




## Article

# An Intra-Day Electricity Price Forecasting Based on a Probabilistic Transformer Neural Network Architecture

Sergio Cantillo-Luna <sup>1</sup>, Ricardo Moreno-Chuquen <sup>2,\*</sup>, Jesus Lopez-Sotelo <sup>1</sup> and David Celeita <sup>3</sup>

<sup>1</sup> Faculty of Engineering, Universidad Autónoma de Occidente, Cali 760030, Colombia; sergio.cantillo@uao.edu.co (S.C.-L.); jalopez@uao.edu.co (J.L.-S.)

<sup>2</sup> Faculty of Engineering and Design, Universidad Icesi, Cali 760031, Colombia

<sup>3</sup> School of Engineering, Science and Technology, Universidad del Rosario, Bogotá 111221, Colombia; david.celeita@urosario.edu.co

\* Correspondence: rmoreno@icesi.edu.co

**Abstract:** This paper describes the development of a deep neural network architecture based on transformer encoder blocks and Time2Vec layers for the prediction of electricity prices several steps ahead (8 h), from a probabilistic approach, to feed future decision-making tools in the context of the widespread use of intra-day DERs and new market perspectives. The proposed model was tested with hourly wholesale electricity price data from Colombia, and the results were compared with different state-of-the-art forecasting baseline-tuned models such as Holt–Winters, XGBoost, Stacked LSTM, and Attention-LSTM. The findings show that the proposed model outperforms these baselines by effectively incorporating nonlinearity and explicitly modeling the underlying data’s behavior, all of this under four operating scenarios and different performance metrics. This allows it to handle high-, medium-, and low-variability scenarios while maintaining the accuracy and reliability of its predictions. The proposed framework shows potential for significantly improving the accuracy of electricity price forecasts, which can have significant benefits for making informed decisions in the energy sector.

**Keywords:** decision making; deep learning; electricity price forecasting (EPF); probabilistic forecasting; time series forecasting



**Citation:** Cantillo-Luna, S.; Moreno-Chuquen, R.; Lopez-Sotelo, J.; Celeita, D. An Intra-Day Electricity Price Forecasting Based on a Probabilistic Transformer Neural Network Architecture. *Energies* **2023**, *16*, 6767. <https://doi.org/10.3390/en16196767>

Academic Editors: Tiago Pinto and Lin Wang

Received: 4 August 2023

Revised: 11 September 2023

Accepted: 15 September 2023

Published: 22 September 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The integration of distributed energy resources (DERs) into modern power systems is becoming increasingly important as new multi-resource electricity transactions emerge [1]. DERs, including residential photovoltaic (PV) systems, demand response (DR), and energy storage systems (ESSs), have the potential to provide numerous technical and environmental advantages [2]. However, there are still a number of transactional and reliability challenges associated with integrating distributed energy resources (DERs) into existing power grids [3]. The inclusion of these energy assets introduces additional complexity in forecasting future electricity prices, as traditional models may no longer be sufficient [4]. Due to their intermittent nature and variability, PV-powered DERs can significantly impact the supply and demand dynamics of the power system, which increases uncertainty in price forecasting.

In this context, accurate electricity price forecasting has an important role in decision making by various stakeholders and the development of innovative business and market models towards a future power grid [5]. Market players such as generators, retailers, and consumers require accurate price forecasting to optimize their operations, manage their energy portfolios, and make informed decisions about energy transactions according to their interests [6]. This also enables innovative business models, such as peer-to-peer energy trading, grid balancing services, and energy aggregation platforms, along with DER

integration. These developments foster a dynamic and decentralized energy landscape, offering opportunities to enhance efficiency and sustainability in the energy sector [7].

Consequently, short-term power price forecasting has gained significant importance and development in recent times [8–10]. The emergence of new market approaches, such as intra-day periods, has also increased the need for accurate forecasts. This allows market participants to adjust their strategies and optimize their energy trading activities in shorter time frames. Such predictions are essential for market participants to effectively manage their assets, hedge risks, and take advantage of opportunities presented by dynamic market conditions [11].

To address these challenges and take advantage of the aforementioned opportunities, the present paper aims to predict intra-day electricity prices using a tuned multi-step forecasting model called Time2Vec Transformer (T2V-TE). This model incorporates a combination of stacked transformer encoders and a special time-varying embedding called Time2Vec. To evaluate the performance of the model, (both point and probabilistic approaches), different comparison baseline forecasting models were considered, including Holt–Winters, XGBoost, and LSTM-based models, to analyze historical hourly electricity prices in the Colombian wholesale market. This analysis was conducted in the context of the increasing integration of DERs and the exploration of new market insights.

### 1.1. Literature Review

The electricity price is a crucial factor in energy markets and current power grids, playing a pivotal role in providing a reliable and economically efficient power supply [1,12,13]. Therefore, precise electricity price forecasting is essential for all stakeholders, as it empowers them to make informed decisions that increase profitability and reduce risks in competitive electricity markets. In addition, it enhances the overall stability and optimal operation of the power grid, even in scenarios involving the inclusion of new energy resources such as DERs.

However, developing a robust electricity price forecasting model (with either a probabilistic or point-wise approach) presents significant challenges. Not only does it display a level of seasonality, but it also exhibits highly nonlinear and time-varying features [14]. Consequently, there has been a strong interest in developing models that can effectively deal with these complex issues. Various approaches have explored advanced regression models to manage the complexities of accurately predicting electricity prices, including statistical time series analysis methods as well as various artificial intelligence (AI) algorithms.

Initially, traditional time series analysis forecasting models (e.g., linear regression, moving averages, auto-regressive models, etc.) were used, including more sophisticated ones (e.g., ARIMA, SARIMA, exponential smoothing, Box–Jenkins, state-space, or hybrid statistical models) [15], to capture patterns and seasonality in electricity price data. Several studies, including [16–19], have examined electricity price forecasting using this approach. These models were the basis for the development of more robust forecasting methods. In [16], an ARIMA model coupled with a neural network (a multi-layer perceptron in this case) was developed. The ARIMA model captured linear patterns, while the MLP modeled the remaining nonlinear residuals. The results suggest that the combined model produces lower forecast errors, measured by the mean absolute percentage error (MAPE) and mean absolute deviation (MAD), than either model used separately.

Likewise, the authors in [17,18] developed some ARIMA, SARIMA, GARCH, and hybrid models for modeling and forecasting electricity prices. They explore various model structures and evaluate their accuracy using statistical measures ( $R^2$  and MAPE). The results highlight the competitiveness of these hybrid models for this task. A hybrid model combining multivariate linear regression with ARIMA and Holt–Winters models is presented in [19]. Tests on data from the Iberian electricity market show superior performance (from MAPE) compared to some benchmark models, with promising results under different scenarios.

There are also studies [20–22] focusing on the comparison of different time series analysis methods for electricity price forecasting. In [20], the authors compare several prediction models, including SARIMA, SARIMAX, and ARIMA, to predict day-ahead electricity prices in Germany. The SARIMAX model with exogenous variables performed the best, enhancing the forecast accuracy. The authors in [21] compare double and triple exponential smoothing for electricity price forecasting from volatility, using elastic net regularization. The results show superior performance for triple exponential smoothing and reduced mean square error with regularization. The findings enable informed decision making for power generation scheduling in the electricity market. Furthermore, accurate forecasting results have been obtained through various auto-regressive statistical models and their derivatives, as presented in [22]. Detailed computational procedures are provided along with numerical results and performance (MAPE), with some promising results and issues to consider.

However, the time series analysis models commonly presented rely on linear relationships and stationarity, hindering their accuracy for data with high variability and seasonality, as well as predicting values multiple steps into the future, even in hybrid models. Nonetheless, these models perform acceptably when the seasonality of the data is low (e.g., week- or month-long patterns with small deviations). Consequently, these models prove more fitting for other energy-related tasks. They also promote the exploration of various forecasting methods. In this regard, machine learning (ML) algorithms have shown promising results in the prediction of electricity prices both in point and probabilistic approaches [23], given their suitability for the high variability (associated with nonlinearity) and seasonality evident in these data. Many researchers have developed predictive models to forecast electricity prices in various countries and scenarios. Some of the ML models widely used for this task include support vector machines (SVMs) [24–26], tree-based models [27–29], k-nearest neighbor (KNN) [30–32], shallow architectures of artificial neural networks (ANNs) [33–35], quantile regressor as a probabilistic forecasting approach [36–38], and different related hybrid models [39–46].

For instance, in [26], an electricity price and short-term load forecasting model is presented using improved SVM and KNN algorithms. The study utilized the New York Independent System Operator (NYISO) dataset for six months, and applied feature selection and extraction techniques. The modified SVM and KNN models are evaluated using metrics such as MAE, RMSE, and MAPE. In a similar vein, in [29], the prediction of electricity prices in Victoria, Australia was analyzed using various tree-based regression algorithms, including gradient boosting, decision tree, and random forest regression models, with the performance evaluated using metrics such as MAE and  $R^2$ . Furthermore, in [46], a hybrid machine learning model for short-term electricity price forecasting is proposed. The model merges linear regression with ensemble tree-based models. Metrics such as MSE and MAE are used to fit and evaluate the performance of the model. The results reveal that the proposed model outperforms other single and hybrid models in terms of prediction accuracy. In [39], the authors used the SVR-based hybrid model alongside various feature selection techniques to forecast electricity price spikes. Likewise, the authors in [45] developed a hybrid forecasting model that combines a seasonal component auto-regressive model with an ANN model. The model was applied to forecast day-ahead electricity prices with acceptable accuracy.

As with statistical time series analysis models, there are also some studies [47–49] that focus mainly on the comparison of different ML models for electricity price forecasting under specific contexts. In [47,48], the authors compare forecasting models for predicting short-term electricity prices in the Italian market. Different regression methods, including SVM, Gaussian process, decision trees, MLP, parametric and non-parametric methods are evaluated using performance metrics such as MAE, RMSE, R, and percentage error anomalies. Likewise, the authors in [49] provide an extensive overview of the current level of advancement in short-term electricity price prediction. They examine the application of single and hybrid machine learning models, assess their effectiveness through evaluation

metrics like MAE, RMSE, and MSE, and identify the influence of distinct features on their forecasting performance. In this case, they use data from the Nord Pool market.

However, the above ML models and comparisons for electricity price forecasting only focus on shallow learning. Since shallow ML models are sensitive to overfitting and gradient vanishing, they are limited in their ability to handle large data and complex non-linear (extreme variant) issues [50]. Therefore, the combination of intelligence optimization theories and advancements in computer technology has led to a growing research interest in predictive models in this area, particularly using deep learning (DL) architectures due to their remarkable performance and broad application scope. There have been developments in electricity price time series prediction models with different architectures, some of which are: convolutional network (CNN) [51–53], recurrent neural network (RNN)-based models [54–58], generative models [59,60], Bayesian networks (BNs) [61,62], and hybrid models (ensembles, signal preprocessing steps, among others) [63–66].

Among the various deep learning architectures used for this task, RNN-based models stand out for their recurrent feedback network framework. Unlike other forecasting models, RNNs consider the temporal and abstract correlations of the time series, enabling a more thorough and comprehensive modeling of the time series data. For instance, in [56,67], the authors present a hybrid forecasting model for the short-term prediction of electricity load and price. The model integrates wavelet transform with feature selection based on entropy and mutual information, utilizing LSTM networks. The results (using MAPE, RMSE, and variance) show in both cases the accuracy of the load and price predictions. Similarly, in [58], electricity price forecasting is explored, focusing on proposed long-term and short-term memory networks using historical prices, timestamps, and additional engineering features. The research results highlight the meaningful impact of feature selection on forecast accuracy and the importance of selecting appropriate test datasets, especially when drastic trend changes occur in historical data.

Despite the variety of objectives and architectures employed in the development of predictive models in this field, most of them have predominantly utilized a point or deterministic approach. However, there are other outstanding concerns. Firstly, raw features containing outliers and high-dimensional data can complicate feature extraction. Hence, feature preprocessing will be critical in obtaining representations that facilitate precise forecasting. Secondly, point forecast models lack the ability to adequately assess the forecast uncertainty in electricity prices, which is detrimental to the risk trade-off, considering the data noise [68]. Both of these factors are crucial to consider when using these models as input for informed decision-making tools. However, there is still a lack of development in these issues.

## 1.2. Key Contributions

The key contributions of this paper are summarized as follows:

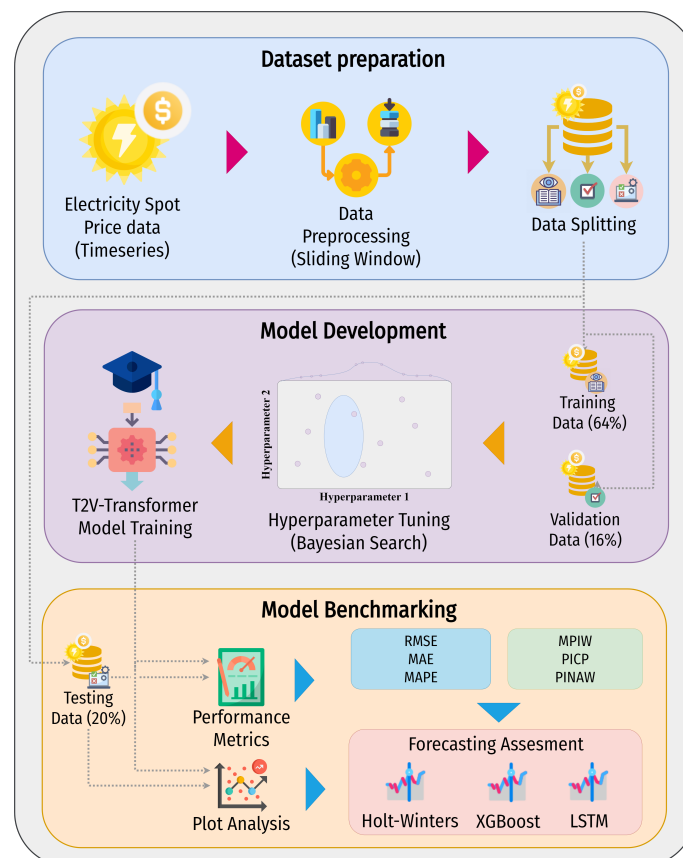
- Firstly, this paper presents a Time2Vec-Transformer encoder model, a hybrid architecture for temporal representation and sequence handling in order to provide inputs for point-wise and probabilistic hourly electricity price time series forecasting for intra-day periods (8 h ahead). This framework aims to provide valuable insights for the development of decision-making tools for different energy market players or new DER-related business models with public data. The hyperparameters of this model were tuned with Bayesian optimization algorithms.
- Subsequently, we provide a benchmark comparison of the proposed model with different time series forecasting approaches such as machine learning, deep learning, and time series analysis (TSA). This comparison was conducted for both point-wise and probabilistic forecasting with different and focused metrics. Therefore, the performance analysis of the prediction is more comprehensive and better focused on the development of new decision-making tools.

This paper is structured as follows: Section 2 presents the methodology, along with the details of the proposed T2V-TE model and a description of baseline benchmark models.

Section 3 showcases the experimental results and provides a comprehensive analysis. Finally, Section 4 summarizes the paper, highlighting key findings, and future research directions with conclusive remarks.

## 2. Methodology

The methodology employed in this study encompasses three distinct phases. In the initial phase (phase 1—dataset preparation), the primary objective is the acquisition and processing of data. This involves transforming the electricity price time series into a dataset that is compatible with the proposed Time2Vec-Transformer probabilistic forecasting model. Subsequently, this resultant dataset is partitioned into training, validation, and test sets. In the second phase (phase 2—model development), the focus shifts to the development and fine-tuning of the proposed T2V-TE model. In this phase, the datasets obtained in the previous phase are used to build and optimize the model. Finally, the third phase (phase 3—model benchmarking) focuses on evaluating the performance of the proposed model. This evaluation is performed in comparison to several state-of-the-art forecasting models and approaches, namely, XGBoost, Holt–Winters models, and LSTM-based models. A variety of performance metrics, both point and probabilistic forecasting metrics, are used for this evaluation. Figure 1 provides an overview of the methodology used in this study.



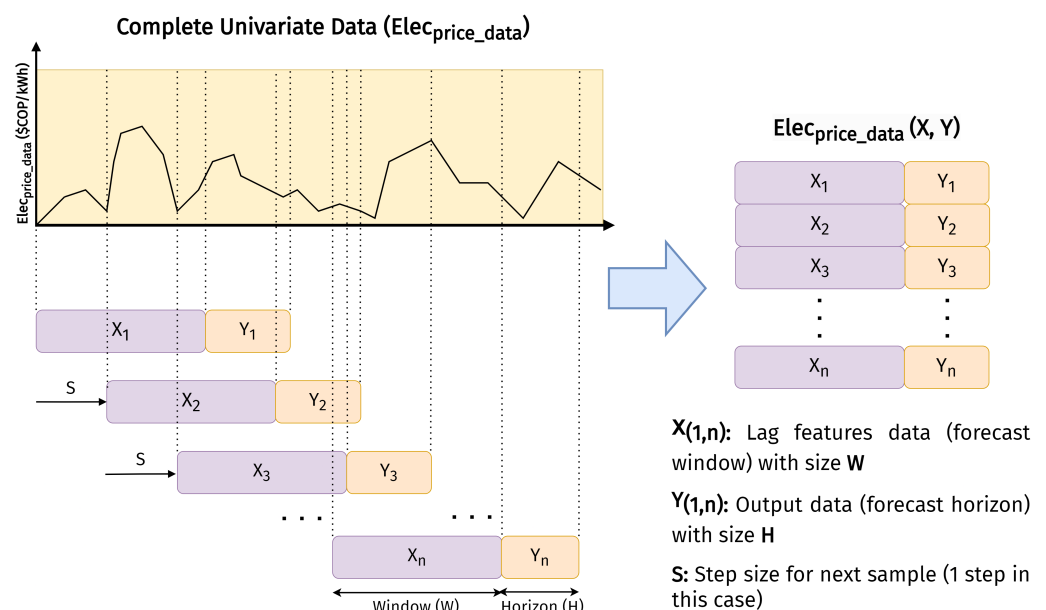
**Figure 1.** Framework of this research, including electricity price time series conversion, data splitting, proposed model and baseline development and further assessment.

### 2.1. Data Collection and Processing

In the field of prediction, time series refer to systematically structured and chronologically ordered datasets collected at consistent intervals, such as hourly, daily, or weekly observations. The inherent regularity of these data sequences introduces the potential for interdependence among the response variables, thereby posing significant hurdles for machine learning (ML) and deep learning (DL) algorithms commonly employed in regression tasks. These challenges primarily take the form of multi-collinearity and non-stationarity issues.

Hence, in order to prepare a dataset suitable for the forecasting model proposed in this paper, the time series is transformed into a reshaped dataset with non-time-dependent input features  $X$  and output variables  $Y$ . The format of the input features is [samples, lags, features], while the output variables have a shape of [samples, horizon], according to the architecture of the model presented below.

To accomplish this, the sliding window technique is applied. This method involves using data for training and forecasting, which is particularly useful for predicting time series data. In this case, a forecasting window of 24 h (last 24 hourly electricity prices, i.e., 24 time lags) and a forecasting horizon of 8 h (i.e., 24 outputs considering upper and lower forecasting intervals) were chosen based on time series analysis tools and extensive experiments conducted to tune the model. These values were chosen to align with the objective of the model, which is to provide intra-day predictions for electricity prices. Figure 2 shows the sliding window technique used for this case.



**Figure 2.** Dataset preparation with sliding window technique [69].

After the data transformation process described previously, we randomly split the dataset into 64%, 16%, and 20% for the training, validation, and test sets, respectively. The 20% test set represents a full year of hourly electricity prices. The remaining data were divided into 80% training (64% of all data) and 20% validation (16% of all data) sets, after extensive performance testing with different training–validation data ratios.

## 2.2. Model Development

This section provides a comprehensive overview of the time series forecasting model proposed in this article. To handle the issues of electricity price forecasting, we introduce a hybrid approach referred to as T2V-TE (deep transformer encoders with Time2Vec embedded time representation). The T2V-TE methodology synergizes the temporal representation capabilities of Time2Vec with the robust encoding capabilities of deep transformer models, thereby improving the reliability and accuracy of electricity price forecasts. The overall method is outlined as follows.

### Time2Vec-Transformer Model (T2V-TE)

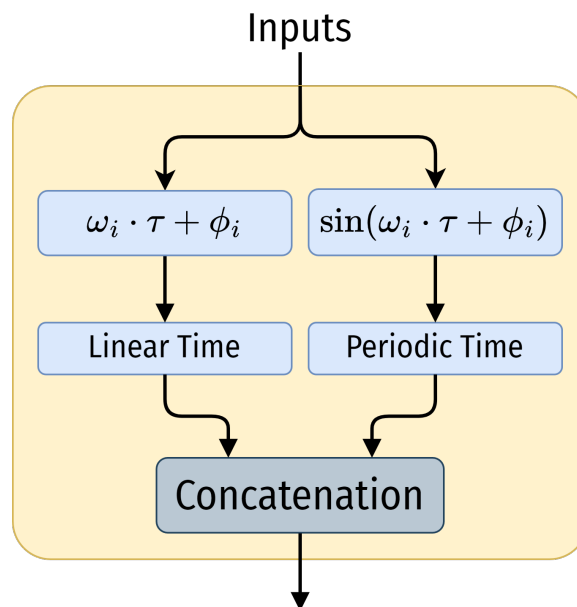
The hybrid T2V-TE model comprises a Time2Vec (T2V) [70,71] embedding processing as an input layer, followed by a stacked transformer encoder architecture [72], and a dense output layer divided into 3 segments. Each dense layer segment presents the number of units corresponding to the 8 h ahead prediction horizon, considering prediction, upper, and lower bounds, respectively. This model architecture requirement is critical, since for

stakeholders and related decision-making tools, the information obtained in a multi-step-ahead prediction is more valuable than a single-step-ahead prediction [73].

Thus, the processed hourly electricity price data enters the input layer corresponding to a Time2Vec block. Here, we obtain a time representation of the data that is both explicit (i.e., unaffected by changes in time scale) and embedded, making it independent of the specific model employed. This representation is well suited for integration into any neural-network-based architecture in order to capture periodic (using periodic signals as sinusoidal signals) and non-periodic input signal behavior. This layer automates the feature engineering of these data and enhances their modeling capabilities. This data processing approach is based on the deployment presented in Equation (1).

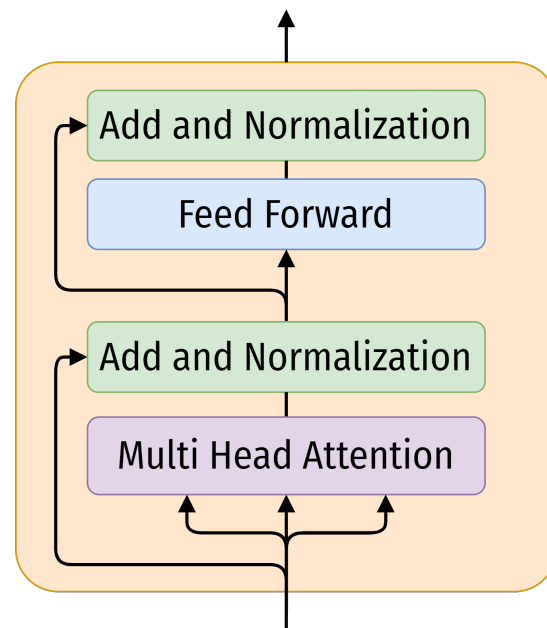
$$\text{Time2Vec}(\tau)[i] = \begin{cases} \omega_i \cdot \tau + \phi_i & \text{if } i = 0, \\ F(\omega_i \cdot \tau + \phi_i), & \text{if } 1 \leq i \leq k \end{cases} \quad (1)$$

In this context, we denote  $\tau$  as the raw data, while  $\omega$  and  $\phi$  represent sets of tunable parameters utilized in the model learning process. The activation function for this layer, denoted as  $F$ , is carefully selected to capture the periodic behavior inherent in the data. In this particular case, a sine function has been chosen due to its well-suited periodic features. Lastly, the hyperparameter  $k$  defines the output dimension of this layer, which constitutes the Time2Vec (T2V) output. The results obtained from the T2V layer encompass a composite representation of sinusoidal patterns learned by this layer, in conjunction with a linear representation according to  $i = 0$ , as depicted in Figure 3.



**Figure 3.** Time2Vec block.

All of this concatenated data feeds a set of stacked transformer encoder blocks. These architectures use only attention mechanisms for sequence-based data processing. This makes them well suited for time series prediction because time series data are inherently sequential. The transformer encoder does not utilize recurrent and convolutional layers, which are commonly used in sequence modeling. Instead, it employs stacked multi-head self-attention and full-connected layers. Both layers are followed by dropout and add (residuals) and normalization layers as shown in Figure 4.



**Figure 4.** Single transformer encoder block.

Each encoder block in this model creates representations (or encodings) of critical information about which sections of the input data are relevant to each other. These encodings are then passed to the next encoder block, where they are used to create even more complex data representations. To achieve this, each encoder block employs multi-head attention. The multi-head attention mechanism measures each input's importance to the overall input sequence and then weights the inputs accordingly.

Moreover, within each multi-head attention module of an encoder block, there are three trainable weight sets: query ( $Q$ ), key ( $K$ ), and value ( $V$ ) weights. Each attention head is responsible for distinguishing the significance levels among the input parameters. The standalone attention outputs, as depicted in Equation (2), are then concatenated and subjected to a linear transformation to obtain the desired dimension. Multi-head attention enables the model to capture many linkages and nuances for each input.

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{Q \cdot K}{d_k}\right) \cdot V \quad (2)$$

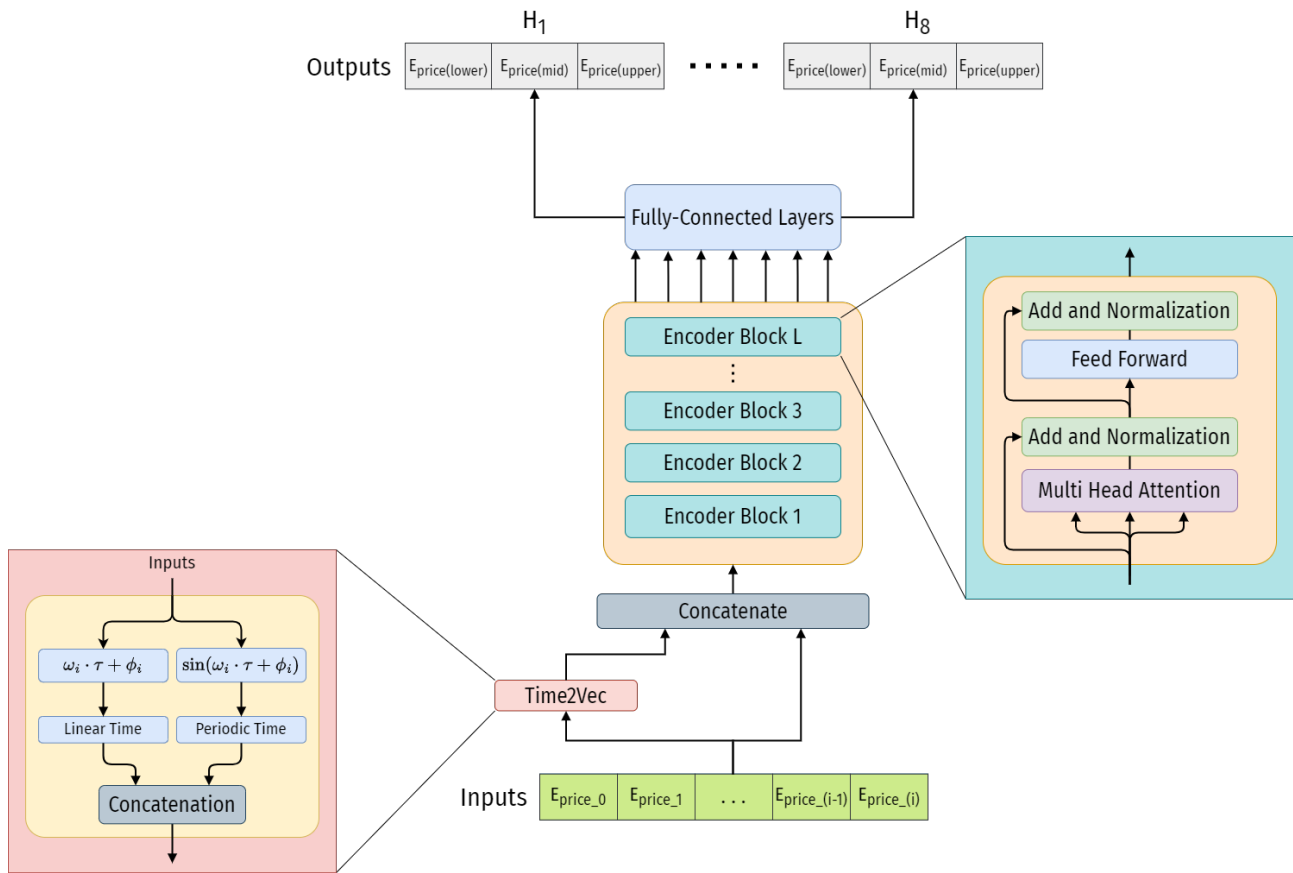
Here,  $d_k$  represents the dimension of the key and value vectors. The multi-head attention score is obtained by merging the output from  $h$  heads, as expressed in Equation (2), and then scaling it by a trainable projection parameter  $W$ , as illustrated in Equation (3). Notably, to avoid overfitting, dropout procedures of 20% and 40% have been incorporated between the attention and dense sub-layers within each transformer block, respectively.

$$\text{Multi-head Attention} = \text{Concat}(Att_1, \dots, Att_h) \cdot W \quad (3)$$

Once the transformer encoders have processed the data, these representations are fed to three different segments of eight dense output layers as described above, linked with different forecasting tasks: lowest electricity price, point prediction, and highest electricity price, respectively, as shown in the model architecture presented in Figure 5. We used the pinball loss function, as shown in Equation (4), as a loss function for each segment output with different quantiles (i.e., quantile regression), in order to obtain both the point-wise prediction and the probabilistic prediction interval.

$$\text{Loss}_\tau(E_p, \hat{E}_p) = \begin{cases} (1 - \tau)(E_p - \hat{E}_p) & \text{if } E_p - \hat{E}_p \geq 0 \\ \tau(E_p - \hat{E}_p) & \text{if } E_p - \hat{E}_p < 0 \end{cases} \quad (4)$$

where  $E_p$  is the observed electricity prices,  $\hat{E}_p$  is the predicted electricity prices, and  $\tau$  is the desired quantile. Finally, each output segment is scaled to its original range, and subsequently evaluated through the use of different performance metrics, as described below.



**Figure 5.** The proposed T2V-TE model for the hourly electricity price probabilistic forecasting.

### 2.3. Assessment of Forecasting Performance

To assess the performance of the proposed forecasting model, two distinct methodologies, namely, point-wise and probabilistic forecasting, will be employed. Performance metrics commonly utilized for this category of tasks have been chosen from the point-wise perspective. These metrics encompass root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE), as defined in Equations (5)–(7).

In these equations,  $E_p^{(i)}$  denotes the  $i$ -th observed electricity price value,  $\hat{E}_p^{(i)}$  represents the  $i$ -th data point predicted by the model, and  $N$  signifies the total number of samples. Importantly, in addition to serving as indicators of forecasting performance, these metrics also offer indirect assessments of the risk associated with the characteristics of the variable being predicted. Specifically, they provide measures of both risk neutrality (MAE) and risk aversion (RMSE).

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (E_p^{(i)} - \hat{E}_p^{(i)})^2} \quad (5)$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |E_p^{(i)} - \hat{E}_p^{(i)}| \quad (6)$$

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \frac{|E_p^{(i)} - \hat{E}_p^{(i)}|}{E_p^{(i)}} \times 100\% \quad (7)$$

From a probabilistic perspective, the performance metrics focus on three critical issues of this type of prediction for decision-making tasks: uncertainty as measured by the mean prediction interval width (MPIW), the forecasting precision as measured by the percentage interval normalized average width (PINAW), and the reliability as measured by the prediction interval coverage probability (PICP).

The MPIW is a metric that indicates the average size of the prediction intervals generated by a model, as shown in Equation (8). It represents the amplitude or amount of uncertainty associated with the model predictions. It is expressed in units of the variable (in this case COP/kWh). A larger MPIW indicates greater variability and, therefore, greater uncertainty in the model predictions.

$$\text{MPIW} = \frac{1}{N} \sum_{i=1}^N (E_{p(1-\alpha)}^{(i)} - E_{p(\alpha)}^{(i)}) \quad (8)$$

where  $\alpha$  is the confidence level,  $E_{p(\alpha)}^{(i)}$  and  $E_{p(1-\alpha)}^{(i)}$  are the lower and upper bounds of the prediction interval for each observation  $i$ , respectively.  $N$  is the total number of observations used to calculate this metric.

The PINAW is a metric that evaluates the accuracy of the model by comparing the average size of the prediction intervals to the size of the observed values. It is calculated as the percentage of the arithmetic mean of the observed values covered by the prediction intervals, as evidenced in Equation (9). A lower PINAW indicates higher accuracy, as it implies that the prediction intervals cover a smaller proportion of the observed data, suggesting better agreement between the predictions and the observed values.

$$\text{PINAW} = \frac{1}{N} \sum_{i=1}^N \frac{E_{p(1-\alpha)}^{(i)} - E_{p(\alpha)}^{(i)}}{\max(E_p) - \min(E_p)} \times 100 \quad (9)$$

where  $E_p$  is the training data, and  $\max(\cdot)$  and  $\min(\cdot)$  are the maximum and minimum functions, respectively.

The PICP is a metric that assesses the reliability of the forecasting model by determining what proportion of the observed values fall within the prediction intervals generated. It is expressed as a percentage, and the higher the PICP, the greater the confidence that the prediction intervals adequately capture the true values. The PICP metric is defined in Equation (10).

$$\text{PICP}(\alpha) = \frac{1}{N} \sum_{i=1}^n \mathbb{I}(E_p^{(i)} \in [E_{p(\alpha)}^{(i)}, E_{p(1-\alpha)}^{(i)}]) \times 100 \quad (10)$$

Here, the expression  $\mathbb{I}(\cdot)$  represents an indicator function, which evaluates to a value of 1 when the specified condition is true and 0 otherwise.

### 3. Results and Discussion

The objective of this study is to forecast electricity prices with an 8 h lead time, employing both point-wise and probabilistic approaches. The experimentation utilized hourly electricity price data from the Colombian wholesale market spanning from 1 January 2018 to 31 December 2022, encompassing a total of 43,824 samples. These data were obtained from the Colombian market operator web platform Sinergox, from XM [74].

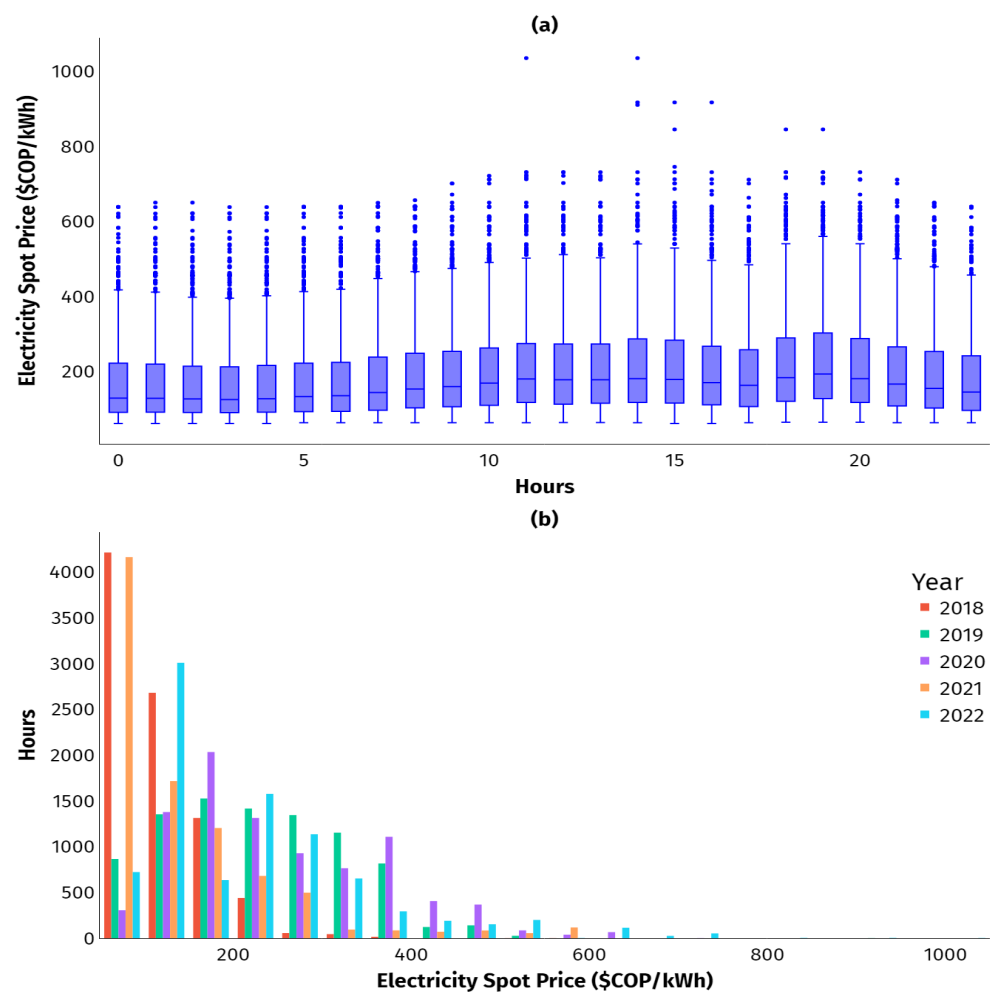
To assess the effectiveness of the proposed hybrid forecasting model it is compared with established techniques in time series analysis, as well as state-of-the-art machine learning and deep learning models commonly employed in similar time series forecasting tasks within the energy sector. Specifically, the baseline methods selected for this analysis include Holt–Winters [19], XGBoost [10,75], and two LSTM-based models [7,66], namely, Stacked LSTM and Attention-LSTM [55,57,67].

All computational simulations and data processing were conducted on a Windows® PC equipped with an Intel® Core i5+ 10300H processor running at 2.5 GHz, and featuring

16.00 GB of RAM. The experiments were performed using Google<sup>®</sup> Colab, and the following Python libraries and APIs were employed for hyperparameter tuning, model training, and performance assessment: Scikit-learn 0.24 [76], Keras backend [77,78], and Statsmodels [79].

### 3.1. Exploratory Data Analysis

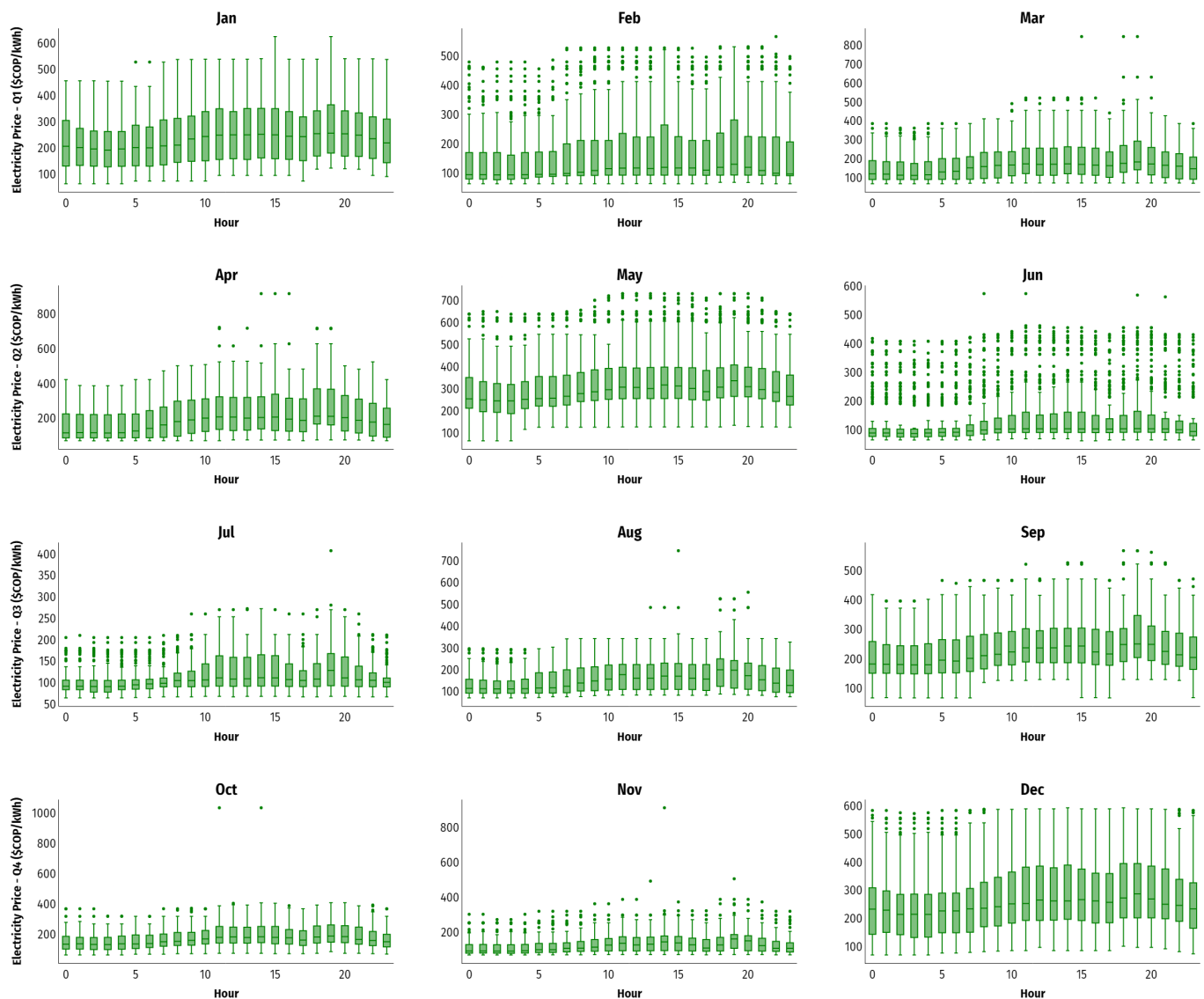
Using the approach described above, the exploratory data analysis (EDA) of the electricity price time series, as illustrated in Figures 6 and 7, affirms a robust correlation between fluctuations in electricity load and supply and corresponding variations in electricity prices. This relationship aligns with previous research findings [13,80–82]. Notably, during peak hours (specifically, hours 19 and 20), when power consumption is at its highest, the price rises to its daily peak. Conversely, during off-peak hours (hours 3 and 4) the reverse trend is observed. Instances of exceptionally high electricity prices (exceeding 650 COP/kWh) occasionally occur, typically attributed to climatic or operational factors. However, such occurrences remain infrequent, comprising less than 2.5% of the dataset.



**Figure 6.** Electricity price time series data analysis: (a) by hour; (b) yearly distribution.

This outlier description is evidenced in the histogram presented in Figure 6b which shows the electricity price distribution for each year from 2018 to 2022. The x-axis shows the price range, while the y-axis displays the frequency of occurrence for each price range. The histogram reveals that electricity prices are concentrated in the range of 120 to 350 COP/kWh, with the highest frequency of occurrence around 130 COP/kWh. Moreover, the histogram demonstrates a slight shift towards higher prices in 2022 compared to previous years, which may indicate an increasing demand for electricity or the influence of various external factors on the market dynamics.

Similarly, this same hourly behavior can be observed month by month throughout the study period, as depicted in Figure 7. When performing a quarterly analysis, it can be observed that Q2 and Q4 are the periods with the highest occurrence of electricity price outliers, with June being the month with the highest frequency. In addition, the months of February, October, and December show the highest variability in electricity prices in the years under study. These patterns can be attributed to multiple factors, including seasonal demand due to weather conditions, and coinciding with periods of increased economic activity, among others.



**Figure 7.** Electricity price time series hourly data analysis by month.

On the other hand, Table 1 presents the statistical outcomes for electricity prices arranged by year and quarter within this study. As noted in Section 2.1, the electricity price data from 2018 to 2021 serves as the training and validation data (80%), while the data from 2022 is the test data (20%). Upon analyzing the data, notable circumstances emerge. The most noteworthy year was 2022, which showed significant changes, especially in the second quarter. There was a substantial increase compared to previous years, making it the most variable year among the data.

**Table 1.** Descriptive statistics for Colombian electricity price.

Year	Quarter	Mean	Max.	Min.	Std.	Skew.	Kurt.
2018	Q1	145.73	244.29	61.76	27.96	0.37	0.67
	Q2	88.10	206.14	61.46	28.11	1.50	1.34
	Q3	92.02	231.37	64.42	34.82	1.74	1.77
	Q4	139.49	555.12	69.86	65.20	1.33	1.95
2019	Q1	283.27	510.57	75.92	63.87	−0.04	−0.36
	Q2	158.55	325.02	69.95	62.76	0.55	−0.65
	Q3	183.70	910.47	70.92	91.05	1.19	1.76
	Q4	288.16	524.80	72.58	89.23	0.22	−0.38
2020	Q1	362.90	646.02	71.34	95.01	0.11	1.13
	Q2	307.17	573.19	133.11	88.10	0.61	−0.46
	Q3	155.65	272.54	81.85	30.92	0.08	0.88
	Q4	180.51	745.50	78.46	63.76	1.56	6.02
2021	Q1	203.62	378.99	85.27	58.26	0.32	−0.68
	Q2	117.67	257.81	85.06	35.99	1.06	−0.02
	Q3	97.67	225.08	84.24	18.40	3.32	12.03
	Q4	182.13	592.91	81.76	150.30	1.63	1.14
2022	Q1	302.23	731.24	93.49	134.44	1.52	1.60
	Q2	117.68	1035.13	89.06	45.57	9.06	153.72
	Q3	169.57	845.12	97.10	90.88	1.89	5.71
	Q4	274.46	917.14	109.35	115.78	0.89	1.61

Therefore, according to the 2022 electricity price data behavior described above, scenarios of low (Q1), moderate (Q3), high (Q4), and very high (Q2) variability are identified based on aspects such as IQR, range of values, and the number of outliers. These operational scenarios will be employed in the performance testing of both point prediction and probabilistic forecasting for the proposed prediction model.

In this context, to assess the suitability of the data series as an input for the prediction model and the proposed baselines, a stationary analysis was conducted using tests such as the augmented Dickey–Fuller test (ADF). The results revealed that the  $p$ -value ( $6 \times 10^{-5}$ ) was significantly lower than the determined confidence interval values (0.05 and 0.01). This result indicates that the time series is stationary. This means that it has no trends or changing patterns. Thus, this makes it easier for the proposed forecasting model to capture the temporal relationships and make more accurate predictions without additional transformations. However, it is important to add that due to the requirements of the proposed architecture, a standardization preprocessing was performed.

### 3.2. Hyperparameter Tuning

In order to obtain the best performance for the proposed model, an exhaustive search for hyperparameters was undertaken, targeting the configuration that yielded the lowest pinball loss, as delineated in Equation (4). The Adam algorithm [83] was chosen as the optimizer for model training. Throughout the training and tuning processes, provisions were made for early stopping and dynamic learning rate reduction, transitioning from 0.001 to 0.00001, with the objective of mitigating overfitting and enhancing forecasting accuracy.

The hyperparameter search was carried out using a Bayesian approach over tunable parameters with flexible but limited ranges. These parameters included fixed values and stochastic uniform (U) variations for the different layer units of the model, excluding the last dense layer. To ensure a fair comparison among the models, the hyperparameter search was performed with 200 training epochs, and batches of 128 samples. The ‘hypetune-keras’ package, based on hyperopt [84], was used for this purpose. A summary of the search space and tuning outcomes for the layers of the proposed model is presented in Table 2.

**Table 2.** Hyperparameter search space and tuning values for the proposed model.

Model Hyperparameter	Hyperparameter Values	
	Value Range	Selected Value <sup>a</sup>
Encoder Blocks	2 + U(0, 6)	6
Number of Heads	2 + U(0, 6)	3
Head Size	128 + U(0, 128)	189
Encoder Dropout	0.1 + U(0, 0.5)	0.2
MLP Units	64 + U(0, 128)	143
Feedforward Dimension	2 + U(0, 2)	3
MLP Dropout	0.1 + U(0, 0.6)	0.4

<sup>a</sup> 373,583 parameters were trained in this model architecture.

Similarly, hyperparameter tuning was performed on each baseline model to ensure a fair and comprehensive comparison considering the conditions and features of each model. This process was conducted through a systematic search also using a Bayesian search approach, along with the development of an exhaustive set of experiments in order to tune them, according to the methods and hyperparameters applied for each model. The results of this joint process are presented in Table 3.

**Table 3.** Hyperparameters of baseline models.

Baseline Model	Hyperparameter	Value
Holt–Winters	seasonal_periods	24
	trend	'add'
	seasonal	'mul'
XGBoost	n_estimators	600
	max_depth	7
	Learning Rate ( $\alpha$ )	0.01
Stacked LSTM	Hidden Layers	3
	Neurons	[94,92,115]
	Learning rate ( $\alpha$ )	[0.001–0.00001]
	Activation Function	['Tanh','Tanh','Tanh']
	Optimizer	'Adam'
Attention-LSTM	num_heads	5
	head_size	253
	Hidden Layers	3
	Neurons	[124,84,94]
	Learning Rate ( $\alpha$ )	[0.001–0.00001]
	Activation Function	['Tanh','Tanh','Tanh']
	Optimizer	'Adam'

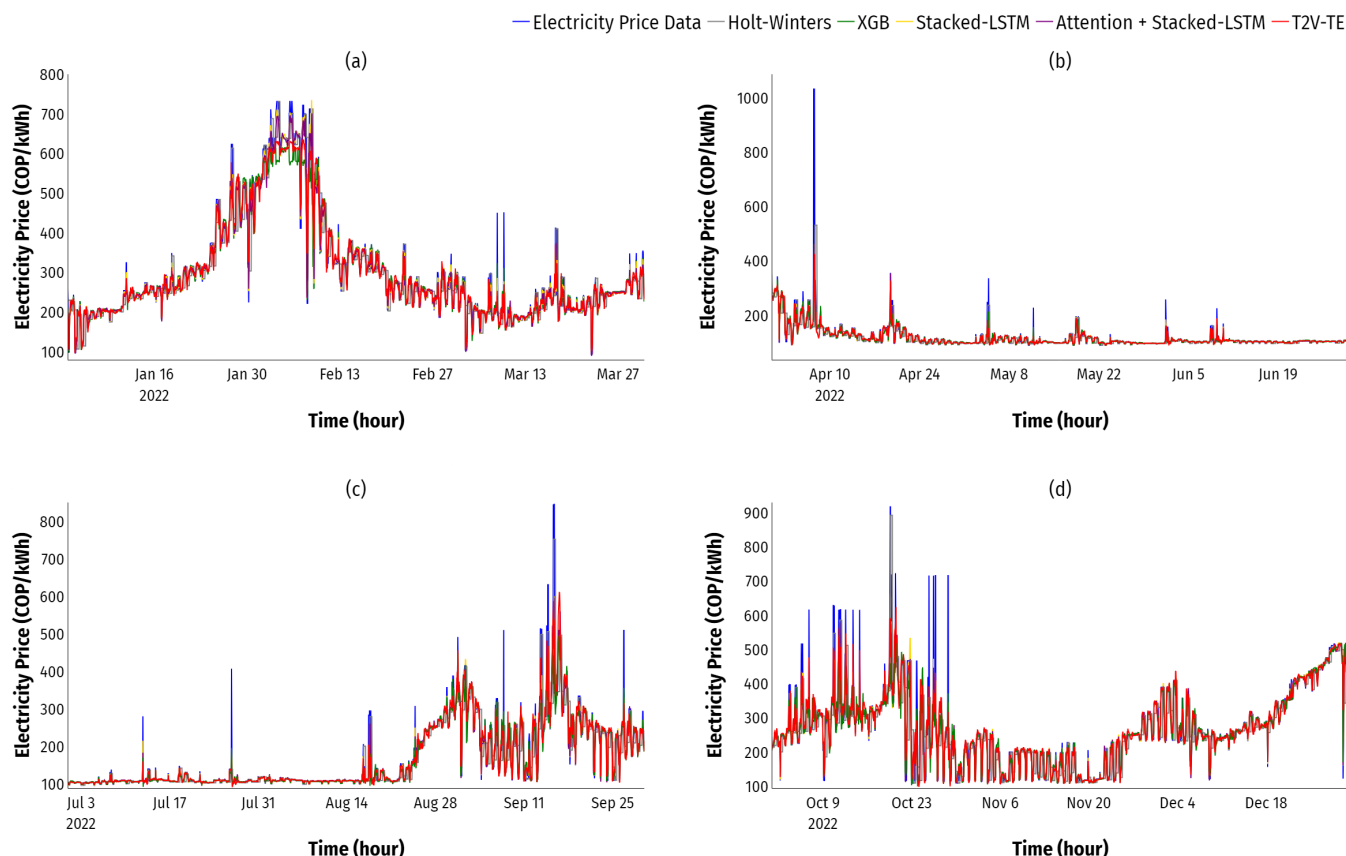
### 3.3. Model Performance Assessment

To compare the overall performance of the intra-day electricity price forecasting, the results of the point and probabilistic forecasting metrics of the model were compared with the baselines indicated above. For this purpose, the four forecast scenarios described in the EDA were evaluated across the entire test set: Q1-2022 (January to March quarter), Q2-2022 (April to June quarter), Q3-2022 (July to September quarter), and Q4-2022 (October to December quarter). All with hourly resolution. The following sections present the results and analysis of each forecast approach.

#### 3.3.1. Point Forecasting Approach

The evaluation of point forecast models for electricity prices during the test period covered the different forecast operational variability scenarios mentioned above (i.e., a specific quarter of the year 2022: Q1, Q2, Q3, and Q4). Figure 8 illustrates the performance of these models, including the proposed hybrid T2V-TE forecasting model, along with the

different baseline approaches. It can be seen that in the scenarios characterized by low to moderate variability (Q1 and Q3, respectively), there is an underlying parity between the forecasting models depicted in the graph. However, the proposed model presents a slightly superior performance compared to the proposed baselines.



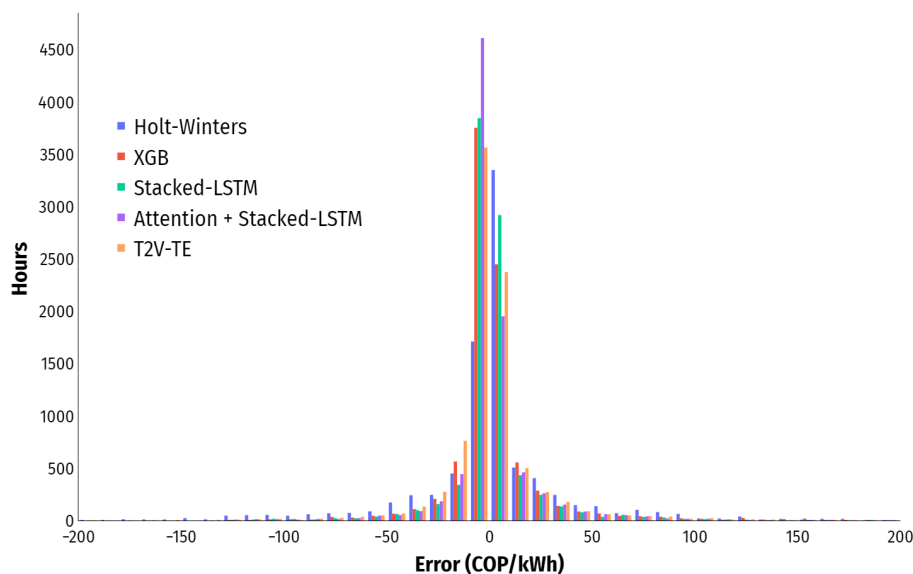
**Figure 8.** Comparative analysis of point forecasting approach performance: T2V-TE vs. proposed baseline models. (a) Q1-2022, (b) Q2-2022, (c) Q3-2022, (d) Q4-2022.

During these periods, where the electricity price presents relatively stable patterns and minor fluctuations, the different baseline models show competitive results, very similar to those presented by the proposed model. However, closer examination reveals that the proposed model consistently achieves slightly better predictive accuracy, capturing subtle nuances and temporal dependencies in the data.

However, this result trend undergoes a meaningful change when examining scenarios characterized by higher variability in electricity prices (Q4, and especially Q2), where deep-learning-based models (including the proposed model and the LSTM-based models) outperform statistical and machine learning models by a wide margin, given their robustness.

In these periods, marked by increased volatility and changing patterns in very short times, benchmark models, which are based on traditional statistical and machine learning techniques, have difficulty capturing the complex dynamics present in the data. In contrast, both T2V-TE and LSTM-based models demonstrate remarkable adaptability, effectively leveraging their deep learning architectures to extract meaningful representations and capture the intricate temporal dependencies inherent in the data.

This noteworthy performance superiority, observed in the face of increased variability, underscores the ability of the proposed model to overcome the limitations of traditional approaches, showing its potential to improve predictive capabilities and robustness in challenging and dynamic environments. To highlight these insights, the frequency plots of the proposed forecasting model and baseline errors in the whole test set are presented in Figure 9.



**Figure 9.** Error distribution for each point forecasting model.

The error distribution of the proposed model showcases a narrower spread (between  $-50$  and  $50$  COP/kWh) and a more symmetric shape, suggesting a better overall fit (avoid overestimating and underestimating electricity prices) to the observed data. On the other hand, the error distributions of the baseline models exhibit wider spreads (in most of the baselines, above the range between  $-70$  and  $70$  COP/kWh), with occasional outliers and skewness indicating deviations from the expected values.

In this context, Table 4 presents a comprehensive comparison of point prediction performance metrics between the T2V-TE model and the baseline models across various operational scenarios, taking into account the entire test dataset. This analysis conclusively demonstrates the enhanced forecasting quality achieved by the T2V-TE model. Specifically, the T2V-TE model exhibits improvements of at least 26.4% and 17.1% in RMSE and MAE values, respectively, when compared to the best-performing baseline (i.e., the Stacked LSTM model). This improvement is particularly pronounced in scenarios characterized by high and consistent hourly variability in electricity prices, such as the 2022-Q4 quarter, where the most substantial differences in performance between the T2V-TE model and the baseline models are observed.

**Table 4.** Comparison of performance metrics of the T2V-TE model with the proposed baselines for point forecasting of electricity prices.

Scenario	Model	Performance Metrics		
		MAE ( $\frac{\$COP}{kWh}$ )	RMSE ( $\frac{\$COP}{kWh}$ )	MAPE (%)
2022-Q1	Holt-Winters	27.21	49.15	10.00
	XGBoost	28.15	45.89	8.96
	Stacked LSTM	19.98	34.97	6.90
	Attention-LSTM	20.11	34.70	6.91
	T2V-TE	<b>15.15</b>	<b>29.38</b>	<b>4.83</b>
2022-Q2	Holt-Winters	12.50	41.38	8.21
	XGBoost	10.84	31.74	7.64
	Stacked LSTM	8.48	31.05	5.55
	Attention-LSTM	8.78	31.13	5.78
	T2V-TE	<b>6.50</b>	<b>28.18</b>	<b>4.12</b>

Table 4. Cont.

Scenario	Model	Performance Metrics		
		MAE ( $\frac{\$COP}{kWh}$ )	RMSE ( $\frac{\$COP}{kWh}$ )	MAPE (%)
2022-Q3	Holt–Winters	21.39	49.09	10.67
	XGBoost	19.47	39.19	10.08
	Stacked LSTM	15.44	35.69	7.60
	Attention-LSTM	15.69	35.79	7.78
	T2V-TE	<b>11.48</b>	<b>29.53</b>	<b>5.39</b>
2022-Q4	Holt–Winters	38.30	72.83	16.46
	XGBoost	31.51	57.47	12.70
	Stacked LSTM	26.91	53.81	11.13
	Attention-LSTM	27.12	53.92	11.22
	T2V-TE	<b>18.68</b>	<b>43.05</b>	<b>7.17</b>
All data	Holt–Winters	25.03	54.85	11.40
	XGBoost	22.62	45.02	9.87
	Stacked LSTM	17.79	40.13	7.82
	Attention-LSTM	18.03	40.18	7.95
	T2V-TE	<b>13.00</b>	<b>33.25</b>	<b>5.39</b>

These findings provide further evidence of the improved performance and robustness of the proposed model compared to the baselines, highlighting its potential as a reliable forecasting tool in practical applications such as decision making.

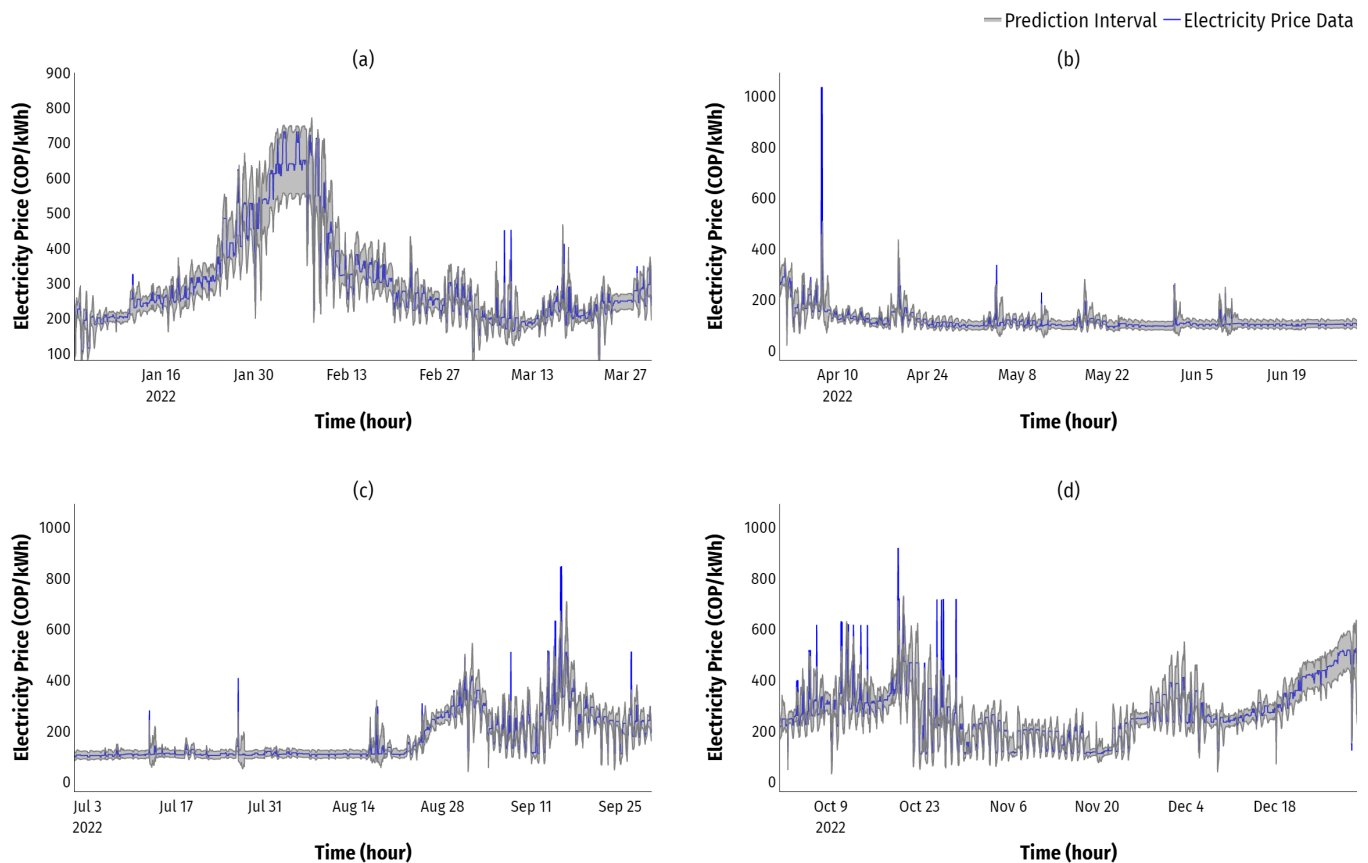
### 3.3.2. Probabilistic Forecasting Approach

The evaluation of probabilistic forecasting models for electricity prices was conducted during the specified test period, encompassing the same operational scenarios as the point forecasting models. The performance of the hybrid T2V-TE model, specifically its prediction intervals, is illustrated in Figure 10.

Firstly, the performance of the prediction intervals of the proposed model is remarkable in all operational scenarios of variability. These intervals effectively encompass the range of possible electricity prices, allowing for informed decision making even in the best- and worst-case scenarios within a certain time range. However, it is important to be aware of the existence of small inaccuracies that remain due to occasional sharp fluctuations, although they represent a negligible percentage of cases where prices exceed 700 COP/kWh between consecutive hours. This observation highlights an opportunity for further enhancement in addressing such discrepancies.

Additionally, it is worth noting that the T2V-LSTM prediction interval aligns with price fluctuations across all operational scenarios. Specifically, the interval widens as the price rises and contracts as the price declines. This trend becomes particularly evident in Q1 and the last 15 days of the Q4 scenario, where the sustained price increase results in a noticeable widening of the interval, thus enhancing the decision-making process by providing a more comprehensive perspective.

Similarly, to evaluate the performance of the proposed prediction model in a probabilistic context, the same baseline models as for the point prediction (i.e., HW, XGB, and LSTM-based models) were benchmarked. This evaluation focused on three critical issues for decision-making tasks: uncertainty (MPIW), precision (PINAW), and reliability (PICP). The results of this assessment are presented in Table 5.



**Figure 10.** T2V-TE probabilistic forecasting performance on the electricity price in test data. (a) Q1-2022, (b) Q2-2022, (c) Q3-2022, (d) Q4-2022.

**Table 5.** Performance metrics comparison of the T2V-TE model with the proposed baselines for probabilistic forecasting of electricity prices.

Scenario	Model	Performance Metrics		
		MPIW ( $\frac{COP}{kWh}$ )	PINAW (%)	PICP (%)
All data	Holt–Winters	180.43	18.53	93.04
	XGBoost	148.11	15.21	93.53
	Stacked LSTM	83.38	8.56	94.19
	Attention-LSTM	79.80	8.20	93.65
	T2V-TE	<b>71.70</b>	<b>7.37</b>	<b>94.87</b>

Considering this context, the Holt–Winters model exhibited an MPIW of 180.43 COP/kWh, which was the widest prediction interval among the proposed models. This indicates that this model had the highest degree of uncertainty in its predictions. Similarly, the PINAW score of 18.53% implies that, on average, the prediction intervals of this model covered only this percentage of the observed data range, indicating a moderate level of accuracy. Furthermore, the PICP score of 93.04% indicates that a large proportion of the observed data fell within the intervals predicted by this model, demonstrating good reliability. However, this model has the lowest performance among the presented models.

On the other hand, the XGBoost model outperformed Holt–Winters, achieving a lower MPIW (17.9% lower). This lower value suggests a higher accuracy in the predictions of this model. The PINAW value of 15.21% shows that, on average, the prediction intervals of XGBoost were narrower than those of the Holt–Winters model, indicating increased accuracy. Additionally, the PICP score of 93.53% indicates a comparable level of reliability.

bility to Holt–Winters, with a significant portion of the observed data falling within the predicted intervals.

The Stacked LSTM forecasting model displayed noteworthy improvements. With an MPIW of 83.78 COP/kWh, it exhibited a significant reduction in uncertainty compared to both Holt–Winters and XGBoost. This decrease in MPIW indicates the effectiveness of the model in producing more accurate predictions. Moreover, the PINAW value saw a substantial decrease from 43.72%. This means that, on average, the prediction intervals covered only a small percentage of the observed data, demonstrating a much higher level of accuracy. Furthermore, the model achieved a PICP score of 94.19%, indicating enhanced reliability. Overall, the performance of this model outperformed the previously mentioned models by a wide margin and was close to that of the proposed model. In contrast to the focal approach, the incorporation of attention layers on top of LSTM-based models, as shown by Attention-LSTM, led to slight improvements in performance metrics within this context compared to the Stacked LSTM model. These improvements were particularly evident in terms of precision and reduced uncertainty (MPIW: 71.70 COP/kWh, PINAW: 8.20%), although reliability (PICP: 93.65%) lagged slightly.

Outperforming all other models, T2V-TE achieved the lowest MPIW of 71.70 COP/kWh (14.4% lower than LSTM), indicating the highest precision and the least uncertainty among all the evaluated models. The PINAW value of 7.37% indicates that the prediction intervals covered an even smaller percentage of the observed data, thus demonstrating superior accuracy. Furthermore, the model attained the highest PICP score of 94.58%, showcasing the best reliability among all the models. This consolidates T2V-TE as the forecasting model with the best overall results among those proposed in this study.

Finally, to determine the statistical significance of the outcomes achieved by the proposed T2V-TE model, Wilcoxon tests were systematically performed at a significance level of 0.05. These tests employed the error between the actual and forecast values to conduct pairwise comparisons between the T2V-TE model and the designated baseline models. The results of these significance tests are outlined in Table 6. Notably, the analysis reveals that the T2V-TE model exhibits statistically significant differences in performance when compared to the Holt–Winters, XGBoost, Stacked LSTM, and Attention-LSTM models. This corroborates the robustness and validity of the findings obtained through the proposed model.

**Table 6.** Results of Wilcoxon test between T2V-TE and forecasting baselines.

Forecasting Models	<i>p</i> -Value	Statistical Significance
T2V-TE—Holt–Winters	$1.32 \times 10^{-15}$	Yes
T2V-TE—XGBoost	$1.76 \times 10^{-9}$	Yes
T2V-TE—Stacked LSTM	$9.16 \times 10^{-7}$	Yes
T2V-TE—Attention-LSTM	$5.07 \times 10^{-8}$	Yes

#### 4. Conclusions

The present research proposes a novel and finely tuned hybrid model known as Time2Vec-Transformer encoder (T2V-TE), for multi-step forecasting. This model is designed to predict the future hourly electricity prices in the Colombian wholesale market, considering future power grid scenarios (i.e., intra-day markets and DERs integration). This model serves as a valuable tool to empower decision-making processes for various market players and facilitates the emergence of new business models. It particularly addresses the challenges posed by the widespread integration of different types of distributed energy resources and the ensuing dynamic nature of electricity prices within this changing landscape.

Hourly electricity price forecasts were made considering the new short-term market perspectives within the policy and operational context of the data used, such as intra-day periods (in this case, 8 h ahead—three time blocks). Two approaches were considered: point-wise and probabilistic forecasting. In the case of probabilistic forecasts, the uncertainty

in the electricity price forecast is quantified by the quantile regression loss function. This allowed for a more comprehensive understanding of the range of potential price outcomes, considering the inherent volatility and unpredictability of the electricity market. In this way, the benchmark results showed that T2V-TE outperformed the traditional time series analysis and ML and DL statistical forecasting approaches presented. This is due to the advantages of obtaining and retaining hidden information in electricity price behaviors provided by the embedded explicit representation of temporal and data features offered by the Time2Vec layer, coupled with the pattern detection in ordered data linked to the attention mechanisms of the transformer encoder.

From a point prediction perspective, the T2V-TE model has shown superior performance compared to the proposed baseline models for the scenarios of this study (overall, the MAE, MAPE, and RMSE metrics linked to prediction errors in the T2V-TE model have been comparatively reduced by an average of 20%). This reduction in prediction errors, particularly those associated with risk aversion (RMSE) and risk neutrality (MAE), showed that the proposed model is a reliable tool for decision-making tool development in this area. This may imply a reduction in risk in decision making, as the model provides more accurate price predictions, minimizing potential losses and maximizing profitability. It can also be shown that the overall performance of deep-learning-based models, including T2V-TE and LSTM-based models, consistently surpasses that of time series analysis and machine learning models within this context.

Likewise, from a probabilistic prediction perspective, based on the comprehensive comparative analysis of these models, it is evident that T2V-TE surpasses Holt–Winters, XGBoost, and Stacked LSTM in terms of MPIW, PINAW, and PICP (overall, these metrics linked to probabilistic prediction errors in the T2V-TE model have been comparatively reduced by an average of 7.2%). T2V-TE has higher accuracy, narrower prediction intervals, and higher reliability (linked to the degree of coverage of observed electricity prices in the forecasting interval) than the baseline models. This is most clearly evident in the quartiles with the highest price volatility (i.e., 2022-Q1 and 2022-Q4).

Overall, this model offers a potentially effective framework for developing electricity market decision-making tools under different scenarios. The risk-mitigating capabilities of this model and the invaluable information it provides provide stakeholders with the opportunity to optimize their energy trading activities and make informed decisions, even in a dynamic and ever-changing market context.

In the scope of this study, the analysis of the electricity price time series for forecasting was conducted without factoring in the influence of highly correlated exogenous variables, such as electricity load. Additionally, another promising avenue for research involves estimating the probabilities associated with the occurrence of upper, middle, and lower values of electricity prices for a given hour. This approach can yield valuable insights for decision-making processes.

In future research endeavors, the objective is to augment the performance of the model for decision support tools by incorporating these variables and addressing these pertinent issues. By integrating highly correlated exogenous variables and delving into probability estimation, the predictive capabilities of the model can be further enhanced, providing more actionable and insightful results for energy market decision makers.

**Author Contributions:** Conceptualization, R.M.-C.; methodology, S.C.-L. and R.M.-C.; software, S.C.-L.; validation, S.C.-L., R.M.-C., J.L.-S. and D.C.; formal analysis, S.C.-L. and R.M.-C.; investigation, S.C.-L. and R.M.-C.; data curation, S.C.-L.; writing—original draft preparation, S.C.-L.; writing—review and editing, S.C.-L., R.M.-C., J.L.-S. and D.C.; visualization, S.C.-L.; supervision, R.M.-C. and J.L.-S.; funding acquisition, D.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** Data for this study is readily available to the public through SINERGOX (url: <https://sinergox.xm.com.co/trpr/Paginas/Historicos/Historicos.aspx>), which is the data analysis website of XM, the operator of the Colombian electricity market.

**Acknowledgments:** The authors would like to thank the support of the Universidad Icesi and Universidad Autónoma de Occidente in Cali, Colombia. As well, part of this work was partially funded by the starting grant IV-TFA056 entitled “Machine learning for Smart Energy Systems” by the Research Direction at Universidad del Rosario. Likewise, we would like to thank to the Center of Resources for Learning and Research (CRAI) at Universidad del Rosario for their help with the heuristic state-of-the-art for this manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

ANN	Artificial neural network
COP	Colombian Peso
DERs	Distributed energy resources
DL	Deep learning
EDA	Exploratory data analysis
HW	Holt–Winters model
LSTM	Long short-term memory
ML	Machine learning
RNN	Recurrent neural network
TSA	Time series analysis
T2V	Time2Vec
T2V-TE	Time2Vec + transformer encoder
XGB	Extreme gradient boosting

## References

1. Cantillo-Luna, S.; Moreno-Chuquen, R.; Lopez-Sotelo, J.A. Intra-day Electricity Price Forecasting Based on a Time2Vec-LSTM Neural Network Model. In Proceedings of the 2023 IEEE Colombian Conference on Applications of Computational Intelligence (ColCACI), Bogota, Colombia, 26–28 July 2023; pp. 1–6.
2. Trivedi, R.; Patra, S.; Sidqi, Y.; Bowler, B.; Zimmermann, F.; Deconinck, G.; Papaemmanouil, A.; Khadem, S. Community-based microgrids: Literature review and pathways to decarbonise the local electricity network. *Energies* **2022**, *15*, 918. [[CrossRef](#)]
3. Cantillo-Luna, S.; Moreno-Chuquen, R.; Chamorro, H.R.; Sood, V.K.; Badsha, S.; Konstantinou, C. Blockchain for Distributed Energy Resources Management and Integration. *IEEE Access* **2022**, *10*, 68598–68617. [[CrossRef](#)]
4. Sridharan, V.; Tuo, M.; Li, X. Wholesale electricity price forecasting using integrated long-term recurrent convolutional network model. *Energies* **2022**, *15*, 7606. [[CrossRef](#)]
5. Burger, S.P.; Luke, M. Business models for distributed energy resources: A review and empirical analysis. *Energy Policy* **2017**, *109*, 230–248. [[CrossRef](#)]
6. Lu, X.; Qiu, J.; Lei, G.; Zhu, J. Scenarios modelling for forecasting day-ahead electricity prices: Case studies in Australia. *Appl. Energy* **2022**, *308*, 118296. [[CrossRef](#)]
7. Zhou, S.; Zhou, L.; Mao, M.; Tai, H.M.; Wan, Y. An optimized heterogeneous structure LSTM network for electricity price forecasting. *IEEE Access* **2019**, *7*, 108161–108173. [[CrossRef](#)]
8. Jiang, L.; Hu, G. A Review on Short-Term Electricity Price Forecasting Techniques for Energy Markets. In Proceedings of the 2018 15th International Conference on Control, Automation, Robotics and Vision (ICARCV), Singapore, 18–21 November 2018. [[CrossRef](#)]
9. Pourdayaei, A.; Mokhlis, H.; Illias, H.A.; Kaboli, S.H.A.; Ahmad, S.; Ang, S.P. Hybrid ANN and Artificial Cooperative Search Algorithm to Forecast Short-Term Electricity Price in De-Regulated Electricity Market. *IEEE Access* **2019**, *7*, 125369–125386. [[CrossRef](#)]
10. Zhao, X.; Li, Q.; Xue, W.; Zhao, Y.; Zhao, H.; Guo, S. Research on Ultra-Short-Term Load Forecasting Based on Real-Time Electricity Price and Window-Based XGBoost Model. *Energies* **2022**, *15*, 7367. [[CrossRef](#)]
11. Lago, J.; De Ridder, F.; Vranx, P.; De Schutter, B. Forecasting day-ahead electricity prices in Europe: The importance of considering market integration. *Appl. Energy* **2018**, *211*, 890–903. [[CrossRef](#)]
12. Tan, Y.Q.; Shen, Y.X.; Yu, X.Y.; Lu, X. Day-ahead electricity price forecasting employing a novel hybrid frame of deep learning methods: A case study in NSW, Australia. *Electr. Power Syst. Res.* **2023**, *220*, 109300. [[CrossRef](#)]

13. Barrientos, J.; Rodas, E.; Velilla, E.; Lopera, M.; Villada, F. A model for forecasting electricity prices in Colombia. *Lect. Econ.* **2012**, *91*–127.
14. de Marcos, R.A.; Bello, A.; Reneses, J. Short-term forecasting of electricity prices with a computationally efficient hybrid approach. In Proceedings of the 2017 14th International Conference on the European Energy Market (EEM), Dresden, Germany, 6–9 June 2017; pp. 1–6.
15. Hyndman, R.J.; Athanasopoulos, G. *Forecasting: Principles and Practice*; OTexts: Columbia, MD, USA, 2018.
16. Skopal, R. Short-term hourly price forward curve prediction using neural network and hybrid ARIMA-NN model. In Proceedings of the 2015 International Conference on Information and Digital Technologies, Zilina, Slovakia, 7–9 July 2015. [[CrossRef](#)]
17. Liu, H.; Shi, J. Applying ARMA–GARCH approaches to forecasting short-term electricity prices. *Energy Econ.* **2013**, *37*, 152–166. [[CrossRef](#)]
18. Rajan, P.; Chandrakala, K.V. Statistical Model Approach of Electricity Price Forecasting for Indian Electricity Market. In Proceedings of the 2021 IEEE Madras Section Conference (MASCON), Chennai, India, 27–28 August 2021. [[CrossRef](#)]
19. Bissing, D.; Klein, M.T.; Chinnathambi, R.A.; Selvaraj, D.F.; Ranganathan, P. A Hybrid Regression Model for Day-Ahead Energy Price Forecasting. *IEEE Access* **2019**, *7*, 36833–36842. [[CrossRef](#)]
20. Abunofal, M.; Poshia, N.; Qussous, R.; Weidlich, A. Comparative Analysis of Electricity Market Prices Based on Different Forecasting Methods. In Proceedings of the 2021 IEEE Madrid PowerTech, Madrid, Spain, 28 June–2 July 2021. [[CrossRef](#)]
21. Banitalebi, B.; Hoque, M.E.; Appadoo, S.S.; Thavaneswaran, A. Regularized Probabilistic Forecasting of Electricity Wholesale Price and Demand. In Proceedings of the 2020 IEEE Symposium Series on Computational Intelligence (SSCI), Canberra, Australia, 1–4 December 2020. [[CrossRef](#)]
22. Shikhina, A.; Kochengin, A.; Chrysostomou, G.; Shikhin, V. Investigation of autoregressive forecasting models for market electricity price. In Proceedings of the 2020 IEEE 20th Mediterranean Electrotechnical Conference, Palermo, Italy, 16–18 June 2020. [[CrossRef](#)]
23. Wang, J.; Zhou, Y.; Li, Z. Hour-ahead photovoltaic generation forecasting method based on machine learning and multi objective optimization algorithm. *Appl. Energy* **2022**, *312*, 118725. [[CrossRef](#)]
24. Chaâbane, N. A novel auto-regressive fractionally integrated moving average–least-squares support vector machine model for electricity spot prices prediction. *J. Appl. Stat.* **2014**, *41*, 635–651. [[CrossRef](#)]
25. Shrivastava, N.A.; Khosravi, A.; Panigrahi, B.K. Prediction interval estimation of electricity prices using PSO-tuned support vector machines. *IEEE Trans. Ind. Inform.* **2015**, *11*, 322–331. [[CrossRef](#)]
26. Ali, M.; Khan, Z.A.; Mujeeb, S.; Abbas, S.; Javaid, N. Short-term electricity price and load forecasting using enhanced support vector machine and K-nearest neighbor. In Proceedings of the 2019 Sixth HCT Information Technology Trends (ITT), Ras Al Khaimah, United Arab Emirates, 20–21 November 2019; pp. 79–83.
27. González, C.; Mira-McWilliams, J.; Juárez, I. Important variable assessment and electricity price forecasting based on regression tree models: Classification and regression trees, Bagging and Random Forests. *IET Gener. Transm. Distrib.* **2015**, *9*, 1120–1128. [[CrossRef](#)]
28. Pórtoles, J.; González, C.; Moguerza, J. Electricity Price Forecasting with Dynamic Trees: A Benchmark Against the Random Forest Approach. *Energies* **2018**, *11*, 1588. [[CrossRef](#)]
29. Orenc, S.; Acar, E.; Ozerdem, M.S. The Electricity Price Prediction of Victoria City Based on Various Regression Algorithms. In Proceedings of the 2022 Global Energy Conference (GEC), Batman, Turkey, 26–29 October 2022. [[CrossRef](#)]
30. Ashfaq, T.; Javaid, N. Short-Term Electricity Load and Price Forecasting using Enhanced KNN. In Proceedings of the 2019 International Conference on Frontiers of Information Technology (FIT), Islamabad, Pakistan, 16–18 December 2019. [[CrossRef](#)]
31. Johannesen, N.J.; Kolhe, M.; Goodwin, M. Deregulated Electric Energy Price Forecasting in NordPool Market using Regression Techniques. In Proceedings of the 2019 IEEE Sustainable Power and Energy Conference (iSPEC), Beijing, China, 21–23 November 2019. [[CrossRef](#)]
32. Khan, S.; Khan, Z.A.; Noshad, Z.; Javaid, S.; Javaid, N. Short Term Load and Price Forecasting using Tuned Parameters for K-Nearest Neighbors. In Proceedings of the 2019 Sixth HCT Information Technology Trends (ITT), Ras Al Khaimah, United Arab Emirates, 20–21 November 2019. [[CrossRef](#)]
33. Pavićević, M.; Popović, T. Forecasting Day-Ahead Electricity Metrics with Artificial Neural Networks. *Sensors* **2022**, *22*, 1051. [[CrossRef](#)]
34. Wang, D.; Luo, H.; Grunder, O.; Lin, Y.; Guo, H. Multi-step ahead electricity price forecasting using a hybrid model based on two-layer decomposition technique and BP neural network optimized by firefly algorithm. *Appl. Energy* **2017**, *190*, 390–407. [[CrossRef](#)]
35. Rafiei, M.; Niknam, T.; Khooban, M.H. Probabilistic forecasting of hourly electricity price by generalization of ELM for usage in improved wavelet neural network. *IEEE Trans. Ind. Inform.* **2016**, *13*, 71–79. [[CrossRef](#)]
36. Bunn, D.; Andresen, A.; Chen, D.; Westgaard, S. Analysis and forecasting of electricity price risks with quantile factor models. *Energy J.* **2016**, *37*, 1–32. [[CrossRef](#)]
37. Hagfors, L.I.; Bunn, D.; Kristoffersen, E.; Staver, T.T.; Westgaard, S. Modeling the UK electricity price distributions using quantile regression. *Energy* **2016**, *102*, 231–243. [[CrossRef](#)]
38. Uniejewski, B.; Weron, R. Regularized quantile regression averaging for probabilistic electricity price forecasting. *Energy Econ.* **2021**, *95*, 105121. [[CrossRef](#)]

39. Zhao, J.H.; Dong, Z.Y.; Li, X.; Wong, K.P. A framework for electricity price spike analysis with advanced data mining methods. *IEEE Trans. Power Syst.* **2007**, *22*, 376–385. [[CrossRef](#)]
40. Zhang, J.; Tan, Z.; Li, C. A novel hybrid forecasting method using GRNN combined with wavelet transform and a GARCH model. *Energy Sources Part B Econ. Plan. Policy* **2015**, *10*, 418–426. [[CrossRef](#)]
41. Akbilgic, O.; Bozdogan, H.; Balaban, M.E. A novel hybrid RBF neural networks model as a forecaster. *Stat. Comput.* **2014**, *24*, 365–375. [[CrossRef](#)]
42. Yang, Z.; Ce, L.; Lian, L. Electricity price forecasting by a hybrid model, combining wavelet transform, ARMA and kernel-based extreme learning machine methods. *Appl. Energy* **2017**, *190*, 291–305. [[CrossRef](#)]
43. Nazar, M.S.; Fard, A.E.; Heidari, A.; Shafie-khah, M.; Catalão, J.P. Hybrid model using three-stage algorithm for simultaneous load and price forecasting. *Electr. Power Syst. Res.* **2018**, *165*, 214–228. [[CrossRef](#)]
44. Grossi, L.; Nan, F. Robust forecasting of electricity prices: Simulations, models and the impact of renewable sources. *Technol. Forecast. Soc. Chang.* **2019**, *141*, 305–318. [[CrossRef](#)]
45. Marcjasz, G.; Uniejewski, B.; Weron, R. On the importance of the long-term seasonal component in day-ahead electricity price forecasting with NARX neural networks. *Int. J. Forecast.* **2019**, *35*, 1520–1532.
46. Alkawaz, A.N.; Abdellatif, A.; Kanesan, J.; Khairuddin, A.S.M.; Ghenni, H.M. Day-Ahead Electricity Price Forecasting Based on Hybrid Regression Model. *IEEE Access* **2022**, *10*, 108021–108033. [[CrossRef](#)]
47. Imani, M.H.; Bompard, E.; Colella, P.; Huang, T. Predictive methods of electricity price: An application to the Italian electricity market. In Proceedings of the 2020 IEEE International Conference on Environment and Electrical Engineering and 2020 IEEE Industrial and Commercial Power Systems Europe (EEEIC), Madrid, Spain, 9–12 June 2020. [[CrossRef](#)]
48. Shah, I.; Bibi, H.; Ali, S.; Wang, L.; Yue, Z. Forecasting One-Day-Ahead Electricity Prices for Italian Electricity Market Using Parametric and Nonparametric Approaches. *IEEE Access* **2020**, *8*, 123104–123113. [[CrossRef](#)]
49. Mubarak, H.; Ahmad, S.; Hossain, A.A.; Horan, B.; Abdellatif, A.; Mekhilef, S.; Seyedmahmoudian, M.; Stojcevski, A.; Mokhlis, H.; Kanesan, J.; et al. Short-term Electricity Price Forecasting Using Interpretable Hybrid Machine Learning Models. In Proceedings of the 2023 IEEE IAS Global Conference on Renewable Energy and Hydrogen Technologies (GlobConHT), Male, Maldives, 11–12 March 2023. [[CrossRef](#)]
50. Pugliese, R.; Regondi, S.; Marini, R. Machine learning-based approach: Global trends, research directions, and regulatory standpoints. *Data Sci. Manag.* **2021**, *4*, 19–29.
51. Li, W.; Becker, D.M. Day-ahead electricity price prediction applying hybrid models of LSTM-based deep learning methods and feature selection algorithms under consideration of market coupling. *Energy* **2021**, *237*, 121543.
52. Khan, Z.A.; Fareed, S.; Anwar, M.; Naeem, A.; Gul, H.; Arif, A.; Javaid, N. Short term electricity price forecasting through convolutional neural network (CNN). In *Web, Artificial Intelligence and Network Applications: Proceedings of the Workshops of the 34th International Conference on Advanced Information Networking and Applications (WAINA-2020)*, Caserta, Italy, 15–17 April 2020; Springer: Berlin/Heidelberg, Germany, 2020; pp. 1181–1188.
53. Zhang, B.; Song, C.; Jiang, X.; Li, Y. Electricity price forecast based on the STL-TCN-NBEATS model. *Heliyon* **2023**, *9*, e13029. [[CrossRef](#)]
54. Ugurlu, U.; Oksuz, I.; Tas, O. Electricity price forecasting using recurrent neural networks. *Energies* **2018**, *11*, 1255. [[CrossRef](#)]
55. Li, Y.; Zhu, Z.; Kong, D.; Han, H.; Zhao, Y. EA-LSTM: Evolutionary attention-based LSTM for time series prediction. *Knowl.-Based Syst.* **2019**, *181*, 104785. [[CrossRef](#)]
56. Memarzadeh, G.; Keynia, F. Short-term electricity load and price forecasting by a new optimal LSTM-NN based prediction algorithm. *Electr. Power Syst. Res.* **2021**, *192*, 106995.
57. Chitalia, G.; Pipattanasomporn, M.; Garg, V.; Rahman, S. Robust short-term electrical load forecasting framework for commercial buildings using deep recurrent neural networks. *Appl. Energy* **2020**, *278*, 115410.
58. Miletic, M.; Pavic, I.; Pandzic, H.; Capuder, T. Day-ahead Electricity Price Forecasting Using LSTM Networks. In Proceedings of the 2022 7th International Conference on Smart and Sustainable Technologies (SpliTech), Split, Croatia, 5–8 July 2022. [[CrossRef](#)]
59. Demir, S.; Mincevic, K.; Kok, K.; Paterakis, N.G. Data augmentation for time series regression: Applying transformations, autoencoders and adversarial networks to electricity price forecasting. *Appl. Energy* **2021**, *304*, 117695. [[CrossRef](#)]
60. Hanif, M.; Shahzad, M.K.; Mehmood, V.; Saleem, I. EPFG: Electricity Price Forecasting with Enhanced GANS Neural Network. *IETE J. Res.* **2022**, 1–10. [[CrossRef](#)]
61. Brusaferrri, A.; Matteucci, M.; Portolani, P.; Vitali, A. Bayesian deep learning based method for probabilistic forecast of day-ahead electricity prices. *Appl. Energy* **2019**, *250*, 1158–1175.
62. van der Heijden, T.; Palensky, P.; van de Giesen, N.; Abraham, E. Day Ahead Market price scenario generation using a Combined Quantile Regression Deep Neural Network and a Non-parametric Bayesian Network. In Proceedings of the 2022 IEEE International Conference on Power Systems Technology (POWERCON), Kuala Lumpur, Malaysia, 12–14 September 2022; pp. 1–5.
63. Zhang, J.; Tan, Z.; Wei, Y. An adaptive hybrid model for short term electricity price forecasting. *Appl. Energy* **2020**, *258*, 114087.
64. Abdellatif, A.; Mubarak, H.; Ahmad, S.; Mekhilef, S.; Abdellatif, H.; Mokhlis, H.; Kanesan, J. Electricity Price Forecasting One Day Ahead by Employing Hybrid Deep Learning Model. In Proceedings of the 2023 IEEE IAS Global Conference on Renewable Energy and Hydrogen Technologies (GlobConHT), Male, Maldives, 11–12 March 2023. [[CrossRef](#)]
65. Zhang, F.; Fleyeh, H.; Bales, C. A hybrid model based on bidirectional long short-term memory neural network and Catboost for short-term electricity spot price forecasting. *J. Oper. Res. Soc.* **2022**, *73*, 301–325.

66. Meng, A.; Wang, P.; Zhai, G.; Zeng, C.; Chen, S.; Yang, X.; Yin, H. Electricity price forecasting with high penetration of renewable energy using attention-based LSTM network trained by crisscross optimization. *Energy* **2022**, *254*, 124212.
67. Peng, L.; Wang, L.; Xia, D.; Gao, Q. Effective energy consumption forecasting using empirical wavelet transform and long short-term memory. *Energy* **2022**, *238*, 121756.
68. Zhang, R.; Li, G.; Ma, Z. A deep learning based hybrid framework for day-ahead electricity price forecasting. *IEEE Access* **2020**, *8*, 143423–143436. [[CrossRef](#)]
69. Cantillo-Luna, S.; Moreno-Chuquen, R.; Celeita, D.; Anders, G. Deep and Machine Learning Models to Forecast Photovoltaic Power Generation. *Energies* **2023**, *16*, 4097.
70. Kazemi, S.M.; Goel, R.; Eghbali, S.; Ramanan, J.; Sahota, J.; Thakur, S.; Wu, S.; Smyth, C.; Poupart, P.; Brubaker, M. Time2vec: Learning a vector representation of time. *arXiv* **2019**, arXiv:1907.05321.
71. Shen, Y.; Jiang, X.; Wang, Y.; Jin, X.; Cheng, X. Dynamic relation extraction with a learnable temporal encoding method. In Proceedings of the 2020 IEEE International Conference on Knowledge Graph (ICKG), Nanjing, China, 9–11 August 2020; pp. 235–242.
72. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 1–11.
73. Su, H.; Peng, X.; Liu, H.; Quan, H.; Wu, K.; Chen, Z. Multi-Step-Ahead Electricity Price Forecasting Based on Temporal Graph Convolutional Network. *Mathematics* **2022**, *10*, 2366. [[CrossRef](#)]
74. XM Colombia. Portal de Variables del Mercado eléctrico Colombiano SINERGOX. Available online: <https://sinergox.xm.com.co/trpr/Paginas/Historicos/Historicos.aspx> (accessed on 15 March 2023).
75. Albahli, S.; Shiraz, M.; Ayub, N. Electricity price forecasting for cloud computing using an enhanced machine learning model. *IEEE Access* **2020**, *8*, 200971–200981. [[CrossRef](#)]
76. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
77. Chollet, F. Keras. 2015. Available online: <https://github.com/fchollet/keras> (accessed on 20 April 2023).
78. Gulli, A.; Pal, S. *Deep Learning with Keras*; Packt Publishing Ltd.: Birmingham, UK, 2017.
79. Seabold, S.; Perktold, J. statsmodels: Econometric and statistical modeling with python. In Proceedings of the 9th Python in Science Conference, Austin, TX, USA, 28 June–3 July 2010.
80. Bello-Rodríguez, S.P.; Beltrán-Ahumada, R.B. Caracterización y pronóstico del precio spot de la energía eléctrica en Colombia. *Rev. Maest. En Derecho Económico* **2010**, *6*, 293–316.
81. Marín, J.B.; Orozco, E.T.; Velilla, E. Forecasting electricity price in Colombia: A comparison between neural network, ARMA process and hybrid models. *Int. J. Energy Econ. Policy* **2018**, *8*, 97.
82. Urbano Buriticá, S.N.; González Pérez, L.F. Proyección de Corto Plazo para el Precio de Bolsa de Energía en el Mercado Colombiano. Universidad de los Andes. 2022. Available online: <http://hdl.handle.net/1992/63441> (accessed on 6 April 2023).
83. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
84. Cerliani, M. Keras-Hypetune. 2023. Available online: <https://github.com/cerlymarco/keras-hypetune> (accessed on 12 April 2023).

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.