prediction accuracy; it further illustrates the influence of social and economic contexts on electricity market prices.

The journal article Poggi, Di Persio, and Ehrhardt (2023) delivers a detailed examination of models, both traditional statistical methods and modern deep learning techniques, used for forecasting electricity prices, with a specific focus on the German electricity market using historical data from 2020 to mid-2022. The authors suggest an innovative approach to predict electricity spot prices through the decomposition of the spot price series into a seasonal trend component and a stochastic one, allowing for highly accurate predictions across varying time frames.

The conclusions of this research underscore the effectiveness of the proposed models in accurately forecasting electricity spot prices; they highlight the potential of machine learning-based solutions to improve the precision and reliability of electricity price predictions. At the same time, the study illuminates the limitations of traditional inferential statistical methods, emphasizing the necessity for advanced techniques to encapsulate the complexities of electricity markets.

In their seminal work, Lago, De Ridder, and De Schutter (2018) introduce an innovative modeling framework for forecasting spot electricity prices utilizing deep learning methodologies. Distinct deep learning architectures (4) are proposed by the authors, their predictive accuracy meticulously compared against 27 prevalent methods used in electricity price forecasting. This investigation addresses a significant gap in the field by thoroughly examining the application of deep learning algorithms in this domain, a relatively unexplored area. The upshots depict that these methods surpass state-of-the-art techniques, delivering statistically significant enhancements in predictive accuracy.

Then, the study reveals that machine learning methods generally exhibit superior accuracy to statistical models; intriguingly, the incorporation of moving average terms does not augment predictive accuracy, and hybrid models do not outperform their simpler equivalents. This research adds to the current body of knowledge by providing a comprehensive benchmark study, underscoring the efficacy of deep learning.

In summary, the historical progression from simple statistical models to complex machine learning and deep learning techniques signifies a marked increase in the complexity of models deployed for collusion detection in energy markets. This evolution in model sophistication presents potential benefits; these include the ability to manage intricate data structures, enhanced accuracy in prediction, and adaptability to non-linear interactions. Yet, these advancements come at the cost of heightened computational demand. The growth in complexity offers a more nuanced understanding of machine learning techniques in the context of the intricate dynamics of GenCos markets, thereby providing invaluable insights for both actors and policymakers.

The literature underscores the efficacy of machine learning and deep learning methodologies in predicting spot prices, highlighting their potential to disrupt competition in liberalized electricity markets. To boot, the research illuminates the limitations of traditional inferential statistical methods and emphasizes the necessity for more sophisticated techniques to capture the complexities and non-linearities inherent in electricity markets.

In conclusion, the synthesis of literature spanning across three distinct yet interconnected axes: the advent of liberalized energy markets, strategic behavior in within these scenarios since the turn of the century, and the employment of machine learning for spot price predictions on markets. In conjunction, these offer a comprehensive understanding of the field, constructing a robust foundation, enabling a more profound comprehension of the multifaceted dynamics of energy markets and allowing for the envisioning of potential market trajectories in the face of emerging technologies and methodologies. This extensive literature review underscores the need for ongoing research in this domain.

3. Methodology

The objective of this study is to present a comprehensive methodological framework aimed at detecting potential anti-competitive behavior within the Colombian wholesale energy market. Our analysis is divided into two primary phases.

Firstly, we employ market concentration indices such as the Residual Supply Index (**RSI**) to identify the key market participants. It is noteworthy that the Herfindahl–Hirschman Index (**HHI**), a traditional measure of market concentration, is not appropriate in our context, for reasons we elaborate upon subsequently. It is inferred that the declared availability from these critical participants may significantly impact the final price.

In the second phase, we leverage a historical dataset comprising variables like past daily offers, supply, demand, and hourly prices, among others, from multiple market agents. This dataset is utilized in conjunction with machine learning techniques to predict energy offers (ask prices) from all agents. Based on the resultant RSI and the predicted offers, potential alerts of anti-competitive practices can be triggered.

In our pursuit of conducting an exhaustive exploration of the Colombian energy market, we rely on historical data directly sourced from the official market administrator, Sinergox (**XM**). Post data acquisition, an extensive cleaning and compiling process is undertaken, which includes translating column names into English, standardizing all variables, conducting descriptive statistical analysis to assess distribution characteristics and ensuring stationarity of the time series (see Predictive Modeling & Results section).

Using data supplied by the market administration, we examine correlations with power plant asking prices. A standard OLS regression for a specific year is executed to verify these associations, as elaborated in the data section of the document.

In accordance with the regulatory norms stipulated in Colombian law (refer to Appendix 1), specific columns in our datasets are utilized that offer pertinent market information, potentially impacting pricing decisions under the assumption of a competitive market scenario. It is important to note that in defining the deep learning architectures, variables are included since these methods account for non-linear correlations.

Outlier detection and removal is an integral part of the pre-processing phase. We exclude instances where non-competitive bids could have been offered, for instance, when a power plant offers an exceedingly high (low) price to guarantee inclusion (exclusion) in the Optimal Economic Dispatch **OED** for the subsequent day. For outlier detection, we implement a rolling yearly interquartile range (**IQR**) calculation method, which offers robustness against outlying values in comparison to standard deviation rules such as the one used by Riascos, Chitiva, and Salazar (2022).

To identify which agents can have a substantial impact on the stock price the Residual Supply Index (RSI) was calculated. The RSI[2], computed as the ratio of the entire supply less one's own to the total demand at the given timestamp, illustrates the relative significance of each observation and provides an essential metric for market entities to exert pricing power, the formula shown below[3]:

$$RSI_{i,t} = \frac{\sum_{j \in N} Q^S_{j,t} - Q^S_{i,t}}{Q^D_t}$$

.

Following the outlines of seminal theoretical considerations from Sheffrin (2002), on the mi-

[2]Due to its formulation, RSI ranges from 0 to ∞, with critical value at 1. Below this, an agent is crucial to meet market demand, over this, the agent is not pivotal.

[3]S and D superscripts in formulas stand for supply and demand, respectively.

croeconomics of electricity markets; a more cross-sector traditional measure as is the HHI

$$HHI_t = \sum_{i \in N} Q^{S^2}_{i,t}$$

, does not adequately measure concentration. There are 2 main issues with it: The first issue stems from the fact a company's capacity could be markedly larger than its market share; it may be contingent on the strategic maneuvers of market participants.

Second, the electrical market is characterized by temporal changes, with an annual total of 8,760 separate hourly markets. Even within the same geographical space, each of these auctions has distinct demand and supply characteristics. As a result, while market concentration may appear to be consistent over time, the underlying competitiveness of the market and the exercise of market power by individual participants might fluctuate significantly. This temporal variation in market conditions hampers the interpretation of traditional market concentration measures, emphasizing the need for a more nuanced understanding of energy market dynamics.

Together with the former, we introduce a 2-version market index for pricing power, i.e., the Lerner index Hansen (2011), calculated in percentage terms: (%):

$$L_{i,t} = \frac{P^S_{i,t} - MC_t}{P^S_{i,t}} = \frac{Markup}{P^S_{i,t}}$$

We apply 2 variants of the Lerner index, both computed and found in our processed data, as detailed subsequently:

- Under a competitive market assumption where $P^D_t = MC_t = P^S_t$.
- Under a structural approach with marginal cost information from the administrator, in the repository as `h_marginalcost.csv`.

This starting stage is substantially pertinent as to whether the regulations outlined in paper do confirm or promote market competition. As mentioned in the introductory part, inherent market features give reasonable doubt to evaluate the reality of these assumptions. Furthermore, we assess the relationship of said market concentration as highlighted in Table 1, where we compute a simple OLS regression of plant mark up, versus agent's RSI:

$$Markup_{i,t} = \beta_0 + \beta_1 \cdot RSI_{I,t} + \epsilon_{i,t}$$
$$Markup_{i,t} = \alpha_t + \beta_3 \cdot RSI_{I,t} + \varepsilon_{i,t}$$

Table 1. Fixed Effects **FE** and Random Effects **RE** coefficients.

Variable	Fixed Effects		Random Effects	
	Coefficient	Standard Error	Coefficient	Standard Error
RSI	-32.80**	13.11	-32.73***	0.47

*** p<0.01, ** p<0.05, * p<0.1

The panel regression results presented in Table 1 clearly indicate the importance of the Residual Supply Index (RSI) in influencing the markup. Both the Fixed Effects (FE) and Random Effects (RE) models showcase a statistically significant and negative relationship between these two variables, which is aligned with the theoretical expectations.

The FE coefficient of RSI is −32.80 ($p < 0.05$), while the RE coefficient stands at −32.73 ($p < 0.01$). Notably, the close alignment of these coefficients from two different model specifications underscores the robustness of our findings. These results essentially highlight that an increase in RSI - indicating less market power by the individual agent - is associated with a decrease in price manipulation, thereby evidencing a pivotal role of RSI in curbing the potential for market power abuse.

The utilization of FE models offers the advantage of controlling for time-invariant unobserved characteristics that are unique to each plant, accordingly focusing on the within-entity variations over time. Consequently, the significant FE coefficient for RSI in our study suggests that changes in a plant's RSI over time have a meaningful impact on its markup behavior.

Our findings are consistent with the results reported by Marin, Orozco, and Velilla (2018) and Riascos, Chitiva, and Salazar (2022), who also found a significant role of RSI in the context of the Colombian first-stage electricity market. Hitherto, our study not only validates the relevance of RSI as an empirical predictor of price markups but also offers compelling evidence for its crucial role in mitigating the potential for market power abuse in the Colombian energy market.

The successive stage of our research methodology entails the utilization of machine learning techniques; our goal is to generate price predictions based on the extant data and subsequently measure the discrepancies between the predicted prices and the actual ask prices. Detailed methodologies for modeling these predicted ask prices are elaborated upon in the forthcoming sections of this report. We also draw parallels with three distinct statistical models, particularly with respect to auto-regressive terms, viewing these models as increasingly encompassing algorithms.

The final stage of our research process involves assessing the disparities between the predicted and actual ask prices, juxtaposing these differences against an agent's market power. The underlying objective here is to discern whether agents with significant market power tend to inflate the markup component of their ask prices; such a trend, if observed, could potentially signal the existence of anti-competitive practices within the Colombian wholesale electricity market. To measure these discrepancies, we need a metric that is sensitive to outlying true values vis-à-vis predictions. Accordingly, we adopt a 2 standard deviation measure to gauge these divergences.

An in-depth discussion of these final two steps is offered in the results section (see Predictive Modeling & Results).

4. Data

4.1 Overview

Our analysis involves a detailed investigation of specific data points; notably, the price at which each power plant offers its generation, the total capacity, the daily market demand, stock prices set hourly in the open market, quantities demanded, marginal cost of plant operation, and the Southern Oscillation Index (SOI[4]). These primary data points provide the foundation for the construction of additional variables such as market share, Residual Supply Index by agent, and individual plant markup.

The frequency of our data is hourly; this is represented by the stock price of energy. It is of crucial significance to recognize that not all power plants necessarily contribute to power generation every day. Power generation is instead determined by those power plants that fall just under the Optimal Economic Dispatch (OED) curve. Accordingly, we refine our analysis by focusing on those power plants and agents that actively participated in the supply for each specific day. This filtering process allows for a more accurate representation of the dynamics within the market and affords a more nuanced understanding of market behavior. However, it is important to understand that whilst stock prices are set hourly; asking prices of plants are daily (at which plants offer supply curves for each hour during the day). We perform adequate propagation of these variables to increase their

[4]Extracted from the National Oceanic and Atmospheric Administratiion (**NOAA**) website: https://www.ncei.noaa.gov/access/monitoring/enso/soi.

granularity to our desired timestamps. Originally, the SOI is also computed at a different frequency:

$$SOI = \frac{SLP_{Tahiti} - SLP_{Darwin}}{\sigma_{monthly}}$$

Specifically, the Southern Oscillation Index is a standardized measure that captures the differential in observed sea-level pressure (**SLP**) between Tahiti and Darwin, Australia. This index provides a valuable gauge of the large-scale oscillations in atmospheric pressure occurring between the western and eastern sections of the tropical Pacific, particularly during the episodes of *El Niño* and *La Niña*.

In essence, the SOI offers a reliable correspondence with changes in the temperature of the ocean across the eastern tropical Pacific. The index's negative phase is characterized by lower-than-average air pressure at Tahiti, coupled with higher-than-average air pressure at Darwin. Durations of negative (positive) SOI values align with unusually warm (cool) ocean waters across the eastern tropical Pacific, which are symptomatic of El Niño (La Niña) events. Given Colombia's tropical location, the El Niño phenomena plays a significant role in shaping the overall energy generation and is hence integral to our analysis; outlined by the dominance of hydro-generating power plants which are directly affected by drought periods present during El Niño.

Conventionally, in forecasting price trends within markets, we frequently rely on standard variables, which encapsulate the dynamics of supply and demand, as well as the prevailing price levels. These key indicators offer insights into the underlying market mechanisms, shaping the trajectory of future prices. Table 2 showcases time-series these variables.

Table 2. Types of variables. From: XM, NOAA.

Variable	Type	Raw Frequency	Unit
Ask Price	Continuous	Daily	$COP/kWh
Bid Price	Continuous	Hourly	$COP/kWh
Technology	Categorical (4)	Plant-dependent	-
Fuel	Categorical (9)	Technology-dependent	-
Capacity	Continuous	Hourly	kW
SOI	Continuous	Monthly	-
Marginal Cost	Continuous	Hourly	$COP
Generation	Continuous	Hourly	kW
Demand	Continuous	Hourly	kW
Supply	Continuous	Hourly	kW

Moreover, we describe the differences for energy-unit based variables; demand, supply, and generation hourly might not be the same for several reasons. Electricity grids balance supply and demand in real-time, there may be unanticipated changes in demand or issues with generation (such as a power plant going offline unexpectedly). On top of this, it's also crucial to account for transmission and distribution losses in the network. The close relationship between categorical variables may (for exclusively statistical models) ensue the existence of the dummy variable trap, causing perfect multicollinearity in the context of a multiple regression model. These are used in further architectures: Recurrent Neural Networks (**RNN**s); and other machine learning methods that provide advanced techniques that can manage high-dimensional categorical data more effectively, potentially mitigating issues:

- Technology, the energy conversion used - thermal, solar, hydro, eolic
- Fuel, raw material used for energy conversion - bagasse, gas, coal, biogas, diesel and marine diesel for thermal; solar radiation; water; wind.

Although a wide range of technology-related information is available and extracted, as shown in the Introduction, 3 distinct generation modules represent the biggest share of Colombia's grid. Besides these variables described, data extracted from the source also includes plant reference, agent (parent ownership), and heat rate (amount of energy used by an electrical generator/power plant to generate one kilowatthour - **kWh** of electricity).

4.2 Exploratory Data Analysis

4.2.1 Descriptive Statistics

Table 3. Descriptive statistics at level. Source: XM. Authors' calculations.

Variable	Measures				Test p-value		
	Mean	SD	Skewness	Kurtosis	JB	ADF	KPSS
Capacity	3613920.52	4427970.12	0.80	-0.90	0.00	0.00	0.01
Ask Price	316.29	332.27	2.55	11.11	0.00	0.00	0.01
SOI	0.36	0.95	0.09	0.03	0.03	0.00	0.01
Bid Price	189.84	159.18	3.82	24.08	0.00	0.00	0.02
Marginal Cost	176.90	157.68	3.77	22.66	0.00	0.00	0.04
Generation	51100.70	5983.16	-0.03	-0.69	0.00	0.11	0.01
Demand	246096.00	32973.70	0.06	-0.62	0.00	0.22	0.01
Supply	168262.00	52643.30	-0.42	-0.73	0.00	0.95	0.01

The descriptive statistics of the dataset, as presented in Table 3, offer several significant insights:

- Capacity: The descriptive statistics indicate substantial variance, a slight right skewness, and a platykurtic distribution. This aligns with our earlier observations regarding the dominant share of hydro plants in the market. In relation to our investigation, this could suggest that the wide range of sizes for resources allows for the optimal dispatch.
- Ask Price: The significant variation and high positive skewness in the ask prices, along with the leptokurtic distribution, underscore the need for robust filtering mechanisms when utilizing this data for future pricing predictions.
- SOI: Despite the near symmetrical distribution and low kurtosis. Nevertheless, understanding the SOI necessitates examining individual values to discern the prevalent phenomena.
- Bid Price: The broad variation and high positive skewness, indicative of peak hour demand prices, in combination with the leptokurtic distribution, underpin the dynamic nature of the demand prices at peak and valley hours.
- Marginal Cost: The considerable variability, high positive skewness, and high kurtosis of marginal costs suggest that they can frequently deviate from the mean, significantly affecting the pricing viability of power plants.
- Generation: The descriptive statistics of power generation signal that a minority of power plants typically generate less power than most others, reinforcing the importance of understanding this distribution for effective management and forecasting, whilst indicating the role of "stronger" resources (plants).
- Demand: The slight right skewness of demand indicates instances of significantly higher power demand, at those aforementioned specific hours; 19:00 to 21:00, and 00:00 to 5:00 respectively.
- Supply: The slight left skewness of supply statistics could point towards periods of significantly lower power supply. Correlating these periods with events such as El Niño, which typically result in higher prices and lower supply, especially for hydro generation will be crucial for day-ahead predictions.

The previous descriptive statistics table outlines the characteristics of the variables, most variables appear to have statistically significant Jarque-Bera (**JB**) tests, registering non-normality; although given the sufficient amount of data points we consider, this does not entail a restriction for our employed methods.

Table 4. Descriptive statistics after log transformation and differencing (1).

Variable	Test p-value		
	JB	ADF	KPSS
Capacity	0.00	0.00	0.10
Ask Price	0.00	0.00	0.10
SOI	0.00	0.00	0.10
Bid Price	0.00	0.00	0.10
Marginal Cost	0.00	0.00	0.10
Generation	0.00	0.00	0.10
Demand	0.00	0.00	0.10
Supply	0.00	0.00	0.10

4.2.2 Stationarity Testing

A meticulous inspection of our datasets, as substantiated by multiple illustrative figures (in the following part); sheds light in a significant degree of non-stationarity inherent in several variables. This non-stationarity introduces an intricate layer of complexity when on forecasting endeavors and employing regression analysis. For this reason, we underscore the necessity of evaluating the stationarity of these data series; a critical step towards ensuring the validity and reliability of our forthcoming predictions.

At levels (Table 3), stationarity testing according points to a possible structural change in the data. The latter is verifiable by the corresponding graphs. In example, for the variable supply exhibiting an upward trend in the total energy supplied to the grid through the years. Which also holds in the case of the total quantity demanded.

We implement the Augmented Dickey-Fuller (**ADF**) test; a prominent statistical tool for identifying non-stationarity in a time series. The ADF test operates under the null hypothesis that a unit root is present in an auto-regressive model representing the time series. A non-significant p-value fails to reject this null hypothesis, thereby indicating non-stationarity.

Moreover, the Kwiatkowski–Phillips–Schmidt–Shin (**KPSS**) test serves as an additional measure to ensure the robustness of our analysis. Unlike the ADF, the KPSS test posits the null hypothesis that the series is stationary around a deterministic trend. Thus, a significant p-value contradicts this hypothesis, signaling the presence of trend-stationarity.

Ideally, for a time series to be deemed stationary, we would aim to reject the H_O of the ADF test whilst failing to reject that of the KPSS.

In a typical time series analysis, the use of both tests in conjunction is quite common. The confluence of these two conditions bolsters confidence in the stationarity of the time series. However, should the tests deliver contradictory results, it necessitates a more detailed examination of the series and possibly the need for further transformations to achieve stationarity. Such is the case of our data, where Table 3 counter-intuitively rejects both. After proper transformations to achieve stationarity are included, the results adjust to the ideal tandem mentioned, as exposed by Table 4.

4.2.3 Graphs

Several intriguing findings surface when analyzing the graphs presented here. Examining the distribution of price offerings presented through Figures 7, and 8; a noteworthy observation is the high concentration of prices at the lower end of the range. This asymmetry suggests that while there are instances of remarkably high price offerings, they are outliers against the backdrop of predominantly low to moderate prices. The relatively scant occurrences of high prices could be attributed to a variety of factors such as peak demand periods, plant operational issues, or (as expected) macro-level influences like adverse weather conditions. Nonetheless, the existence of these outliers emphasizes the complex, dynamic interplay of factors impacting price determination and reinforces the necessity of a nuanced approach to predictive modeling in this domain.

Figure ?? characterizes the monthly behavior of El Niño Southern Oscillation (**ENSO**) or SOI as prescribed before. A key predictor of global weather patterns, it bears considerable implications for weather-dependent industries, including agriculture and energy. Specifically, El Niño events, signalled by negative SOI values, often precipitate drier conditions and increased drought risks. The El Niño petiod of 2016, with the SOI at a striking -2.2, stands as a prime exemplar of these potent climatic shifts. During this time, particularly for the country, cross-industry effects had a lasting impact on the countries macroeconomic outlook as a whole Martınez et al. (2017); as well as in sector-specific measurements Bastianin, Lanza, and Manera (2018).

In the energy sector, the SOI's impact is acutely felt, particularly within hydroelectric and thermal power plants. Hydroelectric plants, reliant on water flow to generate electricity, can face severe operational constraints during El Niño events; drought conditions engendered by such episodes can cause drastic reductions in water flow, impairing the output of these power plants, thereby potentially eroding their profitability and market value. Thermal power plants, though not directly dependent on water flow for power generation, necessitate significant water volumes for cooling purposes. In a drought scenario, induced by an El Niño event, these plants can face operational challenges due to water scarcity, potentially leading to sub-optimal capacity operation, consequently affecting their operational efficiency and market value.

Paramount amongst these findings from our data is the pattern revealed between daily supply pricing (Figure 9) from hydro power plants and the quantity of electricity offered during these periods (Figure 10). A distinct escalation in pricing aligns with low kWh offerings, corresponding, as told, with periods during which the El Niño phenomenon asserts its influence. These stages, characterized by droughts and substantial reductions in water body levels, appear to directly implicate the supply and cost parameters in the power market; underscoring the critical interdependence between environmental phenomena and the energy sector. Specifically, for the 2015-2016 dramatic El Niño presence, a sizeable ascending shift in both bid and asking prices is shown; even for thermal plants (as they do hey do rely on large amounts of water for cooling processes).

Turning our attention to commercial demand, the distribution presents a behavior most closely aligned with a Gaussian process amongst all variables. This could potentially be indicative of the inherent trend of commercial power demand over the years. There is a relative consistency of demand patterns, amidst the volatility of other market variables. In addition we notice particular breaking point in the surge of demand, as coincidental with COVID-19 lockdown onset.

4.3 Data Split

In the pursuit of robust model evaluation, we adhere to the standard practice of chronological splitting in time-series analysis; we segregate our data to preserve the temporal sequence of observations. The training set spans an encompassing timeframe, beginning from the start of 2010 and extending to the end of 2021. This comprehensive 11-year stretch encapsulates various market conditions and key external events influencing energy market dynamics, such as El Niño occurrences; it's worth noting the deliberate inclusion of the year 2020, known for the onset of the global pandemic, amplifying the range of scenarios our model encounters during training.

Commencing from January 2022, our testing phase extends through December 2022, offering the model a robust one-year period of unseen data for evaluation. This period allows us to examine the model's ability to generate forecasts amidst variable and potentially unprecedented conditions. In this stringent testing environment, we truly assess the practical relevance of the models.

4.4 Window Data

Windowing is a technique commonly employed in time-series forecasting, used to structure input data to capture temporal dependencies and variations within defined periods. This is essential for domains such as ours of interest, where outputs at a specific time point are influenced by preceding observations Marcjasz, Serafin, and Weron (2018). Both the `create_windows` function performs this windowing task in a manner that allows for overlapping windows, and, in contrast; the `create_windows_no_overlap`, windows without overlap. This with the aim to lead to less redundant representation of the original data. While being computationally less intensive, and less prone to over-fitting. Our chosen s is 5 for both rolling and non-overlapping methods.

Table 5. Window data example for window size = s, for n variables including target.

Time	Variables				
	x_i	x_{i+1}	$\cdots$	x_{n-1}	y
$t-s$	34.56	76.89	$\cdots$	89.67	34.76
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
$t-1$	23.56	12.34	$\cdots$	34.54	67.89
t	56.76	23.45	$\cdots$	67.89	0

Here, Table 5 exemplifies a basic window, note that the overlapping would fundamentally roll predictions one timestamp forward. Whilst the non-overlapping would create an entirely new window. For the target variable at time t, 0 is masked in order to actually forecast.

5. Modeling, Results, & Anomalies

5.1 Statistical Models

For the purpose of measuring relative improvements for using deep learning models, baseline models were created. Baseline models play a critical role in setting a comparative standard to evaluate the effectiveness and robustness of more advanced models, such RNNs. These are typically simple, well-understood, and widely accepted models that are relatively quick and easy to implement. By comparing the performance of more complex models to these baselines, we can assess whether the added complexity provides a meaningful improvement in prediction accuracy.

The two baseline models in this project are the Auto-Regressive (**AR**) model and the Auto-Regressive Integrated Moving Average (**ARIMA**) model. Both models are employed due to their capacity to predict future points in a time series data based on past observations.

5.1.1 Auto-Regressive Baseline

The first model, AR(1), is a first-order autoregressive model. It forecasts the variable of interest using a linear combination of past values of the variable. The parameter used in this model, order=(1,0,0), indicates that the model is using one lagged value of the time series, with no differencing or moving average component. This is the most accurate naïve we aim to improve on:

$$y_t = \phi \cdot y_{t-1} + \epsilon_t$$

.

5.1.2 Auto-Regressive Integrated Moving Average

The second model, ARIMA(5,1,3), is a more advanced model that combines auto-regressive, differencing, and moving average components. The parameters used in this model, follow the usual ARIMA(p,d,q) set-up. This configuration is a consequence after considering multiple ARIMA(p,d,q) styles and returning the consistently strongest (in terms of predictive value) one across plant and timestamp subsets when validating with our test. This was acheived with the `autoARIMA` approximation available in *Python*; which hinges in the stepwise optimization found in Hyndman and Khandakar (2008).

Precisely, our best ARIMA-fitting model is considerably close to that one achieved by Gao, Lo, and Fan (2017) for the United Kingdom electricity market, with order (4,1,3). Similar conclusions are drawn in Arroyo, Nogales, and Conejo (2003) for the order of ARIMA day-ahead forecasting in first-stage power markets.

5.2 Neural Network Architectures

ARIMA approaches are predicated on the concept that a variable's future values are determined by its past values and the errors made in earlier predictions. Artificial Neural Networks (**ANN**) models, on the other hand, are based on the structure and function of the human brain and employ a network of interconnected processing nodes to learn from input and predict outcomes. When the link between input and output variables is complicated and nonlinear, ANNs may come in handy. This however, must be handled on a case-to-case basis; not as a general rule to approach price forecasting.

Comparatively, the RNN model is a deep learning model capable of capturing non-linear relationships and complex temporal dependencies in time series data. Unlike AR and ARIMA models that are constrained by a pre-specified number of lagged values, RNNs have an inherent ability to "remember" information from past time steps. This makes them more versatile and potentially more accurate for certain types of time series data. Nevertheless, RNNs also come with higher computational cost and complexity.

Moreover, as Preeti, Bala, and Singh (2019) and Kim et al. (2004) argue; financial and non-stationary time series forecasting can learn the non-linear and non-stationary nature in the context of RNNs, concretely for Long Short-Term Memory (**LSTM**) as they show. Unlike conventional statistical approaches, neural networks such as fuzzy logic, support vector machines, and neural networks do not require normal differencing or log transformations for stationarity.

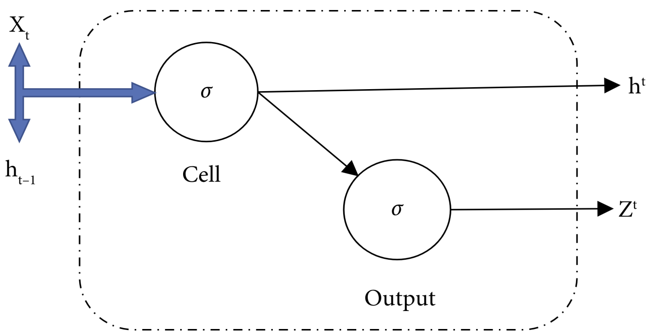

Figure 3. RNN Cell Unit Representation. From: Tomar et al. (2022)

5.2.1 Time-2-Vec

Time2Vec (**T2V**) represents an innovative algorithmic strategy; it leverages the power of vector representations to encode temporal data, thus enhancing the model's predictive capabilities. Serving as a layer within the Recurrent Neural Networks, T2V transforms the temporal information of the input data into a higher-dimensional vector space; each dimension could potentially encode different temporal aspects. This unique capability allows the model to retain both the ordinal and cyclical characteristics of time and to embed additional temporal dynamics; thereby, considerably bolstering the predictive efficacy of the Gated Recurrent Units (**GRUs**) and LSTM cells, especially in the realm of energy spot pricing prediction.

The T2V layer is made up of a linear term and a sum of weighted sine functions; the linear term models non-periodic components and aids in extrapolation, while the weighted sine functions capture the time series' periodic behavior. This is mathematically expressed as follows: Kazemi et al. (2019):

$$a(\tau, k)[i] = \theta_{i,0}(\omega_0\tau + \phi_0) + \sum_{j=1}^{k} \theta_{i,j} \sin(\omega_i\tau + \phi_i)$$

Here, $(a(\tau, k)[i]$ symbolizes the $i - th$ component of the output vector, τ is the input time value, and k stands for the number of sine functions utilized. During the training phase, the T2V layer optimizes the learned parameters—such as the number of sine functions and the weights assigned to each function—by minimizing a loss function. These learned parameters are pivotal in determining the dimensionality of the output vector and the importance of each frequency component. Upon optimizing these parameters, they can be employed to convert time values into vector representations, of temporal patterns; hoping to enhance the predictive capability.

5.2.2 *Gated Recurrent Units*

Gated Recurrent Units are powerful and versatile RNN components designed to retain long-term dependencies in sequence data. GRUs provide a robust learning mechanism that can capture and retain crucial temporal dynamics in the energy spot pricing time series. It helps model complex dependencies between time steps and learns relevant patterns over varying time periods, thereby potentially enhancing the prediction capability of the model. This, in conjunction with the T2V layer, assembles a potent combination for predictive modeling of time-series data. The learned parameters are instrumental in shaping the model's understanding of temporal trends, aiming to achieve superior predictive accuracy.

GRUs, proposed by Cho et al. (2014), address vanishing gradient problems that are often experienced in traditional RNNs by introducing a gating mechanism. This mechanism consists of two types of gates, the update gate and the reset gate. Controlling the gates determines how much past information is to be held for analysis. This allows GRUs to create a balance between past information (long-term dependencies) and new input data (short-term dependencies), which is particularly crucial for time series forecasting.

In this study the RNN, the GRU layer is parameterized with a variable number of units ranging from 64 to 128, with step size 32, as represented by `'hp.Int('rnn_units', 64, 128, step=32)'`. This value determines the dimensionality of the output space and, by extension, the capacity of the layer. The `'return_sequences'` parameter controls whether to return the last output in the output sequence or the full sequence; this is particularly important when stacking multiple recurrent layers. The activation function is set to `'tanh'`, which is a hyperbolic tangent activation function, useful for GRUs due to its properties of returning outputs between -1 and 1 and centering the data.

$$
\begin{aligned}
z(t) &= \sigma(W_z \cdot [h(t-1), x(t)]) \\
r(t) &= \sigma(W_r \cdot [h(t-1), x(t)]) \\
\tilde{h}(t) &= \tanh(W \cdot [r(t) \cdot h(t-1), x(t)]) \\
h(t) &= (1 - z(t)) \cdot h(t-1) + z(t) \cdot \tilde{h}(t)
\end{aligned}
$$

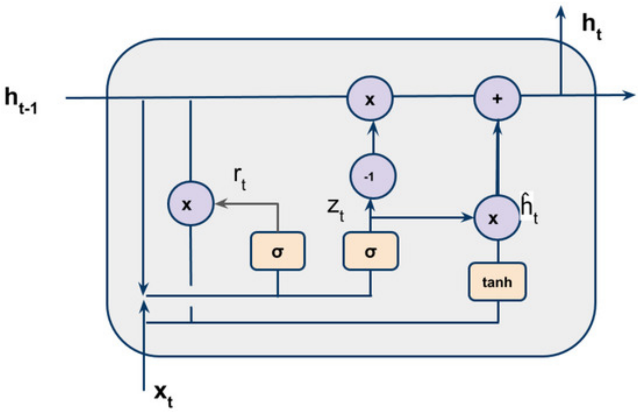

Figure 4. GRU Cell Unit Representation. From: Hosseini et al. (2020)

5.2.3 Long Short-Term Memory

Long Short-Term Memory units are a specialized type of Recurrent Neural Network architecture designed to manage long-term dependencies in sequence data. They are particularly effective in learning the inherent temporal dynamics of time-series data, which is crucial in fields like energy spot pricing prediction. The LSTM units have the unique ability to forget, remember, and carefully expose information, allowing them to model complex dependencies over varying periods, thereby potentially improving the predictive accuracy of the model.

In a general sense, they incorporate a gating mechanism, which includes three gates: the forget gate, the input gate, and the output gate. This mechanism helps control the extent of past information (long-term dependencies) to be remembered and the amount of new input data (short-term dependencies) to be stored in the cell state, thereby creating a balance, crucial for sequence data prediction Goodfellow, Bengio, and Courville (2016).

The LSTM layer in this study is parameterized with a variable number of units. This value determines the dimensionality of the output space and, by extension, the capacity of the layer. The `‘return_sequences‘` parameter controls whether to return the last output in the output sequence or the full sequence, a necessary configuration when stacking multiple recurrent layers. The LSTM units utilize `’tanh’` as the activation function in the cell state, in the same fashion as for GRU.

$$
\begin{aligned}
f(t) &= \sigma(W_f \cdot [h(t-1), x(t)] + b_f) \\
i(t) &= \sigma(W_i \cdot [h(t-1), x(t)] + b_i) \\
\tilde{C}(t) &= \tanh(W_C \cdot [h(t-1), x(t)] + b_C) \\
C(t) &= f(t) * C(t-1) + i(t) * \tilde{C}(t) \\
o(t) &= \sigma(W_o \cdot [h(t-1), x(t)] + b_o) \\
h(t) &= o(t) * \tanh(C(t))
\end{aligned}
$$

where σ denotes the sigmoid function, $h(t-1)$ is the previous hidden state, $x(t)$ is the current input, $f(t)$, $i(t)$, and $o(t)$ are the forget, input, and output gates respectively, $C(t-1)$ is the previous cell state, $C(t)$ is the current cell state, and $\tilde{C}(t)$ is the candidate cell state.

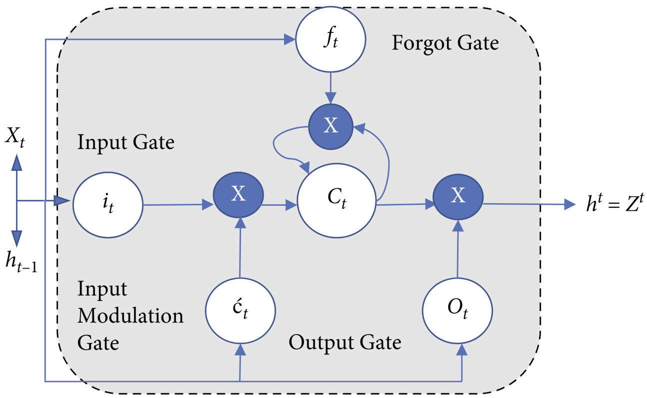

Figure 5. LSTM Cell Unit Representation. From: Tomar et al. (2022)

5.3 *Parameter Tuning*

Here we discuss our comparison approach for all presented statistical and machine learning prototypes. Specifically, we also extend the parametrization of our most relevant settings, them being RNNs. We extensively run through the code to clearly showcase our parameter tuning.

From scratch, we assign the implementation of a TT2V layer, with initialized trainable parameters `W`, `B`, and `P` during the build phase of the layer; these parameters' dimensions are derived from the input shape and a specified output dimension. When we forward propagate inputs through the layer during the call phase, we first produce a linear transformation of the input `x`, which is subjected to dimension expansion to ensure compatibility with the weight matrices. This transformation involves the weight matrix `W` and bias vector `B`, culminating in a resultant tensor, denoted as `original`. Subsequently, the sine function is applied to `original`, yielding a tensor `sin_trans`. As the final step of the transformation, we introduce the position encoding `P` to `sin_trans` and return the resultant tensor. This operation is intended to encode temporal patterns within the tensor representation, thereby enhancing the performance. Essentially, the T2V layer learns a time embedding in a manner analogous to the learning of word embeddings in natural language processing tasks. The model's capacity to learn the optimal values of `W`, `B`, and `P` stems from the process of backpropagation during training.

The `build_model` function facilitates the construction of a recurrent neural network (RNN) architecture with a Time2Vec (T2V) layer, the structure of which is dictated by hyperparameters obtained through the `hp` argument. Initially, an input tensor is defined, the shape of which corresponds to a six-window time-series and a number of features equivalent to `train_w[0].shape[1]-1`. Subsequently, a T2V layer is applied to the temporal feature in the input tensor, yielding a time-embedded vector `t2v` of fixed dimensionality (32 in this instance). This resultant vector is concatenated with the non-temporal input features, forming a hybrid tensor `x`.

The architecture of the RNN is then constructed through an iterative process, where the number of recurrent layers is determined by the hyperparameter `hp.Int('num_rnn_layers', 1, 2)`. Each layer in this architecture is either a GRU or a LSTM layer, as specified by the hyperparameter `hp.Choice('rnn_type', ['gru', 'lstm'])`. The number of units in each layer, alongside the activation function (tanh), is also governed by hyperparameters. Notably, all recurrent layers, except for the last one, return sequences to enable stacking of the layers.

Upon configuring the RNN, a densely connected layer is added to the architecture. The number of units in this layer, as well as the activation function (**ReLU**), are defined by hyperparameters. The final output of the model is produced by a dense layer with a single unit, which represents the prediction target.

Lastly, the entire architecture is compiled into a model with the Adam optimizer (of adaptive learning rates) and the mean squared error (**MSE**) loss function, alongside mean absolute error (**MAE**) as an additional metric. This function, in essence, provides a model that seamlessly integrates temporal embedding with recurrent layers, and the architecture is versatile enough to be tuned for both overlapping and non-overlapping windows.

Finally the Hyperband tuner, provided by O'Malley et al. (2019), is employed to optimize our model's hyperparameters. Hyperband operates on a bandit-based strategy, allowing for significant advantages over traditional grid search methods. As an adaptive resource allocation and early-stopping strategy, it balances the trade-off between exploration and exploitation in the hyperparameter search space. This approach enables us to efficiently navigate through this space, eliminating the exhaustive search often associated with grid search, which can be computationally expensive.

Our model's architecture, which includes multiple flexible hyperparameters, makes it particularly well-suited for Hyperband optimization. The model's performance is highly sensitive to the optimal configuration of these parameters; such as the choice between GRU or LSTM layers, the number of recurrent layers, and the number of units within each layer. Therefore, an adaptive strategy like Hyperband is an optimal choice for this model configuration.

At end, the provided code applies the Hyperband tuner to our model with the primary objective of minimizing the validation mean absolute error (**MAE**). Set to run for a maximum of five epochs, balancing computational efficiency.

Table 7 demonstrates our out-of-sample validation metrics for the defined test period of the last year. To achieve a streamlined computation of these, we designed custom functions that enable us to swiftly calculate the predictive power of our formulations with multiple metrics, here reported. In general, we see appreciably high values across the board. Our methods chosen account for high variance explanation of increasing manner as complexity also rises within the models.

An important caveat is highlighted here, as discerning from articles embodied in the literature. Due to the constraints presented by mean average percent error (**MAPE**) when dealing with negative errors ($\hat{y}_t < y_t$) than positive ones. To avoid said asymmetry we compute the 0-200 ranging of symmetric MAPE (**sMAPE**) as preferred by Makridakis and Hibon (2000) and Hyndman and Athanasopoulos (2018).

We were able to construct 2 search spaces for RNNs for overlapping and non-overlapping windows respectively. Table 6 presents the hyperparameters tuned in both scenarios.

Table 6. Best models for corresponding search spaces.

Model	Parameter	Description	Activation
No-overlap	GRU	2 Layers	tanh
	T2V	32 Unit	sine
	Optimizer	0.0001 Learning	-
	Dense	1 Layer	ReLU
Overlap	LSTM	2 Layers	tanh
	T2V	32 Unit	sine
	Optimizer	0.0001 Learning	-
	Dense	1 Layer	ReLU

5.4 Results & Anomalies

Table 7. Test metrics for each model.

Model	RMSE	MAE	sMAPE(0-200)	R-squared
Overlap RNN	0.1327	0.0854	11.8221%	0.9316
Non-overlap RNN	0.1361	0.0865	12.0120%	0.9281
ARIMA(5,1,3)	0.2065	0.0590	19.6500%	0.9205
AR(1)	0.2175	0.0619	29.2700%	0.9118

In our analysis of the out-of-sample results from day-ahead supply price forecasting, the performance of various models presents intriguing variations across four key performance metrics: RMSE, MAE, sMAPE, and R-squared.

1. RMSE: We observe that the RMSE, being sensitive to outliers, magnifies large errors due to the squaring operation involved in its calculation; hence, the RNN model may exhibit a higher RMSE if a few high-magnitude prediction errors are present.
2. MAE: Contrastingly, the MAE metric, calculated as the mean of absolute differences between actual and predicted values, does not disproportionately penalize larger errors; this could potentially elucidate why the RNN models demonstrate a higher MAE if their average performance is inferior.
3. sMAPE: Regarding the sMAPE metric, which represents the average absolute percent difference between observed and predicted values, we find that the RNN models, due to their superior percentage-wise prediction accuracy, tend to score lower.
4. R-squared: Lastly, the R-squared metric, which quantifies the proportion of variance in the dependent variable that is predictable from the independent variables, tends to be higher for the RNN models; this suggests that they excel in encapsulating the data's variation compared to the ARIMA and AR models.

These variations in the models' performance can be attributed to the unique properties of the error distribution and the models' capacity to discern patterns in the data. On the one hand, the RNN model may proficiently predict the overall trend, thereby achieving a superior R-squared and a lower sMAPE. Yet, it may falter in accurately predicting precise values, resulting in a higher RMSE and MAE. On the other hand, the ARIMA and AR models may consistently predict specific points, leading to a lower RMSE and MAE, even if they do not capture the overall trend as effectively.

Furthermore, we speculate that the presence of outliers or extreme values in the dataset may contribute to these disparities. The RNN model's higher sensitivity to outliers could result in inflated

MAE and RMSE scores. These statements are congruent to what Figure 11 presents, where extreme values for NN approaches are less prone to predict outliers correctly.

6. Conclusions, Limitations & Extensions

This study extends the work of Riascos, Chitiva, and Salazar (2022) and other related literature by employing modern data science techniques to examine the pricing behavior within the Colombian energy market. The market, liberalized in the 1990s and further refined in its formation in 2010, has intrinsic characteristics such as inelastic demand, market concentration, and high storage costs. These attributes pose challenges to achieving competitiveness and efficiency, the original aims of liberalization. In this context, the capacity to predict spot pricing can provide a mechanism for policymakers and stakeholders to understand and potentially rectify market anomalies.

The methodology adopted in this investigation is two-pronged. Firstly, key market participants are identified using market concentration metrics such as the Residual Supply Index (RSI), based on the disclosed availability of these agents, which could indicate significant power in pricing in the market. Secondly, a historical dataset comprising various factors, including past daily offers, supply, demand, and hourly prices paid, is used by machine learning techniques to predict the energy offers of all agents. These predicted offers combined with the market concentration metrics could signal possible imperfect competitive practices and could be used to trigger alerts.

The study deployed 4 different predictive models: a Gated Recurrent Unit, coupled with Time2Vec; a Long Short-Term Memory unit based RNN, with T2V layering; an ARIMA(5,1,3) model; and a naïve AR(1) model. Their performance was assessed based on four different metrics: sMAPE, RMSE, MAE, and R-squared. The GRU + T2V model displayed superior performance in capturing the variance in the data as indicated by its highest R-squared value, but it was more sensitive to large errors, as suggested by its higher RMSE. The ARIMA model showed a lower average percentage and absolute error, but its explanatory power for the variance in the data was less robust. The Naive AR(1) model exhibited the highest average percentage error and the weakest explanatory power for data variance. While some metrics proved challenging to interpret from model to model; overall goodness-of-fit is indeed better for the deep learning architectures.

These diverse results underscore the intricacies inherent in model selection, highlighting the trade-offs involved, and emphasizing the need to align the choice of metrics with the specific objectives of the investigation. As a function of different applications and research contexts, the prioritization of different aspects of model performance may vary, warranting careful consideration in interpreting comparative results.

The findings of this study offer valuable insights into potential market imperfections within the Colombian energy market and provide a robust methodology for similar investigations in other liberalized and regulated markets. As lastly explained, limitations of these models are nowadays constrained at most by hardware and time limitations; trading accuracy for effectiveness.

In conclusion, we find it commonplace to encounter discrepancies across these metrics. Each metric underscores different facets of the errors, and it is plausible for a model to excel in one but underperform in another, contingent upon the characteristics of the data and the nature of the model's errors. While this may be the case in theory, studies such as Siami-Namini, Muhammad, and Fahimullah (2018) reveal better metrics across the board when comparing LSTM and ARIMA; which indicates the need to proceed cautiously with each dataset.

Nonetheless, this methodology proved useful at shedding light on perceived marker deviations from competitiveness. Further evaluation is recommended to strike a robust balance between using market given variables (as the ones used for prediction here) and market concentration indicators; which may in turn increase predicting power while veering off of the theoretical (and regulatory) framework we have started from.

References

Arroyo, Jose M, Francisco J Nogales, and Antonio J Conejo. 2003. Day-ahead electricity price forecasting using the wavelet transform and arima models. *IEEE Transactions on Power Systems* 18 (3): 1014–1020.

Bastianin, Andrea, Alessandro Lanza, and Matteo Manera. 2018. Economic impacts of el niño southern oscillation: evidence from the colombian coffee market. *Agricultural economics* 49 (5): 623–633.

Cho, Kyunghyun, Bart van Merrienboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. 2014. *Learning phrase representations using rnn encoder-decoder for statistical machine translation.* arXiv: 1406.1078 `[cs.CL]`.

Ciarreta, Aitor, Shahriyar Nasirov, and Carlos Silva. 2016. The development of market power in the Spanish power generation sector: Perspectives after market liberalization. *Energy Policy* 96:700–710. ISSN: 0301-4215. https://doi.org/https://doi.org/10.1016/j.enpol.2016.06.029. https://www.sciencedirect.com/science/article/pii/S0301421516303226.

Comisión de Regulación de Energía y Gas. 2021. *Modernización del Mercado de Energía Mayorista: Despacho vinculante, mercados intradiarios y servicios complementarios.* Technical report. Documento CREG-114. Bogotá D.C., Colombia, September.

Correa-Giraldo, Manuel, John J. Garcia-Rendon, and Alex Perez. 2021. Strategic behaviors and transfer of wholesale costs to retail prices in the electricity market: evidence from colombia. *Energy Economics* 99:105276. ISSN: 0140-9883. https://doi.org/https://doi.org/10.1016/j.eneco.2021.105276. https://www.sciencedirect.com/science/article/pii/S014098832100181X.

Esmaeili Aliabadi, Danial, and Katrina Chan. 2022. The emerging threat of Artificial Intelligence on competition in liberalized electricity markets: a deep Q-network approach. *Applied Energy* 325:119813. ISSN: 0306-2619. https://doi.org/https://doi.org/10.1016/j.apenergy.2022.119813. https://www.sciencedirect.com/science/article/pii/S030626192201087X.

Fabra, Natalia, and Juan Toro. 2005. Price wars and collusion in the Spanish electricity market. *International Journal of Industrial Organization* 23 (3): 155–181. ISSN: 0167-7187. https://doi.org/https://doi.org/10.1016/j.ijindorg.2005.01.004. https://www.sciencedirect.com/science/article/pii/S0167718705000123.

Fridolfsson, Sven-Olof, and Thomas P. Tangerås. 2009. Market power in the Nordic electricity wholesale market: A survey of the empirical evidence. New Zealand Energy Strategy, *Energy Policy* 37 (9): 3681–3692. ISSN: 0301-4215. https://doi.org/https://doi.org/10.1016/j.enpol.2009.04.036. https://www.sciencedirect.com/science/article/pii/S0301421509002833.

Gao, Gao, Kwoklun Lo, and Fulin Fan. 2017. Comparison of arima and ann models used in electricity price forecasting for power market. *Energy and Power Engineering* 09 (January): 120–126. https://doi.org/10.4236/epe.2017.94B015.

Garcia Pires, Armando J., and Frode Skjeret. 2023. Screening for partial collusion in retail electricity markets. *Energy Economics* 117:106473. ISSN: 0140-9883. https://doi.org/https://doi.org/10.1016/j.eneco.2022.106473. https://www.sciencedirect.com/science/article/pii/S0140988322006028.

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. 2016. *Deep learning.* Http://www.deeplearningbook.org. MIT Press.

Hansen, Jean-Pierre. 2011. *Market Power: Microeconomics and Electricity Sector.* https://cgemp.dauphine.fr/fileadmin/mediatheque/centres/cgemp/presentations/JP%5C%20Hansen.pdf.

Haselip, James. 2007. Optimising reform: the sustainability of electricity market liberalisation in less developed countries. Optimising Reform: The Sustainability of Electricity Market Liberalisation in Less Developed Countries, *Journal of Cleaner Production* 15 (2): 109–111. ISSN: 0959-6526. https://doi.org/https://doi.org/10.1016/j.jclepro.2006.02.012. https://www.sciencedirect.com/science/article/pii/S0959652606001247.

Hosseini, Majid, Satya Katragadda, Jessica Wojtkiewicz, Raju Gottumukkala, Anthony Maida, and Terrence Lynn Chambers. 2020. Direct normal irradiance forecasting using multivariate gated recurrent units. *Energies* 13 (15). ISSN: 1996-1073. https://doi.org/10.3390/en13153914. https://www.mdpi.com/1996-1073/13/15/3914.

Hyndman, Rob J., and Yeasmin Khandakar. 2008. Automatic Time Series Forecasting: The forecast Package for R. *Journal of Statistical Software* 27 (3): 1–22. https://doi.org/10.18637/jss.v027.i03. https://www.jstatsoft.org/index.php/jss/article/view/v027i03.

Hyndman, Robin John, and George Athanasopoulos. 2018. *Forecasting: principles and practice* [in English]. 2nd. Australia: OTexts.

Kazemi, Seyed Mehran, Rishab Goel, Sepehr Eghbali, Janahan Ramanan, Jaspreet Sahota, Sanjay Thakur, Stella Wu, Cathal Smyth, Pascal Poupart, and Marcus A. Brubaker. 2019. Time2vec: learning a vector representation of time. *CoRR* abs/1907.05321. arXiv: 1907.05321. http://arxiv.org/abs/1907.05321.

Kim, Tae, Kyong Joo Oh, Chiho Kim, and do jong-doo do. 2004. Artificial neural networks for non-stationary time series. *Neurocomputing* 61 (October): 439–447. https://doi.org/10.1016/j.neucom.2004.04.002.

Lago, Jesus, Fjo De Ridder, and Bart De Schutter. 2018. Forecasting spot electricity prices: deep learning approaches and empirical comparison of traditional algorithms. *Applied Energy* 221:386–405. ISSN: 0306-2619. https://doi.org/https://doi.org/10.1016/j.apenergy.2018.02.069. https://www.sciencedirect.com/science/article/pii/S030626191830196X.

Lise, Wietze, Benjamin F. Hobbs, and Sebastiaan Hers. 2008. Market power in the European electricity market: The impacts of dry weather and additional transmission capacity. *Energy Policy* 36 (4): 1331–1343. ISSN: 0301-4215. https://doi.org/https://doi.org/10.1016/j.enpol.2007.12.012. https://www.sciencedirect.com/science/article/pii/S0301421507005526.

Lundin, Erik, and Thomas P. Tangerås. 2020. Cournot competition in wholesale electricity markets: The Nordic power exchange, Nord Pool. *International Journal of Industrial Organization* 68:102536. ISSN: 0167-7187. https://doi.org/https://doi.org/10.1016/j.ijindorg.2019.102536. https://www.sciencedirect.com/science/article/pii/S0167718719300645.

Makridakis, Spyros, and Michèle Hibon. 2000. The m3-competition: results, conclusions and implications. The M3- Competition, *International Journal of Forecasting* 16 (4): 451–476. ISSN: 0169-2070. https://doi.org/https://doi.org/10.1016/S0169-2070(00)00057-1. https://www.sciencedirect.com/science/article/pii/S0169207000000571.

Marcjasz, Grzegorz, Tomasz Serafin, and Rafał Weron. 2018. Selection of calibration windows for day-ahead electricity price forecasting. *Energies* 11 (9). ISSN: 1996-1073. https://doi.org/10.3390/en11092364. https://www.mdpi.com/1996-1073/11/9/2364.

Marin, Jorge Barrientos, Elkin Tabares Orozco, and Esteban Velilla. 2018. *International Journal of Energy Economics and Policy* 8, no. 3 (May): 97–106. https://www.econjournals.com/index.php/ijeep/article/view/6181.

Martınez, Rodney, Eduardo Zambrano, Juan José Nieto, Julián Hernández, and Felipe Costa. 2017. Evolución, vulnerabilidad e impactos económicos y sociales de El Niño 2015-2016 en América Latina. *Investigaciones Geográficas (España),* no. 68, 65–78.

Mastropietro, Paolo, Pablo Rodilla, Lina Escobar Rangel, and Carlos Batlle. 2020. Reforming the colombian electricity market for an efficient integration of renewables: a proposal. *Energy Policy* 139:111346. ISSN: 0301-4215. https://doi.org/https://doi.org/10.1016/j.enpol.2020.111346. https://www.sciencedirect.com/science/article/pii/S0301421520301038.

Mulder, Machiel, and Lambert Schoonbeek. 2013. Decomposing changes in competition in the Dutch electricity market through the Residual Supply Index. *Energy Economics* 39:100–107. ISSN: 0140-9883. https://doi.org/https://doi.org/10.1016/j.eneco.2013.04.006. https://www.sciencedirect.com/science/article/pii/S0140988313000704.

Nagayama, Hiroaki. 2009. Electric power sector reform liberalization models and electric power prices in developing countries: an empirical analysis using international panel data. *Energy Economics* 31 (3): 463–472. ISSN: 0140-9883. https://doi.org/https://doi.org/10.1016/j.eneco.2008.12.004. https://www.sciencedirect.com/science/article/pii/S0140988308001916.

Newbery, David M. 2008. *Predicting market power in wholesale electricity markets.* Cambridge Working Papers in Economics. Faculty of Economics, University of Cambridge. https://EconPapers.repec.org/RePEc:cam:camdae:0837.

O'Malley, Tom, Elie Bursztein, James Long, François Chollet, Haifeng Jin, Luca Invernizzi, et al. 2019. *Kerastuner.* https://github.com/keras-team/keras-tuner.

Poggi, Aurora, Luca Di Persio, and Matthias Ehrhardt. 2023. Electricity price forecasting via statistical and deep learning approaches: The German case. *AppliedMath* 3 (2): 316–342. ISSN: 2673-9909. https://doi.org/10.3390/appliedmath3020018. https://www.mdpi.com/2673-9909/3/2/18.

Prabhakar Karthikeyan, S., I. Jacob Raglend, and D.P. Kothari. 2013. A review on market power in deregulated electricity market. *International Journal of Electrical Power Energy Systems* 48:139–147. ISSN: 0142-0615. https://doi.org/https://doi.org/10.1016/j.ijepes.2012.11.024. https://www.sciencedirect.com/science/article/pii/S0142061512006771.

Preeti, Rajni Bala, and Ram Pal Singh. 2019. Financial and non-stationary time series forecasting using lstm recurrent neural network for short and long horizon. In *2019 10th international conference on computing, communication and networking technologies (icccnt),* 1–7. https://doi.org/10.1109/ICCCNT45670.2019.8944624.

Razmi, Peyman, Majid Oloomi Buygi, and Mohammad Esmalifalak. 2021. A Machine Learning Approach for Collusion Detection in Electricity Markets Based on Nash Equilibrium Theory. *Journal of Modern Power Systems and Clean Energy* 9, no. 1 (January): 170–180. ISSN: 2196-5420. https://doi.org/10.35833/MPCE.2018.000566.

Riascos, Álvaro, Julián Chitiva, and Carlos Salazar. 2022. *Detección de anomalías y poder de mercado en el sector eléctrico colombiano.* Technical report. November. https://doi.org/https://doi.org/10.32468/be.1217. https://repositorio.banrep.gov.co/handle/20.500.12134/10514.

Sheffrin, Anjali. 2002. Predicting Market Power Using the Residual Supply Index. In *Ferc market monitoring workshop.* California Independent System Operator.

Siami-Namini, Sima, Danish Muhammad, and Fahimullah Fahimullah. 2018. Comparison of ARIMA and LSTM in Forecasting Time Series. *Applied Economics and Finance* 5 (5): 23–32.

Suarez, Carlos. 2022. Private management and strategic bidding behavior in electricity markets: Evidence from Colombia. *Energy Economics* 111:106058. ISSN: 0140-9883. https://doi.org/https://doi.org/10.1016/j.eneco.2022.106058. https://www.sciencedirect.com/science/article/pii/S0140988322002250.

Tomar, Dimpal, Pradeep Tomar, Arpit Bhardwaj, G. R. Sinha, and Vijay Kumar. 2022. Deep learning neural network prediction system enhanced with best window size in sliding window algorithm for predicting domestic power consumption in a residential building. *Intell. Neuroscience* (London, GBR) 2022 (January). ISSN: 1687-5265. https://doi.org/10.1155/2022/7216959. https://doi.org/10.1155/2022/7216959.

Vega-Márquez, Belén, Cristina Rubio-Escudero, Isabel A. Nepomuceno-Chamorro, and Ángel Arcos-Vargas. 2021. Use of Deep Learning Architectures for Day-Ahead Electricity Price Forecasting over Different Time Periods in the Spanish Electricity Market. *Applied Sciences* 11 (13). ISSN: 2076-3417. https://doi.org/10.3390/app11136097. https://www.mdpi.com/2076-3417/11/13/6097.

Weinmann, Jens. 2007. The sustainability of electricity sector liberalization in Latin America. Optimising Reform: The Sustainability of Electricity Market Liberalisation in Less Developed Countries, *Journal of Cleaner Production* 15 (2): 112–114. ISSN: 0959-6526. https://doi.org/https://doi.org/10.1016/j.jclepro.2006.02.011. https://www.sciencedirect.com/science/article/pii/S0959652606001235.

Wolak, Frank A. 2003. Measuring unilateral market power in wholesale electricity markets: the california market, 1998-2000. *American Economic Review* 93, no. 2 (May): 425–430. https://doi.org/10.1257/000282803321947461. https://www.aeaweb.org/articles?id=10.1257/000282803321947461.

Appendix 1.
Design of the Colombian energy market

The current regulatory framework for Colombia's electricity sector is outlined jointly between 3 pivotal resolutions put forward during the last 4 decades (Comisión de Regulación de Energía y Gas 2021); CREG-024, CREG-025 from 1995, and CREG-051 from 2009.

As a product of the electric rationing during 1992 and 1993 in the country; the resolutions introduced in the 90s configure both the operational and commercial regulation, respectively. The latter, conceives the *Mercado de Energía Mayorista - Wholesale Energy Market*: **MEM**.

A competitive market split into two segments, short-term and long-term markets: the producers and distributors of electrical energy conduct transactions in the short-term market to meet demand in real time; a spot rate is set in this market based on the supply and demand for electric energy for each day. On the other hand, in the long-term energy market, electricity producers and sellers enter into long-term contracts for the supply of electricity. These contracts, which are negotiated with several years' notice, enable the producers and marketers to guarantee the sale and purchase of electrical energy at stable prices.

The latest resolution mentioned institutes the current price formation in the MEM, done via a 2 part valuation: variable costs and start-stop cost (start and stop of power generation activity from the power plant). This aims to ensure fair pricing for consumers. A single power plant offers its available power supply for all hours of any given day at a specific price point[5]. The MEM is regulated and operates through auctions, in which the energy generators' agents offer prices and declare their availability. The cost is determined by market force interaction. The agents or conglomerate owners have the option of adjusting their offerings or imitating other companies within the market. These key characteristics are monitored for antitrust purposes to preserve fair competition.

At heart, this aims to guarantee that the framework of operation for the short-term market return the minimum cost for electricity consumers in a step-by-step system:

I. Price offerings and declaration of availability: Agents inform daily price offerings and hourly available supply by plant/unit.

II. Price settlement: With the usage of ideal generation output, the spot price for electricity is set by organizing price offerings in an ascending manner, ensuring commercial demand and unforeseen demand events are met at the lowest pricing; this is referred to as *Optimal Economic Dispatch*.

III. Price Reconciliation: Discrepancies between actual generation and ideal generation output or dispatch are resolved for each agent, where deviations are evaluated and positive compensations are determined when actual generation exceeds ideal generation output.

The 2-part system comes in play through the figure of *Precio de Reconciliación Positiva de los Generadores Térmicos e Hidraúlicos - Positive Reconciliation Price for Thermal and Hydroelectric Generators*: **PR**.

$$PR = \min[(CSC + CTC + COM + OCV) + \frac{PCAP}{GSA};\ Supply\ Price + \frac{Par}{GSA}]$$

Where each component is in part calculated or set in the following manner[6]:

- *CSC*: Fuel Supply Cost.
- *CTC*: Costo de Transporte del Combustible - Fuel Transportation Cost.
- *COM*: Costo de Operación y Mantenimiento - O&M Cost.

[5]Before this, resolution CREG-025 indicated hourly rates, essentially 24 price points in a day.

[6]Calculations for these costs can be found in resolutions CREG 034/2001, 071/2006, and 038/2001.

- OCV: Otros Costos Variables - Other Variable Costs.
- GSA: Generación de Seguridad de Arranque - Total MW generated for safety outside of ideal generation output during the day.
- Par: Precio de arranque-parada ideal - Ideal start-stop price for each type of fuel used in power generation.
- $PCAP$: Costo de arranque-parada - Start-Stop Cost offered by the agent.

Where the last 2 components are 0 if the generation from the power plant is continuing from days prior. Besides; the PCAP, or the initial price offering the agent the is defined as:

$$PCAP_t = PCAP_0 \cdot \frac{PPI_{m-1}}{PPI_0}$$

Where:

- $PCAP_0$: Initial $PCAP$ of the agent.
- PPI_{m-1}: Producer Price Index on capital goods for the United States of America at the previous month.
- PPI Current Producer Price Index for the same indicator.

These formulations guide financial considerations for the agents and are considered inputs for price offerings through expectations of these players. The price formation within the short-term market is shown below.

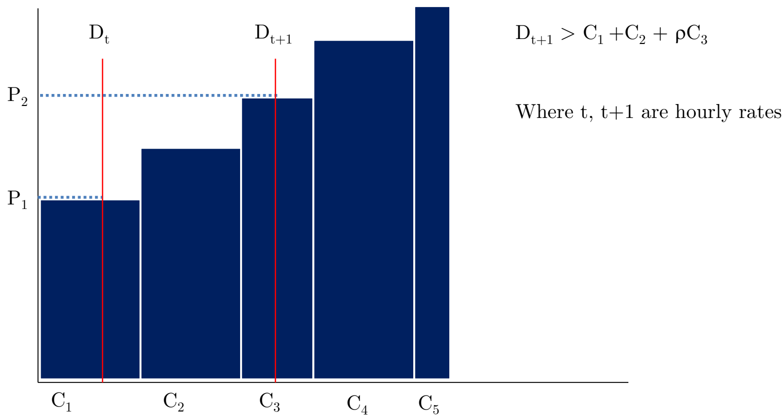

Figure 6. Price Formation - Optimal Economic Dispatch. Source: Authors' elaboration.

Appendix 2.
Code

All relevant code for this research is found in this public **GitHub** repository.

Appendix 3.
Figures

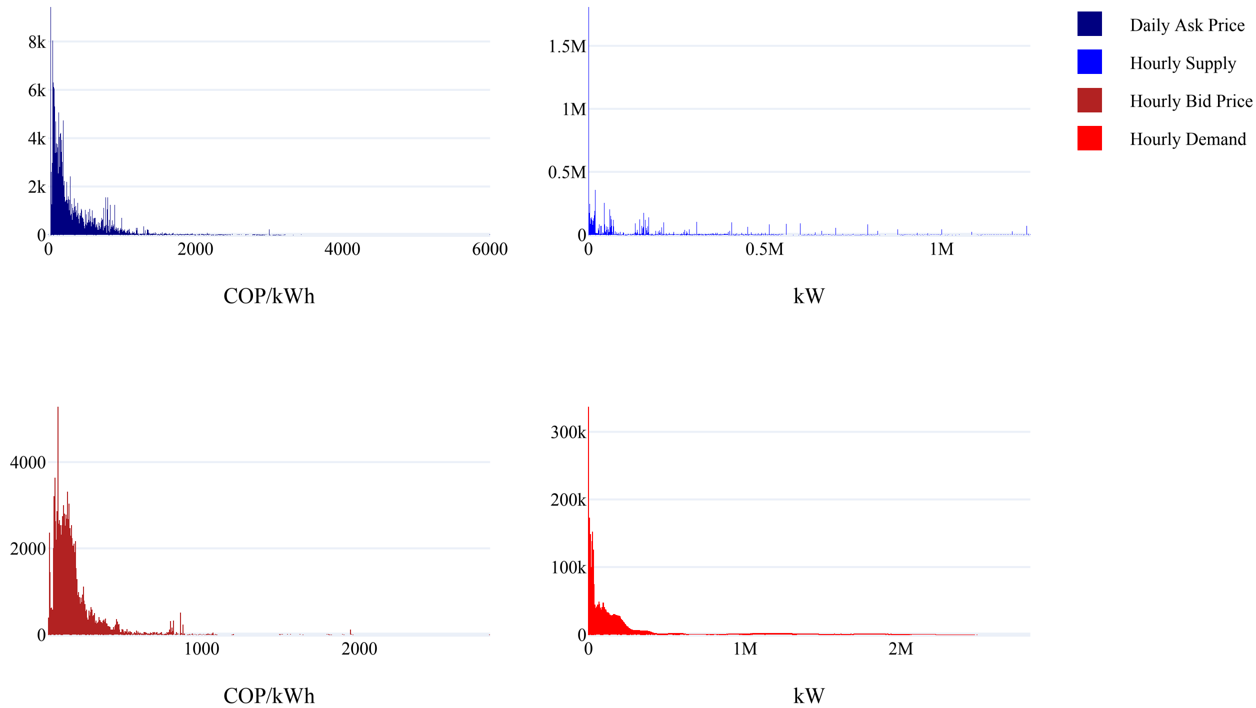

Figure 7. Distribution of defined variables. Source: XM. Authors' calculations.

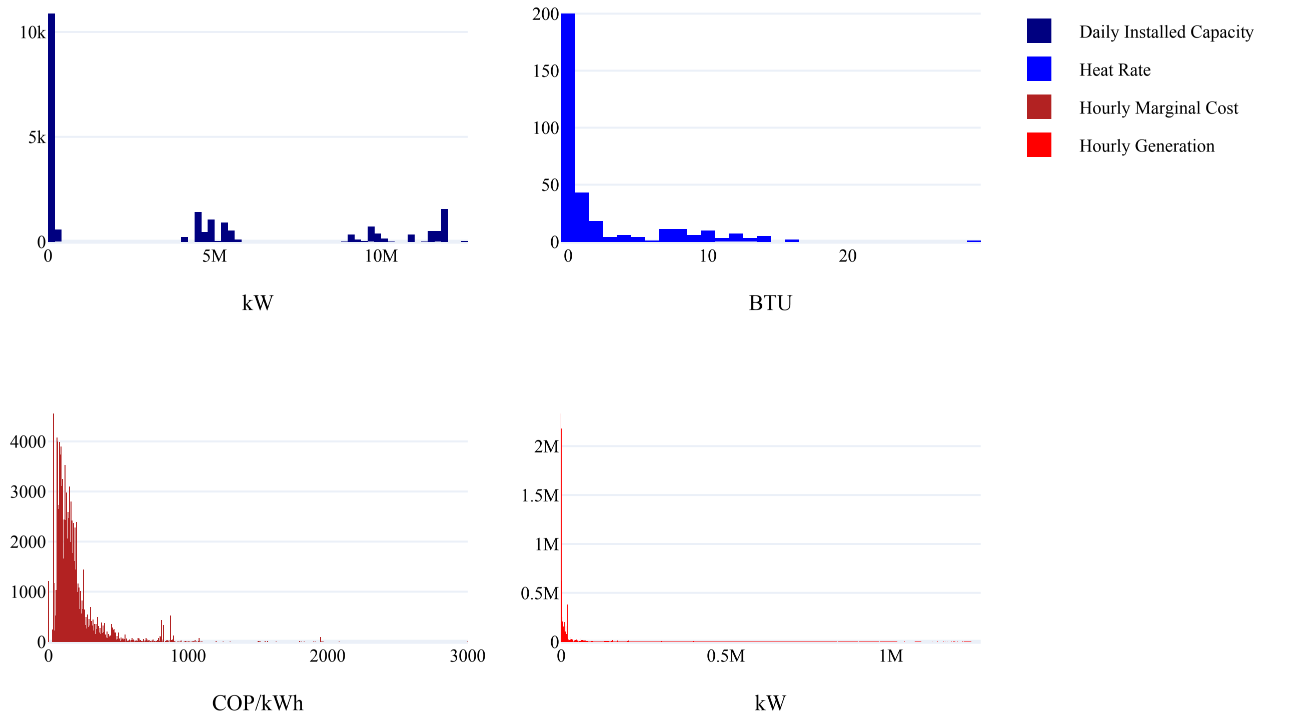

Figure 8. Distribution of defined variables. Source: XM. Authors' calculations.

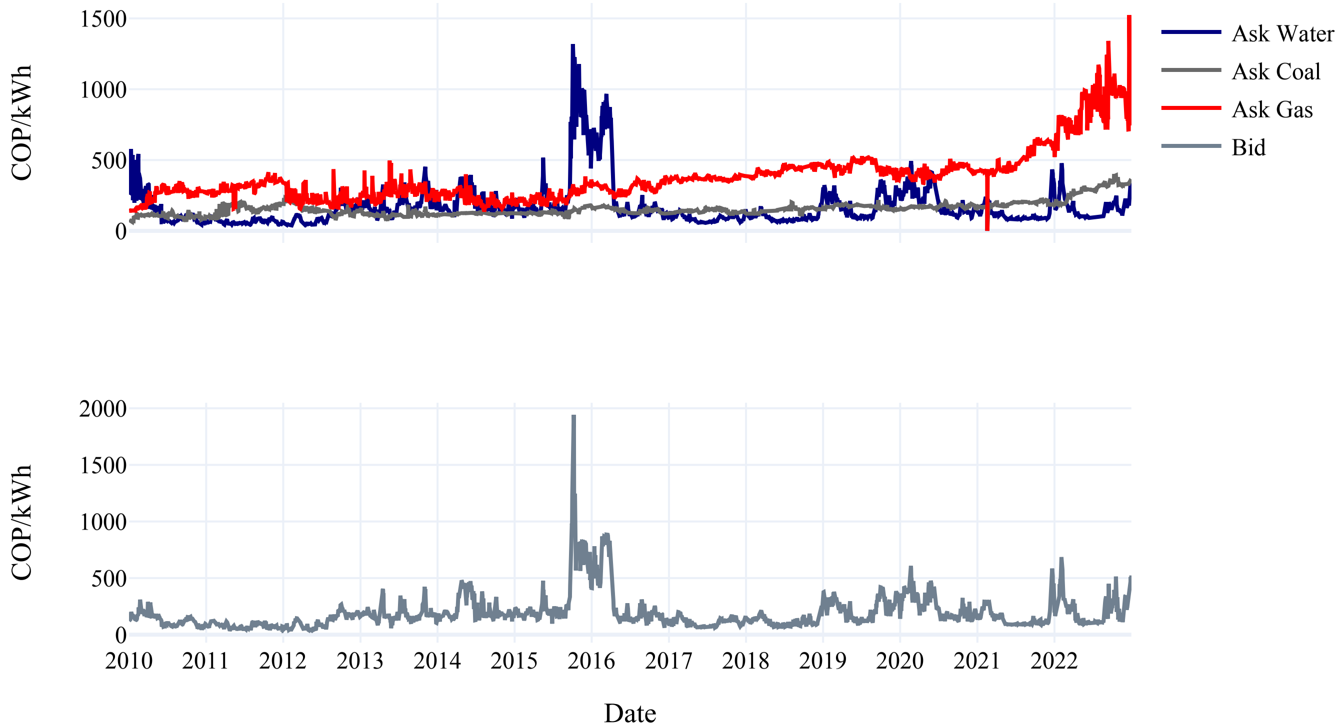

Figure 9. Weighted average prices - supply price by resource and demand price. Source: XM. Authors' elaboration.

Figure 10. Daily supply by resource and demand. Source: XM. Authors' elaboration.

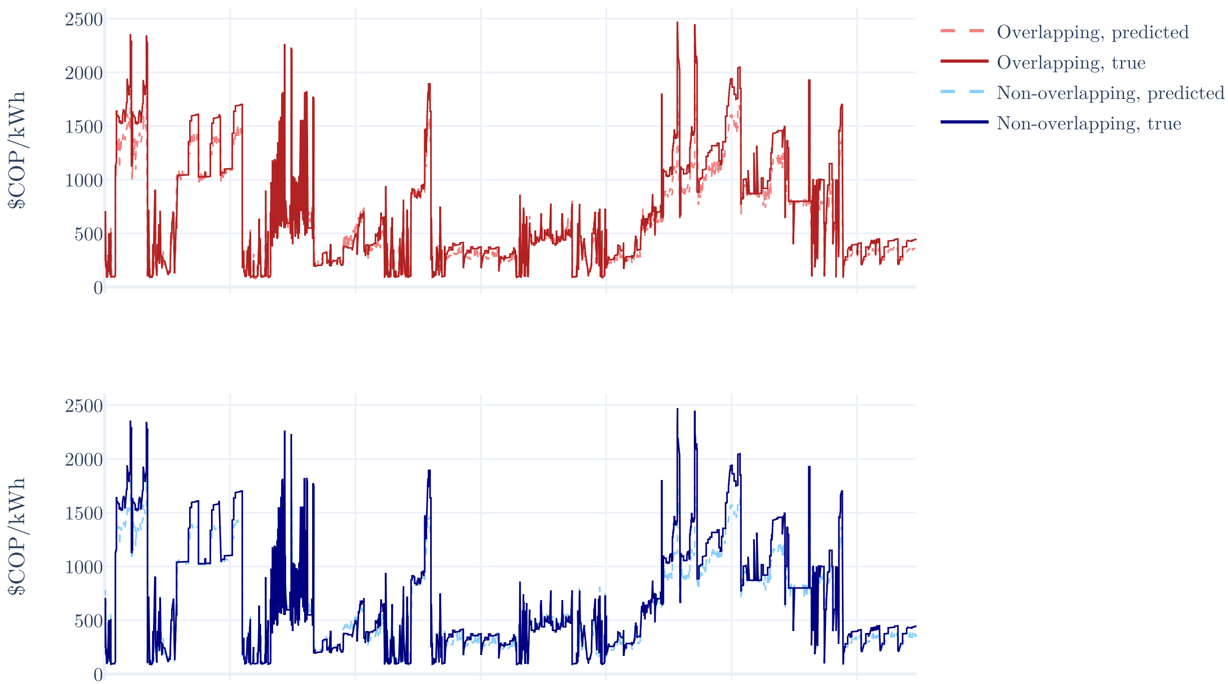

Figure 11. Overlapping and Non-overlapping windows predictions with s=5. Out-of-sample results, 2022 hourly data. Source: Authors' elaboration.