International Journal of Energy Economics and Policy

ISSN: 2146-4553

available at http: www.econjournals.com

International Journal of Energy Economics and Policy, 2024, 14(5), 527-533.

Transforming Mexico's Electric Load Infrastructure: A Quantile Transformer Network Deep Learning Approach, 2019-2020

Wellcome Peujio Jiotsop-Foze, Adrián Hernández-del-Valle, Francisco Venegas-Martínez*

Instituto Politécnico Nacional, México. *Email: fvenegas1111@yahoo.com.mx

Received: 26 April 2024 **Accepted:** 31 July 2024 **DOI:** https://doi.org/10.32479/ijeep.16671

ABSTRACT

To transform Mexico's electric load infrastructure, accurate electric load forecasts are required, which are crucial to efficiently allocate resources, maintain system stability, and manage energy. The purpose of this study is to use the Quantile Transformer Network (QTN) as a novel approach for a deep learning framework for load forecasting, emphasizing its potential and practical consequences in enhancing the accuracy of load forecasting in real-world energy systems. Moreover, it is shown that QTN efficiently captures complex patterns, temporal relationships, and interconnections among factors that influence electric load. In this research it will be shown that the common Quantile Regression (QR) outperforms QTN in capturing dependencies in sequential data. The dataset utilized consists of past records of energy consumption in the Baja California System in Mexico. It includes several factors such as electricity demand, marginal prices, temporal characteristics, temperature-related variables, seasonal patterns, and holidays. Moreover, QTN is combined with the Rainbow Technique (RT) to manage categorical variables, resulting in the creation of a unified feature called category. RT examines the connections between descriptive phrases that reflect distinct combinations of categorical factors depending on the load values. Finally, several recommendations for promoting Mexico's Electric Load Infrastructure are provided.

Keywords: Electric Load Forecasting, Quantile Transformer Network, Temporal Dependencies, Non-linear Patterns, Rainbow Technique, Quantile Regression

JEL Classifications: Q41, Q42, C45, C53

1. INTRODUCTION

Following the decentralization of Mexico's power market in 2014, there has been a significant increase in the capacity of renewable energy. As of 2022, the total power generation capacity of environmentally friendly energy plants exceeded 31,000 MW. This demonstrates a steady increase in capacity each year and highlights the country's dedication to adopting sustainable energy practices (Dieck-Assad and Carbajal-Huerta, 2017).

Accurate electrical load projections are vital in the dynamic field of energy forecasting, as they optimize resource allocation, improve grid stability, and facilitate efficient energy management. To address this difficulty, we introduce an innovative method in this study, where we use a newly constructed Quantile Transformer Network (QTN) to predict the electrical load in the Mexican Baja California Sur (BCS) system.

The QTN approach is an advanced deep learning architecture specifically developed to model the intricate correlations present in energy data. The QTN architecture utilizes deep neural networks to apply their computational capabilities in capturing complex relationships, temporal dependencies, and intricate interconnections among the different components that influence electric demand. This study aims to evaluate the efficacy of QTN by comparing its performance with that of Quantile Regression (QR), a commonly utilized technique in several domains for forecasting purposes.

The Mexican BCS system is an exemplary subject for a case study because of its distinctive attributes and difficulties. Given the varied climatic conditions and increasing energy demand in the area, precise load forecasting is crucial for efficient resource planning, grid stability, and cost-effective power generation. One of the objectives of this investigation is to assess the effectiveness of QTN compared to QR in improving the accuracy of load forecasting and its practical consequences for the BCS system.

The main objective of this study is to provide a thorough examination of the QTN approach as it is used in electric load forecasting, emphasizing its superior performance in comparison to QR. Through a quantitative evaluation and comparison of the results produced from both techniques, our objective is to provide insights into the exceptional capabilities of the QTN framework and its potential to bring about a revolutionary change in the field of energy forecasting. This paper will provide several recommendations to enhance load forecasting techniques and facilitate the implementation of the QTN framework in practical energy systems. QTN uses deep learning to provide energy stakeholders with precise load estimates, allowing them to make well-informed decisions and maximize the operation of the Mexican BCS system.

The ways in which this research differs from the current literature are: (1) it extends the work of Jiotsop-Foze et al. (2024) by including QTN as a novel approach for a deep learning framework for load forecasting, (2) it shows that the ordinary QR outperforms QTN, (3) it provides a set of recommendation in energy policy to improve grid stability and efficient energy management.

This paper is organized in the following way: Section 2 offers an elaborate examination of the relevant literature and current forecasting approaches; section 3 presents a detailed explanation of the procedures for collecting data, preprocessing techniques, and the methodology used for load forecasting. This involves the utilization of QTN and QR; section 4 is dedicated to presenting and analyzing the outcomes of our trials, with a particular focus on highlighting the outstanding performance of the QTN model; section 5 provides a concise overview of the main findings analyzing their significance, provide a set of recommendation in energy policy, and propose prospective directions for future investigation.

2. A SHORT LITERATURE REVIEW

An extensive examination of the current body of literature on electric load forecasting algorithms uncovers a substantial amount of study and progress in the topic. This section presents a concise summary of the research in three important domains: QR, Transformer Networks (TN), and feature creation methods.

Transformer Neural Networks (TNN), which were first introduced in (2017) by Vaswani et al. are sophisticated deep learning models specifically built to handle sequential data processing tasks such as natural language processing and machine translation. Transformer models distinguish themselves from conventional recurrent neural networks (RNN) through the utilization of self-attention processes in place of recurrent connections. This allows them to effectively capture extensive dependencies across long distances. Transformers have demonstrated outstanding proficiency in diverse Natural Language Processing (NLP) tasks.

The essential elements of the transformers design, including the encoder, decoder, self-attention, and positional encoding, are crucial for its operation. The encoder transforms input sequences into encoded representations, while the decoder uses these representations to generate output sequences. The self-attention mechanism allows items in a sequence to selectively concentrate on other elements, therefore capturing interdependencies. Positional encoding provides positional context to the sequences.

Transformers have been integrated into well-known models such as Bidirectional Encoder Representations Transformers (BERT) and Generative Pre-trained Transformers (GPT). TNN excel in electric load forecasting by effectively identifying and understanding long-term relationships in time series data. Their self-attention processes exhibit exceptional efficacy in managing temporal interactions, resulting in remarkable outcomes. Multiple studies, such as those conducted by (Wang et al., 2022; Yao, et al., 2000; Khotanzad et al., 1995), and others, have examined the use of TNN in electric load forecasting. These studies have highlighted the networks' ability to identify patterns over time and improve the accuracy of predictions.

Moreover, the literature also explores the topic of electric demand forecasting by employing LSTM models and other architectures such deep residual networks (ResNet). These models demonstrate the utilization of LSTM and ResNet in identifying crucial dependencies for load forecasting jobs. The use of meteorological data enhances the precision of these models. (Choi et al., 2018 and Jiao et al., 2018) have conducted research that specifically emphasize the effectiveness of these models in the field of load forecasting.

QR is a statistical method that allows for the estimate of different parts of the conditional distribution of a dependent variable, based on a given set of predictors. Unlike traditional regression models that focus on estimating the average value of the dependent variable, QR provides a comprehensive understanding of variable relationships by modeling and analyzing numerous quantiles of the dependent variable. This approach has attracted interest because it is robust against extreme values and data with heavy tails, it can model different quantiles, and it can effectively uncover non-symmetrical correlations between predictors and the dependent variable (Huang, 2012).

Although there is limited literature specifically combining TNN with QR, there are research that explore the use of transformers for time series forecasting, which can be indirectly linked to QR. This research highlights the remarkable ability of transformers to identify relationships in sequential data. Salinas et al. (2019) utilized transformers to do time series forecasting, highlighting their ability to accurately identify long-term relationships and achieve exceptional outcomes. Furthermore, works conducted by Zhang et al. (2019) and He and Li (2018) have examined the capability of transformer-based architectures in predicting future events, emphasizing its effectiveness in capturing complex temporal patterns. Although the primary focus of these investigations is the application of transformers in time series forecasting, they offer valuable insights about the ability of transformers to effectively capture dependencies in sequential data.

Furthermore, there are investigations that explore the utilization of QR in combination with several neural network structures, outside from the research on transformers. Choi et al. (2018) proposed combining LSTM-based recurrent neural networks with QR to anticipate integrated volatility in financial domains. In a similar manner, Jiao et al. (2018) introduced a Quantile Regression Neural Network (QRNN) that combines feedforward Neural Networks (NN) with QR, providing fast and accurate predictions for time series data. In their study, Huuskonen et al. (1997) utilized QRNN to make predictions about the water solubility of organic compounds3. Likewise Meinshausen (2006) proposed the concept of Quantile Regression Forests (QRF), which combines ideas from stochastic gradient boosting and random forests to perform QR. Furthermore, Rodrigues et al. (2016) proposed a sophisticated approach to QR that utilizes deep neural networks to accurately estimate conditional quantiles.

These publications offer valuable insights into the specific approaches employed in QR challenges using Neural Networks (NN). Additional investigation is required to examine the precise hybrid structure that combines transformers and QR, and to evaluate its efficacy in capturing the conditional quantiles of the response variable.

Feature creation methods are essential for enhancing the accuracy of load forecasting by extracting pertinent information or modifying the original characteristics. These strategies have been extensively investigated in numerous investigations that specifically examine load predictions. In this sense, Hong et al. (2014) explored techniques for generating features in load forecasting. They demonstrated their capacity to detect seasonal patterns, trends, and other important qualities that are naturally present in load data. Their findings supported the use of feature engineering techniques, such as Fourier analysis and wavelet decomposition, to extract meaningful features from the load data. Likewise, Gao et al. (2016) conducted additional research on feature generation in load forecasting, emphasizing the crucial role these methods play in improving the accuracy of forecasts. Their study presented a methodology for selecting and extracting features using principal component analysis (PCA) and empirical mode decomposition (EMD). The aim was to identify the inherent structure and patterns in the load data.

On the other hand, Tsanas and Xifara (2016) conducted a thorough evaluation of different techniques for generating features in the specific domain of load forecasting. Their comparative investigation encompassed methodologies such as Fourier analysis, wavelet decomposition, and statistical characteristics, assessing their impact on forecast accuracy. The study emphasized the importance of carefully choosing and designing characteristics to contain relevant information in load data.

These studies highlight the importance of feature generation approaches in load forecasting, emphasizing their ability to discover important load data features such as seasonal patterns and trends. Utilizing these techniques can enhance the precision of load forecasting models, guaranteeing more reliable predictions.

Finally, the increase in renewable energy generation is attributed to various factors, such as legislative incentives, technological progress, social transition towards sustainable energy sources and reduction of CO_2 emissions. In particular, the attempt to reduce CO_2 emissions has led to several measures related to the increase of renewable energy sources, requiring more accurate prediction of electrical load to maintain stability in the electrical system in the transition towards green energies (Alotaibi et al., 2020; Ruiz-Alemán et al., 2023: Hussain et al., 2022; Mendoza-Rivera et al., 2023; Wang et al., (2019); Salazar-Núñez et al., 2020; Salazar-Núñez et al., 2022; Rathor and Saxena, 2020; Santillán-Salgado et al., 2020; Aslam et al., 2020; Aslam et al., 2021; and Valencia-Herrera et al., 2020).

3. NATURE OF DATA

This section provides a concise summary of the dataset used in the study, which was acquired from the Baja California Sur region. The collection consists of historical data on power use for each hour, covering the period from January 01, 2019, to September 30, 2020. The National Energy Control Center (CENACE, Spanish acronym of Centro Nacional de Control de la Energía) supplied the data for electric demand, specifically referred to as Total_Demand, and Local Marginal Price (PML).

Table 1 displays the details of the features employed in load forecasting, obtained from the BCS System Dataset. The table provides a comprehensive overview of the feature names, their corresponding data types, the quantity of values, and the distinct values associated with each feature.

The dataset comprises 15,336 observations, encompassing many elements pertaining to load forecasting. The variable Total_Demand represents the amount of electricity demanded in megawatt-hours (MWh) every hour. Another variable, Average_Pml, represents the mean local marginal price (PML) for the relevant hour.

The dataset includes various nominal variables, specifically Day_week, Day_of_month, Month, Year, and Hour_of_day. The variable "Day_week" denotes the specific day of the week, ranging from Monday to Sunday. "Day_of_month" indicates the numerical day of the month, ranging from 1 to 31. "Month"

Table 1: Features used in load forecasting from the BCS system dataset

Feature	Data type	Number of values	Unique values
Total_Demand	Real	15,336	
Average_Pml	Real	15,336	
Day_week	Nominal	7	7
Day_of_month	Nominal	31	31
Month	Nominal	12	12
Year	Nominal		2
Hour_of_day	Nominal	24	24
CDD	Real	15,336	
HDD	Real	15,336	
Season	Binary	2	2
Holiday	Binary	2	2

Own elaboration with electrical load data from the Baja California Sur (BCS) system signifies the specific month, ranging from January to December. "Year" represents the year, which can be either 2019 or 2020. Lastly, "Hour_of_day" denotes the specific hour of the day, ranging from 0 to 23.

In addition, the dataset contains two variables linked to temperature: CDD (Cooling Degree Days) and HDD (Heating Degree Days). CDD is the count of cooling degree days for the given hour, quantifying the amount of cooling needed during the summer months. HDD stands for heating degree days, which measures the amount of heating needed during winter months. These variables are frequently utilized in the energy sector to estimate energy demand and can be highly beneficial in energy planning, weather prediction, architectural design, and energy efficiency studies.

The dataset additionally contains two binary variables, Season and Holiday. The Season variable categorizes the hour as either part of the dry season (November to April) or the wet season (May to October). The Holiday variable indicates whether the day is a holiday or not. The dataset consists of 18 variables and 15,335 observations in total. The data was partitioned into a training set and a test/validation set, employing an 80/20 ratio. The test/validation set has 3067 observations. This dataset offers a comprehensive and extensive collection of data for doing load forecasting analysis. It includes a wide range of elements such as power demand, marginal pricing, temporal aspects, temperature-related variables, seasonality, and holiday effects.

4. RAINBOW TECHNIQUE (RT)

This section outlines the implementation of the Rainbow Technique (RT), which is used to handle categorical data and generate a unified feature known as category. The RT utilizes various categorical factors, including Day of the Week, Month, Season, Holiday, and Hour, to generate descriptive phrases that accurately depict their combinations. For example, a phrase could read "Tuesday January Low No one."

In order to use the RT, we generate these phrases by merging the categorical variables. There are a total of 2537 distinct sentences in this feature. We determine the exact occurrence of each sentence, and the maximum occurrence is 20. From the sentences that occur most frequently, we can distinguish four unique ones: "Sunday March Low Holiday one," "Thursday January Low Holiday one," "Wednesday July High Holiday one," and "Saturday August High Holiday one." Subsequently, we examine the correlation between these statements and the load. Upon investigation, we ascertain that the phrase with the lowest load has a frequency of 1 and corresponds to "Saturday April Low No one," with a load of 166.75 MW. This sentence is categorized as label 0. Conversely, the phrase "Tuesday August High Holiday seventeen" signifies the peak demand, achieving a load of 500.81 MW. The statement is repeated 8 times in the feature and is categorized as label 2536. In order to include the RT into our research, we generate a Python dictionary and combine it with the load data frame using the Pandas package. In addition, we investigate the correlation between load and various quantiles. Table 2 displays the load values for different quantiles, spanning from the 10^{th} to the 90^{th} percentiles. The load values corresponding to each quantile are as follows:

Table 2: Load by quantiles

Load (MWh)	Quantile (%)
209.38	10
252.34	30
286.53	50
347.92	70
419.13	90

Own elaboration with electrical load data from the BCS system

In addition, we examine the correlations between load and various characteristics. Table 3 displays correlation matrices for each quantile, focusing on the 10%, 30%, 50%, 70%, and 90% quantiles, as well as CDD, HDD, Average PML, and the Category feature obtained using the RT.

Based on the correlation matrices, it is evident that the Category feature created by the RT is the most significant feature for all quantiles, except for the 10% quantile. This emphasizes the significance of the Category characteristic in load forecasting analysis.

To summarize, the RT allows for the efficient modeling of categorical variables in a single feature, offering important insights into their correlation with the load. The category characteristic continuously exhibits significance across different quantiles, emphasizing its value for load forecasting analysis.

5. QUANTILE TRANSFORMER NETWORK (QTN)

The suggested design, named Transformer Model with Long Short-Term Memory (LSTM) and Multi-Head Attention synergistically integrates the benefits of LSTM layers with a Multi-Head Attention mechanism, therefore augmenting the model's capacity to represent information and its ability to generalize. Further enhancements are achieved by incorporating additional layers such as Dense, Dropout, and Layer Normalization.

In accordance with the conventional framework of a transformer

Table 3: Correlation matrices by quantile

Feature	Load, quantile 10%	Load, quantile 30%	Load, quantile 50%	Load, quantile 70%	Load, quantile 90%
CDD	0	1	10	56	18
HDD	−4	2	−6	−27	−2
Average PML	29	16	17	21	23
Category	11	28	28	58	60

Own elaboration with electrical load data from the Baja California Sur system model, the design substitutes the self-attention mechanism in the encoder with LSTM layers, proficiently capturing sequential dependencies. Subsequently, the output of the encoder undergoes a Multi-Head Attention layer, which allows the model to effectively capture interdependencies among the pieces in the sequence. Dense layers facilitate additional transformation. In order to get the intended result, the model compresses the processed output and feeds it through a decoder comprised of dense layers. This extensive metamorphosis guarantees meticulous analysis of the input sequence.

During the training process, a specialized loss function named "quantile_loss" is employed, and the model is optimized using the Adam optimizer. The training procedure encompasses crucial parameters, like batch size, number of epochs, validation split, and early stopping callbacks. The model is a subclass of tf.keras.Model and requires input parameters including input_shape, num_layers, d_model, num_heads, dff, and dropout_rate. The system comprises an encoder with LSTM layers, a Multi-Head Attention layer, dense_transform and dense1 layers for transformation, and dropout and layer normalization layers. The decoder consists of densely connected layers, and the call method performs the forward pass of the model. The architecture employs a bespoke model design that integrates LSTM layers, a Multi-Head Attention mechanism, and thick layers to effectively handle and forecast the incoming data. Below is an elaborate exposition of the architectural structure:

a. Encoder: The encoder is a sequential model implemented using the tf.keras library. The function Sequential (·) is called. The model comprises of two LSTM layers with d_model units and utilizes a ReLU activation function. The LSTM layers analyze the input sequence and generate sequences for each individual time step.
b. Transformer: The transformer layer is constructed utilizing layers. The function MultiHeadAttention (·) is called. The encoded sequence is used as both the query and the key. The num_heads option specifies the quantity of attention heads. The key_dim parameter determines the number of dimensions in the attention mechanism.
c. Dense transformation: Following the attention method, the output is fed into a dense layer (layers). Dense layer. The dense layer modifies the output of the attention mechanism. Dropout and Layer Normalization: Dropout layers (layers. Dropout(·)) are utilized to mitigate overfitting by applying them to the modified output. Layer normalization is a technique used to normalize the activations of individual layers in a neural network. The activations are normalized by applying Layer Normalization on the output.
d. Dense layers: The processed and standardized output is subsequently fed into a dense layer (layers.Dense(·)) with dff units and a ReLU activation function. Regularization is achieved by applying an additional dropout layer. The output undergoes layer normalization once again.
e. Flatten layer: The output is compressed using layers. Apply the Flatten (·) function to reformat it for the decoder.
f. Decoder: The decoder is a sequential model implemented using the tf.keras library. The function Sequential (·) is called. The architecture includes a compact layer with 512 units, a ReLU activation function, and L2 regularization. A dropout layer is utilized to implement regularization.

The number of units in the final dense layer is determined by the length of the quantiles array, which matches the desired output shape. The call (·) function does the forward pass of the model. The system processes the given inputs and carries out the required calculations. The input sequence is transmitted through the encoder, and the resulting output is utilized as both the query and key in the transformer layer. The output of the attention mechanism is subjected to a deep transformation layer. After applying dropout and layer normalization, the query tensor is reshaped to align with the dimensionality of X and then added to X. The output is subsequently propagated through the other levels until the ultimate output is achieved.

The model is constructed using a user-defined loss function named quantile_loss, which calculates the quantile loss based on the y_true and y_pred variables. The Adam optimizer is employed with a predetermined learning rate. The model is trained by utilizing the fit(·) function, which requires the training data (X_train and y_train) as input. Additional options encompass the batch size, epoch count, validation split, callbacks (e.g., early_stopping), and shuffle option. To summarize, the code establishes, builds, and trains a TransformerModel that adeptly integrates LSTM layers, Multi-Head Attention, and dense layers to efficiently interpret and forecast input data.

6. EXPERIMENTAL SETUP

This section outlines the experimental configuration and assessment criteria employed to measure the effectiveness of the QTN and conventional QR methods across five quantiles: 10%, 30%, 50%, 70%, and 90%. In order to carry out QR, we employed the statsmodels package, which enables the fitting of distinct QR models for each specified quantile and gives model summaries for examination of the results.

The processed data frame comprises a single label, Demand, and four features: CDD, HDD, Average PML, and Category. In order to divide the dataset into training and testing sets, we adhered to a ratio of 80/20. The implementation of QR was carried out using the Statsmodels library. The QR estimations may be found in the Appendix part of this research. The average quantile loss for each model on the test set is shown in the QR column of Table 4.

Table 4 presents a juxtaposition of quantile loss values for

Table 4: Quantile loss comparison

Quantile (%)	QTN	QR	Relative difference (%)
10	5.01	39.00	−87
30	8.92	25.01	−64
50	8.11	21.50	−62
70	8.95	24.28	−63
90	6.37	34.34	−81

Own elaboration with electrical load data from the BCS system. QTN: Quantile transformer network various quantiles, comparing the QTN and QR techniques. The comparative disparity between the two techniques is also computed.

The important findings on the quantiles derived from various approaches are as follows:

a. The 10th percentile was 5.01 for the QTN and 39.00 for the QR. The relative difference is −87%, signifying that the QTN yielded a markedly lower value in comparison to the QR.
b. The QTN obtained a value of 8.92 at the 30th percentile, while the QR provides a value of 25.01. The QTN approach yielded a value that was 64% lower than the value obtained using the QR, as indicated by the relative difference.
c. The 50th percentile (median) was 8.11 for the QTN, and 21.50 for the QR. The QTN approach yielded a value that was 62% lower than the value obtained using the QR, as indicated by the negative relative difference.
d. The QTN obtained a value of 8.95 at the 70th percentile, whereas the QR gives a value of 24.28. The QTN approach yielded a value that was 63% lower than the value obtained using the QR, as indicated by the negative relative difference.
e. The 90th percentile was 6.37 for the QTN, while the QR provides a score of 34.34. The relative disparity between the two approaches was −81%, suggesting that the QTN yielded a substantially lower value in comparison to the QR.

In summary, the findings indicate that the QTN consistently yielded lower values for each quantile in comparison to the QR. The relative discrepancies varied from −54% to −87%, suggesting a substantial discrepancy between the two techniques in predicting the quantiles.

7. CONCLUSION

Precise forecasts of electrical demand are crucial for optimizing resource distribution, ensuring grid stability, and facilitating efficient energy administration. The QTN presents a novel method for load forecasting. QTN surpasses conventional QR by capturing non-linear patterns, temporal dependencies, and interactions among elements that impact electric load.

Although QR offers a thorough comprehension of variable associations by modeling several quantiles of the response variable, there is a scarcity of research that integrates transformers with QR. Nevertheless, these investigations illustrate the exceptional efficacy of transformers in capturing interconnections among sequential data.

This study employs a dataset comprising historical energy usage statistics from the Baja California Sur region. The dataset includes factors such as power demand, marginal pricing, temporal aspects, temperature-related variables, seasonality, and holidays. The study uses the RT to manage categorical information and create a unified feature called category.

The RT combines category variables to generate descriptive phrases that depict their combinations. The technique effectively analyzes the links between sentences and the load by producing phrases that capture unique combinations of category factors in the dataset. Category labels are determined by load levels, and the Category feature produced by the RT shows a strong association with the load at various quantiles.

The paper intended to introduce and compare QTN, a deep learning architecture, with QR for load forecasting. Moreover, it demonstrates the efficacy of the RT in capturing the associations between categorical factors and the load. The results highlight the potential of QTN and its practical consequences, leading to enhanced accuracy in load forecasting for real-world energy systems.

REFERENCES

Alotaibi, I., Abido, M.A., Khalid, M., Savkin, A.V. (2020), A comprehensive review of recent advances in smart grids: A sustainable future with renewable energy resources. Energies, 13(23), 6269.

Aslam, A., Khalid, A., Javaid, N. (2020), Towards efficient energy management in smart grids considering microgrids with day-ahead energy forecasting. Electric Power Systems Research, 182, 106232.

Aslam, S., Herodotou, H., Mohsin, S.M., Javaid, N., Ashraf, N., Aslam, S. (2021), A survey on deep learning methods for power load and renewable energy forecasting in smart microgrids. Renewable and Sustainable Energy Reviews, 144(C), 110992.

Choi, H., Ryu, S., Kim, H. (2018), Short-Term Load Forecasting based on ResNet and LSTM. In: 2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm). p1-6.

Dieck-Assad, F. A., Carbajal-Huerta, E. (2017). Solar Energy: Peaceful Hope among Nations, Journal of Business Case Studies, 13(4), 99-108.

Gao, F., Tindula, G., Zhang, L. (2016), Feature generation in load forecasting: A comprehensive review. IEEE Transactions on Power Systems, 31(3), 2304-2313.

He, Y., Li, H. (2018), Probability density forecasting of wind power using quantile regression neural network and kernel density estimation. Energy Conversion and Management, 164, 374-384.

Hong, T., Pinson, P., Fan, S. (2014), Global energy forecasting competition 2012. International Journal of Forecasting, 30(2), 357-363.

Huang, A. (2012), Volatility forecasting by quantile regression. Applied Economics, 44, 423-433.

Hussain, A., Khan, Z.A., Hussain, T., Ullah, F.U.M., Rho, S., Baik, S.W. (2022), A hybrid deep learning-based network for photovoltaic power forecasting. Complexity, 2022, 7040601.

Huuskonen, J., Salo, M., Taskinen, J. (1997), Neural network modeling for estimation of the aqueous solubility of structurally related drugs. Journal of Pharmaceutical Sciences, 86(4), 450-454.

Jiao, R., Zhang, T., Jiang, Y., He, H. (2018), Short-term non-residential load forecasting based on multiple sequences LSTM recurrent neural network. IEEE Access, 6, 59438-59448.

Jiotsop-Foze, W.P., Hernández-del-Valle, A., Venegas-Martínez, F. (2024), Electrical load forecasting to plan the increase in renewable energy sources and electricity demand: A CNN-QR-RTCF and deep learning approach. International Journal of Energy Economics and Policy, 14(4), 186-194.

Khotanzad, A., Hwang, R., Abaye, A., Maratukulam, D. (1995), An adaptive modular artificial neural network hourly load forecaster and its implementation at electric utilities. IEEE Transactions on Power Systems, 10, 1716-1722.

Meinshausen, N. (2006), Quantile regression forests. Journal of Machine

Learning Research, 7, 983-999.
Mendoza-Rivera, R.J., García-Pérez, L.E., Venegas-Martínez, F. (2023), Renewable and Non-renewable energy consumption, CO_2 emissions, and responsible economic growth with environmental stability in North America. International Journal of Energy Economics and Policy, 13(4), 300-311.
Rathor, S., Saxena, D. (2020), Energy management system for smart grid: An overview and key issues. International Journal of Energy Research, 44(6), 4067-4109.
Rodrigues, T., Dortet-Bernadet, J., Fan, Y. (2016), Pyramid quantile regression. Journal of Computational and Graphical Statistics, 28, 732-746.
Ruiz-Alemán, M.E., Carbajal-De-Nova, C., Venegas-Martínez, F. (2023), On the nexus between economic growth and environmental degradation in 28 countries classified by income level: A panel data with an error-components model. International Journal of Energy Economics and Policy, 13(6), 523-536.
Salazar-Núñez, H.F., Venegas-Martínez, F., Lozano-Díez, J.A. (2022), Assessing the interdependence among renewable and non-renewable energies, economic growth, and CO_2 emissions in Mexico. Environment Development and Sustainability, 24(11), 12850-12866.
Salazar-Núñez, H.F., Venegas-Martínez, F., Tinoco-Zermeño, M.A. (2020), Impact of energy consumption and carbon dioxide emissions on economic growth: Cointegrated panel data in 79 countries grouped by income level. International Journal of Energy Economics and Policy, 10(2), 218-226.
Salinas, D., Flunkert, V., Gasthaus, J., Januschowski, T. (2019), DeepAR: Probabilistic forecasting with autoregressive recurrent networks. International Journal of Forecasting, 36(3), 1181-1191.
Santillán-Salgado, R.J., Valencia-Herrera, H., Venegas-Martínez, F. (2020), On the relations among CO_2 emissions, gross domestic product growth, energy consumption, electricity use, urbanization, and income inequality for a sample of 134 countries. International Journal of Energy Economics and Policy, 10(6), 195-207.
Tsanas, A., Xifara, A. (2016), Feature engineering in short-term load forecasting: An empirical study. Energy, 115(Part 2), 1621-1633.
Valencia-Herrera, H., Santillán-Salgado, R.J., Venegas-Martínez, F. (2020), On the interaction among economic growth, energy-electricity consumption, CO_2 emissions, and urbanization in Latin America. Revista Mexicana de Economía y Finanzas Nueva Época, 15(4), 745-767.
Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Polosukhin, I. (2017), Attention is all you need. In: Advances in Neural Information Processing Systems, NeurIPS Proceedings. p5998-6008.
Wang, C., Wang, Y., Ding, Z., Zheng, T., Hu, J., Zhang, K. (2022), A transformer-based method of multienergy load forecasting in integrated energy system. IEEE Transactions on Smart Grid, 13, 2703-2714.
Wang, H., Lei, Z., Zhang, X., Zhou, B., Peng, J. (2019), A review of deep learning for renewable energy forecasting. Energy Conversion and Management, 198, 111799.
Yao, S., Song, Y., Zhang, L., Cheng, X. (2000), Wavelet transform and neural networks for short-term electrical load forecasting. Energy Conversion and Management, 41, 1975-1988.
Zhang, W., Quan, H., Srinivasan, D. (2019), An improved quantile regression neural network for probabilistic load forecasting. IEEE Transactions on Smart Grid, 10, 4425-4434.