Energy Science & Engineering

ORIGINAL ARTICLE OPEN ACCESS

Forecasting Green Energy Production in Latin American Countries and Canada via Temporal Fusion Transformer

Muhammad Shoaib Saleem[1,2] | Javed Rashid[3,4] | Sajjad Ahmad[1] | Ali M. Al-Shaery[5] | Saad Althobaiti[6] | Muhammad Faheem[7,8]

[1]Department of Mathematics, University of Okara, Okara, Pakistan | [2]Center for Theoretical Physics, Khazar University, Baku, Azerbaijan | [3]MLC Research Lab, Okara, Okara, Pakistan | [4]Information Technology Services, University of Okara, Okara, Pakistan | [5]Department of Civil Engineering, College of Engineering and Architecture, Umm Al-Qura University, Makkah, Saudi Arabia | [6]Department of Science and Technology, University College Ranyah, Taif University, Ranyah, Saudi Arabia | [7]School of Technology and Innovations, University of Vaasa, Vaasa, Finland | [8]VTT Technical Research Center of Finland Ltd., Espoo, Finland

Correspondence: Muhammad Faheem (muhammad.faheem@uwasa.fi)

Received: 28 November 2024 | **Revised:** 6 January 2025 | **Accepted:** 13 January 2025

Keywords: deep autoregression (DeepAR) | deep learning (DL) | electricity prediction | gated recurrent units (GRUs) | green electrical production | long-term projections | temporal fusion transformer (TFT)

ABSTRACT

Forecasting green energy is crucial in diminishing dependence on fossil fuels and fostering sustainable development. However, it encounters notable challenges, such as variable demand, restricted data availability, the integration of various datasets, and the necessity for precise long-term projections. This study thoughtfully examines these issues using the temporal fusion transformer (TFT) model to project green energy production across five Latin American nations (Argentina, Brazil, Chile, Colombia, and Mexico) and Canada, drawing on data from 1965 to 2023. The performance of the proposed TFT is more authentic as compared with the gated recurrent unit (GRU), the long short-term memory (LSTM), deep autoregression (DeepAR), and the meta graph-based convolutional recurrent network (MegaCRN). The TFT has a mean square error (MSE) of 0.0003, root mean square error (RMSE) of 0.0173, mean absolute error (MAE) of 0.0112 and mean absolute percentage error (MAPE) of 1.76%. From the preceding results, it is clear that the proposed TFT model can identify dynamic energy patterns that will contribute towards achieving sustainable development goals by the end of 2040.

1 | Introduction

Green energy is renewable energy or energy from inexhaustible natural sources like solar, wind, geothermal, and hydropower. These apply to greenhouse gas (GHG) and climate change, ensuring a steady supply of primary energy in the future. Electricity generation, transport and heating are among its uses as more people search for power sources other than oil [1].

Worldwide, the generation of green energy sources has seen substantial development in the past few years, but there are still some bottlenecks. Based on the International Renewable Energy Agency (IRENA), the report on renewable energy supply showed that in 2022, the renewable energy sector produced almost 30% of the global power generation, with wind and solar electricity generation at the helm [2]. However, they have not made equal progress in guaranteeing all regions' transition to green energy. Whereas some countries are ambitious in their commitment and invest significantly in renewable energy, many rely on fossil fuels because of economic, political, or even technological realities that hamper the world's ability to meet sustainability goals effectively [3]. These challenges are thus well tackled by concerted efforts.

Modern countries such as Argentina, Brazil, and Canada help us with significant progress in generating renewable energy in

America. Brazil, the largest bioenergy and hydropower producer, has created a high standard where renewable energy sources comprise 80% of the total electricity production [4]. Hydropower is Canada's primary electricity source, and it generates almost 60% of its electricity from this source, making it one of the cleanest electricity producers globally [5]. Argentina plans for renewables to contribute 20% of the country's energy mix by 2025, with significant investments in wind and solar energy [6]. These countries are the hope as the world transforms to green energy.

Likewise, many countries in Latin America, including Chile, Colombia, and Mexico, have shown immense progress in generating green energy. The government has steadily turned to solar power sources, especially in the Atacama Desert, which is among the places that can benefit from maximum solar power [7]. Although Colombia depends entirely on hydropower, cooperative investments in wind energy are significant [8]. However, recent policy issues have provided support for the usage of fossil fuels. Mexico is gradually increasing wind and solar power installations due to the commitments made by private investors [9]. The countries which are studied in this article are shown in Figure 1.

However, problems with the world's energy include pollution, population growth, and more people living in towns. Extra power is needed when more people live in cities. It makes carbon emissions and damage to the earth worse. The world must quickly switch to green energy sources to keep it safe and help the economy grow [10].

Modern energy forecasting approaches are based on applying advanced machine learning (ML) and deep learning (DL) methodologies. Whereas ML operates with data organized in a specific model to make advanced predictions, DL applies artificial neural networks with unorganized data sets with much better results. Recurrent neural networks (RNNs) and long short-term memory (LSTM) networks are specifically efficient in modeling time sequences in the energy datasets and frequently produce better results than traditional machine learning algorithms [11]. Other models similar to gated recurrent unit (GRU) and temporal fusion transformer (TFT) are being newly considered for intricate time series forecasting [12]. Deep autoregression (DeepAR), which is noted for its probabilistic forecasting strengths, has produced a very accurate power map in multivariate time series learning [13]. Meanwhile, meta graph-based convolutional recurrent network (MegaCRN) incorporates graph-based learning and recurrent structures and achieves state-of-the-art performance in energy prediction over space [14]. These models have been shown to work in energy forecasting and other related fields like climate change modeling, smart grid control and energy conservation [15].

Solar, wind, and hydro-energies are green energy forecasting using deep learning models. Solar energy generation can be described using the photovoltaic power equation:

$$P_s(t) = G(t) \cdot A \cdot \eta, \tag{1}$$

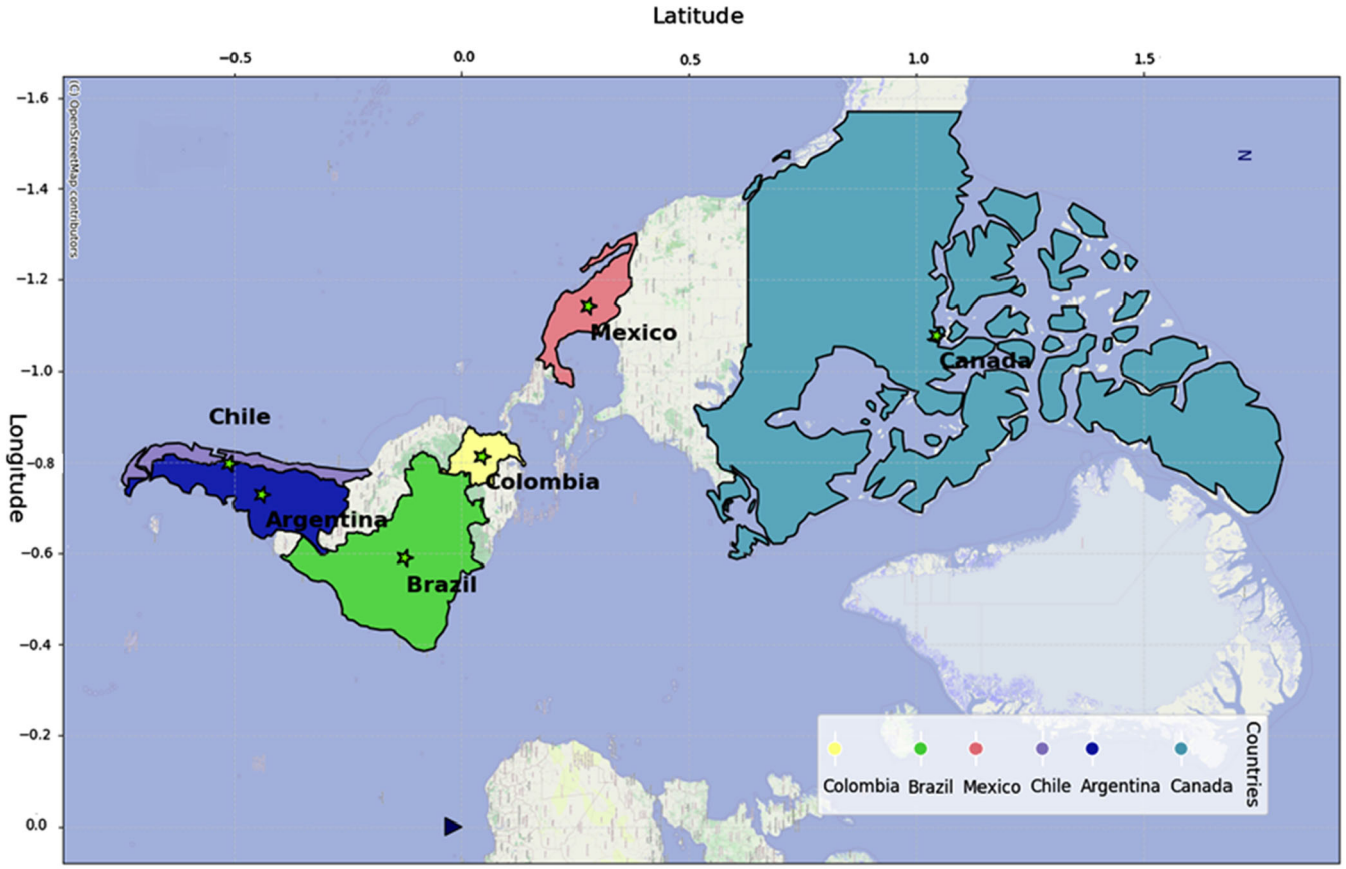

FIGURE 1 | Map of study area of Latin American countries and Canada.

where $P_s(t)$ is the power output, $G(t)$ is the solar irradiance, A is the area of the solar panel and η is the efficiency of the solar cells. In this regard, Awais et al. [16] have focused primarily on short-term solar forecasting using LSTM with a spatio-temporal attention mechanism; gaps remain in long-term prediction methodologies.

Wind energy generation can be modeled with the wind power equation:

$$P_w(t) = \frac{1}{2} \cdot \rho \cdot A \cdot v^3(t) \cdot C_{p,} \tag{2}$$

where $P_w(t)$ represents the power output, ρ is the air density, A is the swept area of the wind turbine, $v(t)$ is wind speed and C_p is the power coefficient. In this regard, Lv et al. [17] have been made in short-term wind forecasting using the VMD-GRU-GSRCV model, but there is still a notable lack of research focused on long-term forecasting.

Hydropower generation relies on the potential energy equation:

$$P_h(t) = \eta \cdot \rho \cdot g \cdot h \cdot Q(t), \tag{3}$$

where $P_h(t)$ is the power output, η is turbine efficiency, ρ is water density, $g$is the gravitational constant, h is water height and $Q(t)$ is flow rate. In this regard, Kaewarsa et al. [18] forecast short-term hydroelectric output using bidirectional LSTM, but long-term forecasting remains underexplored, and the model novelty is limited.

This study will fill the following gaps identified from the existing literature on the forecasts of green electricity output. First, there is a lack of emphasis on the accuracy of long-term forecasts, as such problems demand that the prerequisites for prediction be found by 2040. Although various research works have been done on specific sources such as solar, wind, and hydro energy, there is a lack of adequate forecast encompassing total green energy generation. Furthermore, many models under consideration do not offer great innovativeness in minimizing errors concerning sustainable long-term forecasts. The problem is even worsened by using outdated datasets, which significantly limits the creation of accurate predictions. Hopefully, this study can devise novel models and utilize updated datasets to refine approximation to fill the research gaps identified in this paper. In this regard, the following are the primary contributions of this study:

- This article proposes the TFT model to enhance the accuracy of sustainable electricity forecasts. Thus, using the model's superior ability will enable the prediction of more accurate amounts of green electricity generated.
- The study improves the ability to forecast long-term for green electricity, reduces errors in forecasting, and reduces the risk of making wrong decisions.
- This study employs the updated data set up to 2023 and contributes by projecting total green electricity production up to 2040.

The outline for the remaining portion of the article is as follows: Section 2 of this article contains a literature review. Section 3 presents materials, data set splitting, data processing, methodology, algorithm, experimental design, and assessment metrics. Section 4 provides the model comparison and the contextual evaluation of earlier studies in addition to their results and discussions. Section 5 is for the research's conclusion and contains future plans.

2 | Literature Review

Emerging methods in deep learning are crucial to generating green power, which is the enduring clean electricity for contemporary society. These algorithms are very good at forecasting how many megawatts of green energy will be produced in the future. Hence, they make the power system more efficient, advance its resourcing, and facilitate the identification of demand trends. They have achieved these goals due to their ability to learn from complicated data and discover other hidden trends. RNNs are the most popular deep learning algorithms in the world today. Application areas include midterm prognosis of future electricity market situations, demand side management, error detection, load benchmarking, and green energy prognosis. This means the new green electricity generation assessment techniques are more precise and reveal usage energy patterns and factors [19].

This section intends to conduct a comprehensive systematic literature review of the deep-learning methods for predicting and generating green energy, mainly targeting Latin American countries and Canada's research in the field. Over the last few years, most countries have understood that they must transition from reliance on conventional energy sources that harm the environment to green energy sources.

In the context of solar power generation prediction, Al-Ali et al. [20] proved that hybrid convolutional neural network (CNN)–LSTM–transformers strategic models for forecasting far outperform traditional models. However, these models are confined to only one type of energy input and hence have a somewhat limited usage. Anu Shalini et al. [21] used solar and wind power as different energy inputs to predict energy output using a based Bi-LSTM model in a new work of particular interest. Although it is a step forward, it only works within different specialized methods and can hardly be applied to various datasets. The following study by Mystakidis et al. [22] aimed to use machine learning to manage renewable power generation and output prediction to assist distribution system operators in handling demand response. However, their research does not thoroughly discuss the issues of using multiple renewable sources to provide effective forecasting [22]. Different source integration was done using artificial intelligence by Nikulins et al. [23]. However, their work is more specific to small-scale systems, which is impractical for general energy prediction.

The study by Pasandideh et al. [24] focused on wind power and population diversity issues, which can have actionable insights into renewable energy forecasting. However, their high reliability of finding is relatively narrow in evaluating the practicality of wind energy, and it ignores the complex system of generating power and distributing it through various sectors

[24]. The SSA–CNN–LSTM established by Venkateswaran and Cho [25] became a powerful combined deep learning model that can make forecasts promptly for utilizing greenhouses. Hourly green power production forecasting is one of the most relevant aspects in the cities. The bi-LSTM model has shown a minimum mean absolute percentage error (MAPE) of 7.7256% by applying the feature set, including min-max normalization. However, this can distort the temporal variability of energy consumption profiles despite using normalization techniques to address this issue [26].

According to Marques et al. [27], the use of deep learning approaches in solar energy forecasting is gradually emerging. Their work shows that LSTM-GRU is an effective model for looking for sustainable energy in the Amazon Basin. However, geographic limitations to their work limit their capacity to adequately capture the various aspects of energy forecasting across multiple geographic locations.

However, this body of research also has gaps and limitations in the present interpretation. The existing models focus on one type of renewable energy or limited-time forecasts without considering the complexities and requirements of multiple energy systems for extended forecasts. For example, Abubakar et al. [28] applied LSTM to increase the accuracy of solar power production forecasts in intelligent grids by up to 97%. However, they fail to align the different renewable power types with one another, so their findings can hardly be used for overall energy forecasting [28]. Alike, Unsal et al. [29] found that LightGBM is the best for predicting renewable energy in Turkey. Still, their analysis does not capture a more comprehensive array of energy outputs or ease the trade-offs for cross-border predictions [29]. Hasan et al. [30] used the KNN and Decision Tree (DT) working models to anticipate the accurate source-wise power generation for countries such as Australia, the UK, and the USA, enriched with high classifier efficiency and low mean absolute error (MAE). Yan et al. [31] proposed the GP-Ensemble comprised of CNNs, GRUs, and FNNs to boost renewable energy predictions in ten Asian countries over prior methods with the MSE (0.0631), the MAE (0.1754), and the RMSE (0.2383), which can support energy administration and policy.

Because of the flexibility of the TFT model, many studies have been conducted to investigate it. Wu et al. [32] presented an interpretable wind speed prediction model based on variational mode decomposition, adaptive differential evolution, and TFT for higher accuracy and temporal analysis. Wu and Wang [33] proposed a two-stage decomposition of JADE optimization and TFT to enhance the accuracy and stability of the wind speed forecast. Lim et al. [34] introduced the TFT as an attention-based architecture for multi-horizon forecasting with interpretability in temporal patterns and outperformed benchmarks on various real-world datasets. Subsequently, Wu et al. [35] developed a wind speed forecasting model based on two-stage decomposition, meteorological feature engineering, and optimized TFT that exhibited outstanding forecasting accuracy and can be applied to decision-making. Due to its efficient features, the TFT model can also predict green energy efficiently.

This study seeks to fill gaps in green power production forecast studies. First, long-term forecast accuracy, particularly for 2040 estimates, is neglected. Previous studies have examined solar, wind, and hydro energy, but complete estimates of overall green energy generation are scarce. Numerous models lack innovation, especially in error reduction for long-term forecasts. The use of old datasets significantly reduces prediction accuracy. This work develops new models and uses updated datasets to solve essential research gaps and enhance accuracy. Table 1 lists the significant literature.

3 | Material and Methods

The data gathering and analysis process is described here. It analyzes green energy generation using the TFT model. Our study methods are explained in the following analysis. Figure 2 shows the study's research flow.

3.1 | Data Set

This data set examines the growth of green electricity sources over time, according to five Latin American countries and Canada's total green electricity production (GEP). The data span for this study is 1965–2023. It offers a detailed picture of renewable energy production using data from six countries. The countries involved are—Mexico, Argentina, Brazil, Canada, Chile, and Colombia. Solar, hydro, wind and other renewable energy sources and bioenergy provided TWh in the data set containing the year 1965–2021 data set fetched from the Kaggle data source [36]. Our work integrated all types of renewable resources in a country into a single gigawatt hour (GWh) while updating the data set. Since most of our data is available in the month-wise format, the files are being converted to the year-wise from 2022 to 2023 as available at [37]. As can be observed and as depicted in Table 2 below, which contains our inputs, this update contiguously depicts renewable power production in six countries. By such a revision, the comparative proportions of each nation in the context of the renewable energy market could be adequately evaluated. We want to avoid the notion that renewable energy and other sustainability endeavors across the selected countries are lagging.

Figure 3 highlights the correlation matrix of clean energy production between distinct Latin American countries; this is important as it provides an understanding of the analysis relating to trends and connections of green energy production in the selected countries. The high positive correlation coefficients, especially between Brazil, Chile, and Colombia, imply that these nations follow parallel growth in renewable energy, most probably due to similar environmental climate, political stability and energy policies. This regional aggregation signifies that these nations may possess similar renewable energy resource endowments and use similar technologies or strategies correctly. At the same time, somewhat lower coefficients with Argentina and Canada, as well as other countries, show that Mexico's renewable energy indicators have slightly different trends, most likely because of the difference in resources, policies, or the speed of development. It is significant because, although the matrix offers insightful information about regional patterns and connections, the forecasting models only use data from

TABLE 1 | Literature work.

[Ref], Year	Methodology	Country	Accuracy/error	Cons
[20], 2023	CNN-LSTM-transformer	Finland	MAE (0.393), RMSE (0.344)	The research is exclusively centered on solar energy.
[21], 2023	CNN based BILSTM	India	Solar Power: MSE (0.0884), MAE (0.0219), R^2 (0.993), Wind Power: MSE (2.1432), MAE (1.0125), R^2 (0.992)	The exclusive attention is on wind energy and solar power.
[22], 2023	Ensemble approach	Cyprus	RMSE (1.9993), MAE (0.8306), MSE (3.9972), R^2 (0.8913)	The study is entirely focused on solar energy.
[23], 2024	FCN, CNN	Latvia	FCN: MSE (1.77), CNN: MSE (1.36)	The main focus is on wind and solar energy.
[24], 2024	LSTM	—	RMSE (0.47)	The research is solely concentrated on steam energy.
[25], 2024	SSA-CNN-LSTM	South Korea	1 h ahead: MAE (0.1202),2 h ahead: MAE (0.1400), Day ahead: MAE (0.1774)	The study is entirely focused on solar energy.
[26], 2024	Bi-LSTM	—	MAPE (7.7256%), RMSE (0.12346), R^2 (0.6151)	There is limited model novelty.
[27], 2024	MLP, LSTM2, LSTM-GRU	Brazil	MLP: MAE (0.61), MAPE (19.5), RMSE: (1.05) LSTM2: RMSE (1.05), MAE (0.81), MAPE (22.77), LSTM-GRU: RMSE (0.75), MAE (0.89), MAPE (19.2)	The study is entirely focused on solar energy systems.
[28], 2024	RNN, LSTM, GRU	—	LSTM: MAE (41.0046), RMSE (65.892)	The main focus is on solar energy systems.
[29], 2024	LightGBM	Turkey	RMSE (1.7479), MAE (1.1072)	There is limited model novelty.
[30], 2024	KNN and XGBoost	USA, Australia, France, Germany, UK	MAE (0.578, 0.659, 1.383, 0.738, 1.02)	Outdated data sets are used, and there is limited model novelty.
[31], 2024	GP-Ensemble	Asian countries	RMSE (0.2383), MAE (0.1754), MSE (0.0631)	There is limited model novelty.

Abbreviations: CNN, convolutional neural network; GRU, gated recurrent unit; LSTM, long short-term memory network; MAE, mean absolute error; MLP, multilayered perceptron; RMSE, root mean square error.

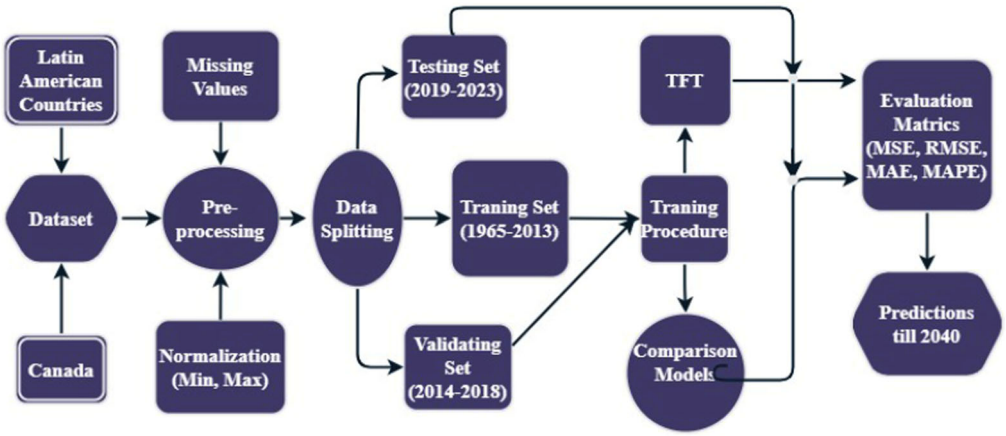

FIGURE 2 | Research flow. MAE, mean absolute error; MAPE, mean absolute percentage error; RMSE, root mean square error; TFT, temporal fusion transformer.

TABLE 2 | Data set of GEP.

Year	GEP in Argentina	GEP in Brazil	GEP in Canada	GEP in Chile	GEP in Colombia	GEP in Mexico
1965	1225.1163	25,515	118,088	3954	3543.9494	8863
1966	1240.9302	27,906	131,000	4168	3937.2194	10,118
1967	1270.6977	29,189	133,924	4255	4389.2536	11,017
—	—	—	—	—	—	—
—	—	—	—	—	—	—
2023	37,579.19	50,7876	427,375	44,389	63,060	102,088

Abbreviation: GEP, green electricity production.

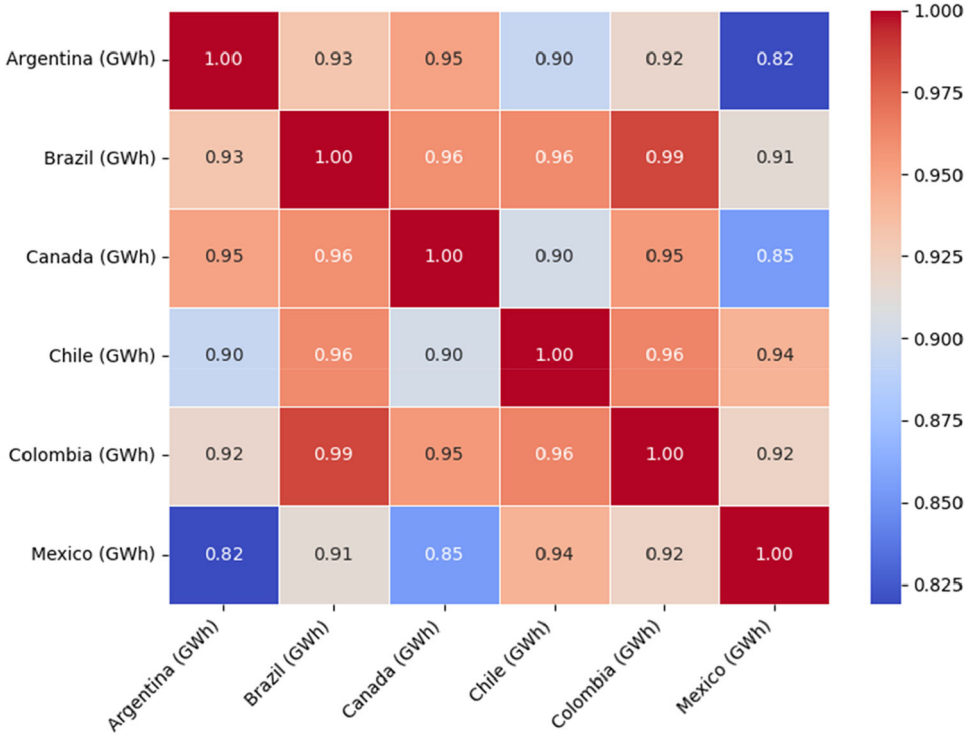

FIGURE 3 | Correlation matrix.

individual nations to estimate the generation of green energy. The correlation matrix, though not used in the actual modeling of the forecast, can be handy for insights into the other regions and can help policymakers in their decision-making process if they are to do so. For example, countries with similar trends might benefit from working together as a region, while countries with different trends might need national policies more tailored to their needs. In the end, this grid gives us a valuable picture of the green energy situation in these countries. It shows how important it is to consider regional factors when planning and predicting energy use.

3.1.1 | Data Preprocessing

The model is convenient to learn from as it has been processed from the raw data to get a clean version. This approach is simplified by arranging the data in a time series format, with each year as the first key. The preprocessing step is easily accessible and varies annually. Here, we start by cleaning the data, an essential and crucial step performed mainly by our data science and machine learning engineers. It is essential to guarantee that the data set is secure and to prepare the data correctly to train the model during this step. Data analysis for errors, irregularities, and missing values can be rigorous. We rigorously analyzed our annual electricity generation data set (in GWh) for Argentina, Brazil, Canada, Chile, Colombia, and Mexico from 1965 to 2023 to ensure accuracy and reliability. The missing values are treated based on the potential they will have on the integrity of the data set. For numerical data, such as for the GWh measurements, we used median imputation, replacing missing entries with the median value of the respective country and year so that temporal trends are consistent. Statistical methods used in time series data, such as Z-scores and the interquartile range (IQR), are used to identify outliers. Carefully examined outliers detected are considered to be due to actual fluctuations in electricity generation in data entry. The outliers from errors are corrected to the nearest plausible value by historical pattern, and legitimate outliers are retained to maintain the integrity of the data set.

Normalization is performed so that all features are equal in importance in the learning process so that no feature is scaled to dominate. It is essential to check that the number of columns in the data set using the cleaning method remains the same, and this is achieved using the Min-Max normalization described by Equation (4). This implies that no part of the data can have a large effect on the conclusion when the data is normalized to the range [0–1] [38]. Measures are taken to increase the signal strength of all characteristics, making sure each contributes equally to the model's learning process. It also proves very helpful in enhancing the DL model for data-driven learning since it brings into resonance all scales of the qualities contained in the given data set. The normalized information can also be seen in Table 3.

$$h' = \frac{h - h_{\min}}{h_{\max} - h_{\min}}. \tag{4}$$

The h is a set of entries, the h' is normalize entry, $h_{\max}$ is a maximum entry and $h_{\min}$ is a maximum entry.

3.1.2 | Data Splitting

Data splitting is required in deep learning model training and testing to ensure it can spread and not over-optimize. We may use this approach to evaluate the model efficiency in the different data sets. It has been familiar with training-to-testing splits to identify the ratio of the best RNN model for the prediction [39]. The selection of the best-split ratio depends on various factors, such as the structure of a particular model, the type of data used, and the model's prediction period, also known as the horizon. The country data we analyzed in this study is from Latin American countries and Canada. Data from these six countries are divided into each data set as training, validation, and testing datasets. The training data set comprises records from 1965 to 2013 because the background information required to train a model is essential. Validation data set containing data for 2014–2018 used for tuning the coefficients to reduce overfit; test records involving power generation from 2019 to 2023 to determine model functionality with new, unseen datasets.

3.2 | Proposed Methodology

Numerous deep learning models exist, although they lack any prediction accuracy. In this research, we employ the TFT model, a newly proposed method, to enhance the precision of our predictions.

TABLE 3 | Normalizing data set.

Year	GEP in Argentina	GEP in Brazil	GEP in Canada	GEP in Chile	GEP in Colombia	GEP in Mexico
1965	0	0	0	0.0095	0	0
1966	0.0004	0.0048	0.0407	0.0148	0.0066	0.0135
1967	0.0012	0.0074	0.0499	0.0169	0.0142	0.0231
—	—	—	—	—	—	—
—	—	—	—	—	—	—
2023	0.9709	0.9755	0.9746	1	1	1

Abbreviation: GEP, green electricity production.

3.2.1 | TFT

The proposed model, TFT, is a deep learning model specializing in multivariate time series forecasting, which inherently incorporates temporal dependencies and co-dependencies of various input features. They identified that the significant application of TFT is to address long-term forecasting issues, using mechanisms such as attention and gating to promote the interpretability of the model and the flexibility that enables it to address the complexity of datasets. This approach is one of the hallmarks of TFT as the method allows for the integration of fixed or structural covariates, features closely relating to data's short- and long-term dependencies, and time-varying inputs that are either known or unknown. It makes it possible for the model to incorporate both short-term and long-term dependency in the data [40].

TFT is made up of several components. A significant constituent in our work is the LSTM-based encoder-decoder architecture that captures the temporal features of the input time series. The encoder transforms the previous inputs, and the decoder uses these learned representations to generate future values and the current future inputs. Another equally important module is the variable selection network (VSN), whose role is to select the relevant static and time-varying inputs at any given time step in real time. It aids the model to gain concentration on the appropriate features, thus enhancing the understanding of its performance and operation. TFT uses a gating mechanism (gating of residual network [GRN]), incorporated into the model at every stage to minimize little extra detail. Residual connections allow the GRNs to manage the convolutions between input and output while increasing interpretability. The applicability of temporal self-attention is also utilized in the long-range dependencies and relations in the different time steps. The self-attention mechanism is masked so that at each step, only information from the past is used to make predictions, as using information from the future would be advantageous. A special kind of encoder used in the TFTs is called static covariate encoders (SCEs), which deals with the time-invariant features by supplementing the dynamically encoded data with static covariate data. These covariates are then processed through the GRN and static enrichment layers to shape effective temporal representations for forecasting. At the output stage, TFT uses a position-wise feed-forward network (FNN) and dense layers, accompanied by an attention-based output mechanism to output the forecasted values. The masked, multi-head attention layer gives an interpretable view of which previous time steps and features were critical in making the model predictions, as shown in Figure 4.

Mathematically, for the input sequence X_t, let the encoder process the input as:

$$h_t = \text{LSTM}(X_t), \tag{5}$$

where h_t represents the hidden states at each time step t.

The attention mechanism in the transformer encoder is provided by:

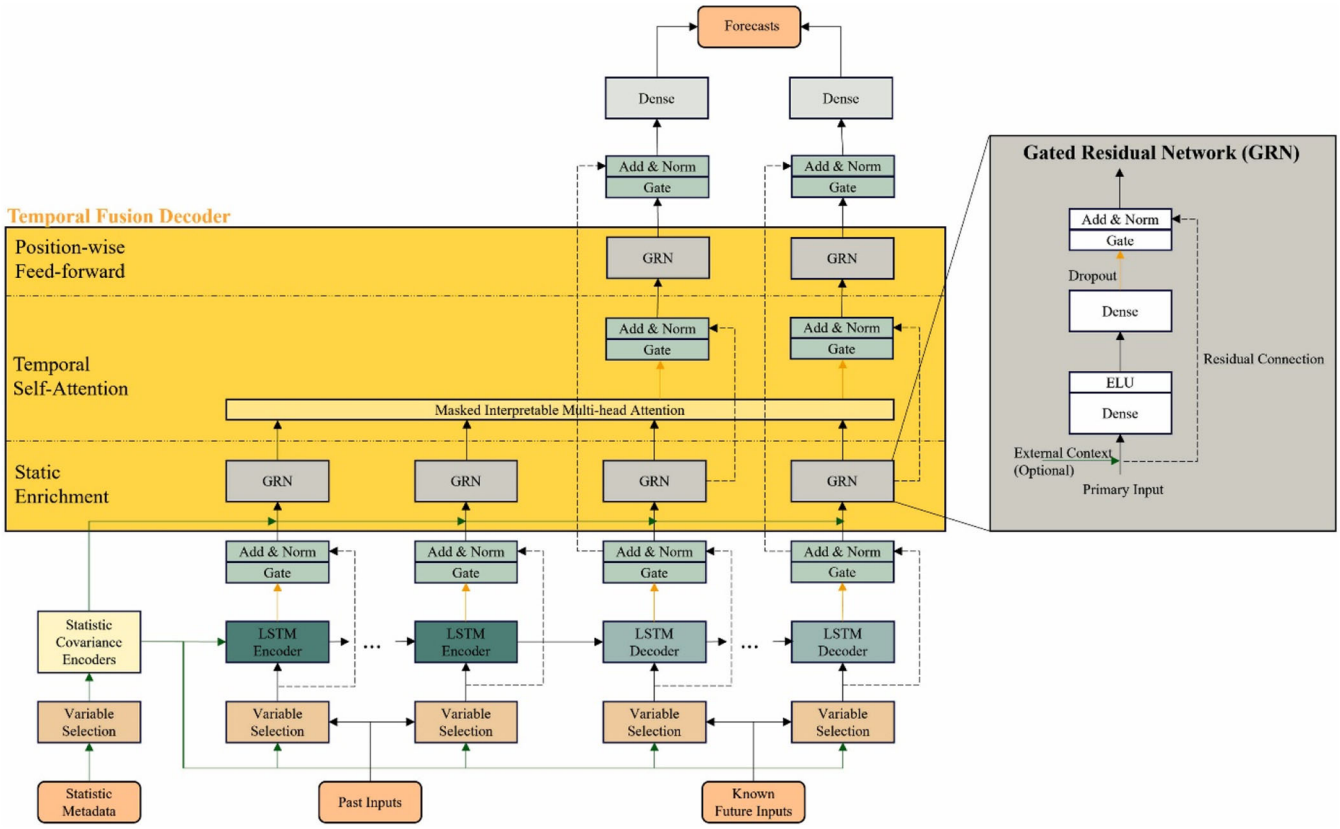

FIGURE 4 | TFT architecture. LSTM, long short-term memory network; TFT, temporal fusion transformer.

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V, \tag{6}$$

where Q, K, and V are the query, key matrices, and value matrix, respectively, whereas d_k is the dimension for key vectors.

Hence, the multi-head attention [41] comes to be:

$$\text{MultiHead}\,(Q, K, V) = \text{Concat}\quad(\text{head}_1, ..., \text{head}_h)W^o, \tag{7}$$

and,

$$\text{head}_j = \text{Attention}\left(QW_j^Q, KW_j^K, VW_j^V\right), \tag{8}$$

where w_j^Q, w_j^K, and w_j^V are the linear transformation matrices used in each attention head. Here, FFN is:

$$\text{FFN}(x) = \max(0, xW_1 + b_1)W_2 + b_2, \tag{9}$$

where x is the input, W_1, W_2 are weight matrices, and b_1, b_2 are bias vectors. Following this, the residual connection is: where b_1, b_2 are bias vectors, x is the input, and W_1, W_2 are weight matrices. The residual connection comes next:

$$\text{Residual}(x) = \text{LayerNorm}(x + FFN\,(x)), \tag{10}$$

where LayerNorm scales the input over the final dimension.

The TFT's fusion component appears after the encoding phase. Additionally, it combines concealed representations from several timesteps to model spatial dependency and trends. This fusion process allows the model to comprehend how previous time series values impact future outcomes. These fused representations' output is provided to a decoder to forecast the following values of the given time series [42]. TF layer is given by:

$$\text{TF}(x) = \text{LayerNorm}(x + \text{FFN}(x)), \tag{11}$$

which resembles the residual connection used in the transformer encoder layer. The temporal fusion gating is:

$$\text{Gating}(x, g) = \text{sigmoid}\,(x) \odot g, \tag{12}$$

where the gating vector is denoted by g. The temporal fusion equation may, therefore, be explained as follows:

$$\text{TF\_eq}(x_1, ..., x_t, g) = \text{TF}(x_1) \oplus \cdots \oplus \text{TF}\,(x_t) \oplus g, \tag{13}$$

where $\oplus$ represents concatenation.

Lastly, the encoder (decoder) output is several hidden states that represent temporal and contextual information of an input sequence. The output layer can be described as:

$$\text{Output}\,(x) = xW_{\text{out}} + b_{\text{out}}\,, \tag{14}$$

where W_{out} and b_{out} are the weight matrix and bias vector of the output layer, respectively.

The gating mechanism is embedded to detect the influence of each input variable at various time steps to make the TFT more interpretable. It enables users to differentiate the importance of multiple elements that may be used in the forecast development process. TFT offers the advantage of processing multivariate data, modeling long-term dependencies, and offering interpretability, which makes this model plausible in time series analysis and prediction [43].

GRN is one of the updates of the residual gated mechanism in the network that enables the network to skip the nonlinear transformation of an input r in the specific context s. Gating in GRNs primarily takes place to regulate the input and the residual connection, enhance model capacity or density, and avert the difficulty of vanishing gradients on deeper networks. The residual block is given by:

$$\rho_1 = \text{ELU}\,(\text{W}_1\text{r} + \text{W}_2\text{s} + \text{b}_1), \tag{15}$$

$$\rho_2 = \text{W}_2\rho_1 + b_2, \tag{16}$$

$$\text{GRN}\,(\text{r}, \text{s}) = \text{LayerNorm}\,(r + \text{GRU}\,(\rho_2)), \tag{17}$$

where ELU is the exponential linear unit, the term GLU stands for the gated linear unit, which has the added benefit of keeping the unnecessary circuits of the GRN [44]. Consider GRN's output τ, then GLU transformation is defined by:

$$\text{GLU}(\tau) = \sigma(\text{W}_4\tau + b_4) \odot (\text{W}_5\tau + b_5). \tag{18}$$

The TFT model is chosen for its ability to work with multivariate time series, which is necessary when predicting green energy and implying short- and long-term temporal dependencies. It has static and dynamic attention, and its gates process data and unknown inputs and combine features. The fact that TFT decision-making is made more effective is due to an emphasis on the features that should be interpreted; however, the general structure of the model is optimal for long-term forecasts. TFT is more suitable than others for handling heterogeneous data and long-range dependencies in the actual and practical world, so it is effective for predicting renewable energy generation, including Latin American nations and Canada for 2024–2040.

3.2.2 | Experimental Setup

This research evaluates the TFT model's predictive performance in estimating green power generation across Canada and five Latin American countries. The data set was preprocessed by applying MinMaxScaler to remove or clean

away impurities so that every feature was within the same scale. Another essential feature is the year range and temporal/stationary parameters. Sub-modules about position encoding, word vectors based on pre-trained embedding and temporal self-attention layers are employed to extract the input. The gating mechanisms (GRN) in the TFT model to filter or prioritize aspects increase its capability to calculate the quantity of green energy produced. The characteristics are limited to only what can help determine green energy production. The static enrichment layers are time-invariant factors like countries' energy policies and structures, which enrich the data set by providing constant relevant contextual information for forecasting the long-run trends in green energy. LSTM layers help capture long-term temporal dependencies. The model is trained to expand the temporal context range of the data set by using knowledge of historical trends in energy production, such as seasonal variations and increased trends in renewable energy generation. With the help of the multi-head attention mechanism, a model can pay attention to different time steps and features simultaneously, which results in a better accuracy of green energy production forecasting. It enables the registration of intricate relations such as temporal and trend effects. In the model, information from different heads of attention is combined to provide better predictions of future energy trends. When incorporated in this way, the model can manage the challenge of green energy production, meet sustainability objectives, and increase the data set's capacity to predict future energy trends. The model contains six hidden layers; each layer has 64 neurons, a sequence-specific feature mapping layer, an embedding layer representing the embedding of feature vectors, a sequence-specific positional encoding layer, a multi-head attention layer, and eight attention heads along with a GLU layer and finally a fully connected layer used for output generation. The intended design contains two successive layers, LSTM and multi-head attention, that progressively improve sequential and contextual components. The architecture effectively fuses attention with sequential methods for robust time series forecasting using a sequence length of three-time steps. We applied the Adam optimizer to train for 100 iterations with a learning rate and norm value of 0.001 and an MSE between the training and validation stages. Algorithm 1 shows the pseudocode of the deep learning technique in TFT, and Table 4 summarizes the experimental setup.

The TFT model dynamically selects relevant characteristics, identifies long-range relationships, and adapts to data set structure patterns. Static covariate encoders better represent time-invariant characteristics, whereas temporal embeddings and masked self-attention layers capture seasonal and periodic energy production patterns. The model generates projections for chosen nations using 16-batch processing. The results reveal future energy generation, proving the model's capacity to handle vast and complicated information. The research evaluates the model predictions accurately using MSE as the significant loss function. This approach shows the TFT's promise in green energy forecasting and lays the groundwork for future prediction accuracy and scalability research.

TABLE 4 | Experimental setup of the TFT model.

Component	Specification
Model architecture	Temporal fusion transformer
Feature count	Corresponds to the number of countries
Neurons per layer	64 (hidden size)
Hidden layers	6 (Embedding (1), Positional Encoding (1), Multi-head Attention (1), GLU (1), LSTM (1), FC (1))
Continuous layers	2 (LSTM, multi-head attention)
Attention heads	8
Sequence length	3
Optimization	Adam optimizer
Learning rate	0.001
Training epochs	100
Norm value	0.001
Loss function	MSE
Batch size	16
Model outputs	Predictions of energy production for five Latin American countries and Canada

Abbreviations: LSTM, long short-term memory network; TFT, temporal fusion transformer.

3.3 | Evaluation Metrics

Our strategy is evaluated using several assessment measures, such as MAE, MSE, RMSE, and MAPE, to determine its efficiency. The values are given by Equations (19–22). MAE measures the difference between the actual and predicted values as much as it is a measure of error. The RMSE value here highlights the deviation from the expected and actual values. The MSE measures the sum of squared amounts by which observed and anticipated values diverge. MAPE is the average percent error between the observed and anticipated value [22, 45].

$$MSE = \frac{1}{s}\sum_{1}^{s}(g - \hat{g})^2, \tag{19}$$

$$MAE = \frac{1}{s}\sum_{1}^{s}|g - \hat{g}|, \tag{20}$$

$$RMSE = \sqrt{\frac{1}{s}\sum_{1}^{s}\left(g - \hat{g}\right)}, \tag{21}$$

$$\text{MAPE} = \frac{1}{s}\sum_{1}^{s}\left|\frac{g - \hat{g}}{g}\right| \times 100. \tag{22}$$

where g is the actual power entry, $\hat{g}$ is the forecasting value, and s is the total number of entries.

Algorithm 1. *Pseudocode for green energy forecasting using deep learning models*

1. Initialize Data set:
–Load GEP data set (GWh) from 1965 to 2023 for five Latin American countries (Argentina, Brazil, Chile, Colombia, and Mexico) and Canada.
–Features: Year, GEP for each country.
–Targets: Future GEP values for forecasting.
2. Data Preprocessing:
a. Ensure no missing values or outliers in the data set.
b. Using Min–Max normalization to align data set values between 0 and 1.
c. Split the data:
–Training set: 1965–2013.
–Validation set: 2014–2018.
–Testing set: 2019–2023.
3. Define Models:
a. Temporal fusion transformer (TFT) Model:
–Input: Sequential time-series data.
–Process: Use temporal self-attention, static enrichment layers, LSTM, and GRN to capture long-term dependencies and temporal dynamics.
–Output: Predicted GEP values.
b. Other Models (for comparison):
–MegaCRN: For capturing spatial dependencies in data.
–GRU: For sequential modeling. DeepAR: For time-series data.
–LSTM: For sequential predictions.
4. Model Compilation:
–Function for loss: MSE for all models.
–Optimizer: A learning rate of 0.001 for Adam's optimizer to rapid convergence.
5. Train the Models:
–Input: Data set for training from 1965 to 2013.
–Epoch size: 100 (use early stopping to prevent overfitting).
–Batch size: 16.
–Norm value: 0.001.
–Validate models using the validation set (2014–2018).
6. Model Evaluation:
–Evaluate the models on the testing set (2019–2023).
–Metrics: Find out how accurate your predictions are by using MAE, RMSE, MSE, and MAPE.
–Compare the performance of TFT with MegaCRN, GRU, DeepAR, and LSTM models.
7. Prediction Future GEP:
–Input: updated data set.
–Predict GEP from 2024 to 2040 using the best-performing model (TFT).
8. Output:
–Show the GEP projections for five Latin American countries and Canada (2024–2040).
–Showing performance analysis (MSE, MAE, RMSE, MAPE) of TFT with other models.

4 | Results and Discussion

All of the DL models are developed in Python using the Keras package. The experiments are done in an environment with Python 3.8.5 and TensorFlow 2.6.0 on a computer with an Intel Core i7-10700K@3.8 GHz and Windows 10, version 21H1. The training and testing are carried out using Google Colab [46], which permits the high utilization of graphic processing units (GPUs) to train models. The data set is divided into training (1965–2013), validation (2014–2018), and testing (2019–2023) sets to ensure the robustness of the evaluation. This division allows the TFT model to train on sufficient historical data, validate its hyperparameters on unseen data, and assess its generalization performance in the most recent years, as mentioned in Subsection 3.1.2. Data leakage is avoided by strictly maintaining the separation between these datasets, ensuring reliable and accurate long-term predictions [47–54]. Here are the key takeaways from the results of the models:

1. We evaluated the TFT model's predictive and improving capabilities using GEP data from several Latin American countries and Canada.
2. Based on the results from several Latin American countries and Canada, we assessed the TFT's interpretability by employing GEP information.
3. To evaluate the TFT model's efficacy on the data set, we contrasted it with MegaCRN, GRU, DeepAR, and LSTM.
4. We experimented on five Latin American countries and Canada datasets for the ablation research.
5. We compared its outcomes with other studies to assess the TFT model.

4.1 | TFT Model Performance on Data Sets of Latin American Countries and Canada

Figure 5 shows the performance of the TFT model using the data set from five Latin American countries and Canada, where the loss metric is MSE. The plot illustrates both the training and validation MSE loss over 100 epochs. The validation MSE loss is relatively high, peaking around 1.2 initially, but rapidly decreases within the first 10 epochs. The training MSE loss starts at a lower value than the validation loss and steadily declines as the epochs progress. Both losses stabilize after approximately 30 epochs, with the training loss converging near zero, while the validation loss remains slightly higher but stable. This behavior indicates that the model is well-trained and effectively avoids overfitting.

Table 5 presents the performance of the TFT model on the testing data set for six countries: Argentina, Brazil, Canada, Chile, Colombia, and Mexico. The model achieves an MSE of 0.0007 for Argentina, RMSE of 0.0264, MAE of 0.0261, and MAPE of 2.50%. Brazil has the MSE of 0.0023, RMSE of 0.0479, MAE of 0.0963, and MAPE of 4.30%. Canada shows the best performance with the MSE of 0.0003, RMSE of 0.0173, MAE of 0.0112, and MAPE of 1.76%. Chile records the MSE of 0.0046, RMSE of 0.0678, MAE of 0.1098, and MAPE of 6.30%. For Colombia, the MSE is 0.0033, RMSE is 0.0574, MAE is 0.1019, and MAPE is 5.40%. Mexico shows the highest error values, with the MSE of 0.0082, RMSE of 0.0905, MAE of 0.1211, and MAPE of 7.60%. This highlights the varying performance of the TFT model across countries, with Canada having the best accuracy and Mexico having the lowest.

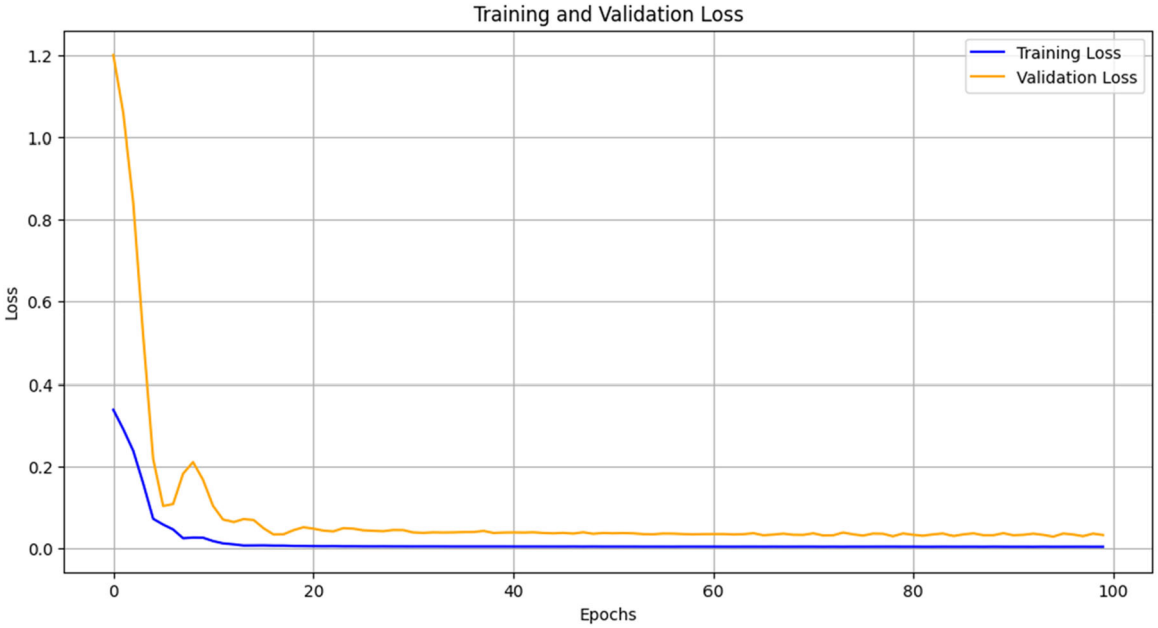

FIGURE 5 | Training and validation losses with temporal fusion transformer (TFT) model.

TABLE 5 | Performance of the TFT model on a testing set.

Country	MSE	RMSE	MAE	MAPE (%)
Argentina	0.0007	0.0264	0.0261	2.50
Brazil	0.0023	0.0479	0.0963	4.30
Canada	0.0003	0.0173	0.0112	1.76
Country	MSE	RMSE	MAE	MAPE (%)
Chile	0.0046	0.0678	0.1098	6.30
Colombia	0.0033	0.0574	0.1019	5.40
Mexico	0.0082	0.0905	0.1211	7.60

Abbreviations: MAE, mean absolute error; MAPE, mean absolute percentage error; MSE, mean square error; RMSE, root mean square error; TFT, temporal fusion transformer.

The reliability of the annual TFT model for predicting green energy production in Latin American countries and Canada is based on the following essential factors. The management of yearly trends and dependencies is, therefore, correct. Failure to correctly integrate or interpret data from countries such as Argentina, Brazil, Canada, Chile, Colombia, and Mexico distorts the nature of energy market interactions. Further, it is possible to get less accurate predictions if the attention mechanism and hyperparameters of the model are not chosen for annual data. Failure to correctly manage missing data and incorrect application of normalization techniques, such as Min-Max normalization, also affect forecast accuracy, hence bringing about errors. Therefore, controlling these factors is crucial to achieving high accuracy in forecasts.

From 2024 to 2040, Table 6 provides a detailed forecast of GEP in GWh for six countries in the Americas using the TFT model. The forecast for Argentina starts at 38,812 GWh in 2024 and will increase to 46,175 GWh by 2040. Brazil's production is projected to rise from 526,397 GWh to 640,715 GWh, showing the largest increase among the listed countries. In Canada, the GEP starts at 437,509 GWh and climbs to 487,520 GWh. Chile's forecast begins at 46,538 GWh, escalating to 66,359 GWh. Colombia starts at 62,149 GWh and is expected to grow to 80,779 GWh. Lastly, Mexico's production begins at 97,203 GWh and reaches 155,091 GWh by 2040, indicating significant growth in electrical output across the years.

The detailed projections for GEP from 1965 to 2040 in six countries are depicted in Figure 6 for Argentina, Brazil, Canada, Chile, Colombia, and Mexico, illustrated through bar charts. These charts detail annual production figures, with the x-axis representing the years and the y-axis indicating GEP in gigawatt-hours (GWh). Historical data from 1965 to 2023 shows varying trends in electrical production, and forecasted data from 2024 to 2040 indicate expected growth. Starting with Argentina, the GEP in 1965 is modest at about 1225 GWh and has seen various ups and downs, with a significant rise forecasted to reach 46,175 GWh by 2040. Brazil, starting at 25,515 GWh in 1965, exhibits a steep upward trajectory, anticipated to surge to the highest among the group at 640,715 GWh by 2040. Canada's initial figure is the largest at 118,088 GWh in 1965, with projections showing a consistent increase. In Chile, the GEP started at 3954 GWh in 1965, with projections suggesting a growth to 66,359 GWh by 2040, reflecting robust development. Colombia began at a slightly lower production level of 3543 GWh in 1965, with forecasts indicating a substantial rise to 80,779 GWh, showcasing significant development in energy capacity. Lastly, Mexico's GEP in 1965 was 8863 GWh, with a solid upward trend forecasted to escalate to 155,091 GWh by 2040. These bar charts visualize the past and projected future of electricity production, reflecting the increasing industrialization and energy demand in these regions over 75 years.

Using the TFT model, Figure 7 depicts the relationship between actual and predicted GEP values from 1965 to 2023 for six countries: Brazil, Canada, Chile, Colombia, Mexico, and Argentina. The fitted values of actual and predicted GEP over time

TABLE 6 | GEP Forecasting on a test set using a TFT Model.

Year	GEP in Argentina	GEP in Brazil	GEP in Canada	GEP in Chile	GEP in Colombia	GEP in Mexico
2024	38,812	526,397	437,509	46,538	62,149	97,203
2025	38,272	533,542	440,010	47,339	63,313	95,196
2026	38,732	540,687	462,511	52,141	65,477	99,189
2027	40,192	547,832	465,011	53,942	64,642	97,182
2028	39,653	574,977	447,512	55,743	63,806	100,175
2029	42,113	582,122	470,013	57,545	64,971	105,168
2030	40,573	599,266	482,514	59,346	68,135	125,161
2031	39,033	576,411	475,014	57,147	70,300	119,154
2032	43,493	563,556	467,515	60,949	71,464	123,147
2033	41,953	600,701	460,016	59,750	72,628	127,140
2034	44,414	597,846	462,516	62,551	70,793	151,133
2035	42,874	604,991	455,017	63,353	78,957	135,126
2036	45,334	612,136	447,518	64,154	72,122	129,119
2037	44,794	619,280	450,018	65,955	75,286	143,112
2038	42,254	646,425	482,519	66,757	78,451	147,105
2039	44,715	653,570	495,020	64,558	79,615	161,098
2040	46,175	640,715	487,520	66,359	80,779	155,091

Abbreviations: GEP, green electricity production; TFT, temporal fusion transformer.

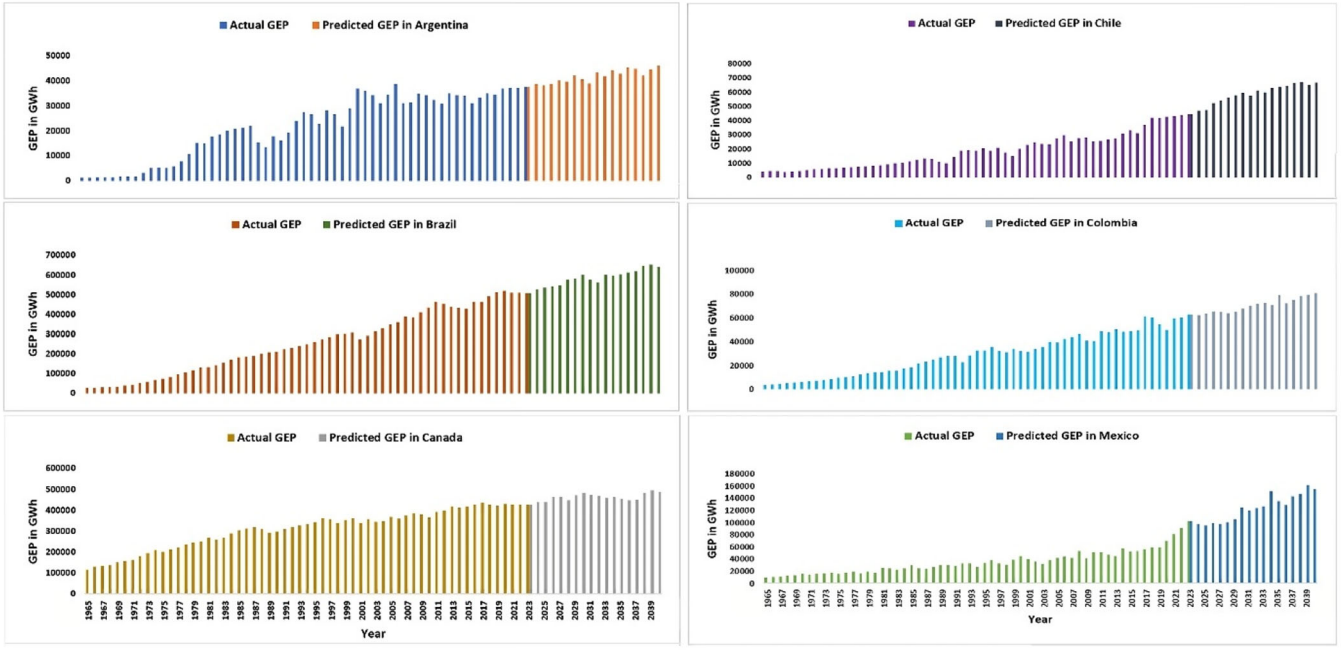

FIGURE 6 | GEP forecasting on a test set using a TFT model. GEP, green electricity production; TFT, temporal fusion transformer.

are plotted, providing insights into the model's forecasting accuracy.

The actual GEP data also supports the fitted curve for Argentina, which shows a steady increase. For example, in 2000, the real GEP was about 28,890 GWh, which is nearly equal to the estimated value of 28,499. However, small fluctuations are observed during periods such as the 1980s. This trend is best captured in the model for Brazil, where actual and predicted GEPs exhibit a strong linear relationship. For instance, the actual GEP in 2020 was 520,010, and the predicted GEP in 2020 was 540,014, which means a very slight difference. Canada reveals exponential growth, and actual and predicted curves are nearly coincident. For 2022, the actual GEP was 427,135, and

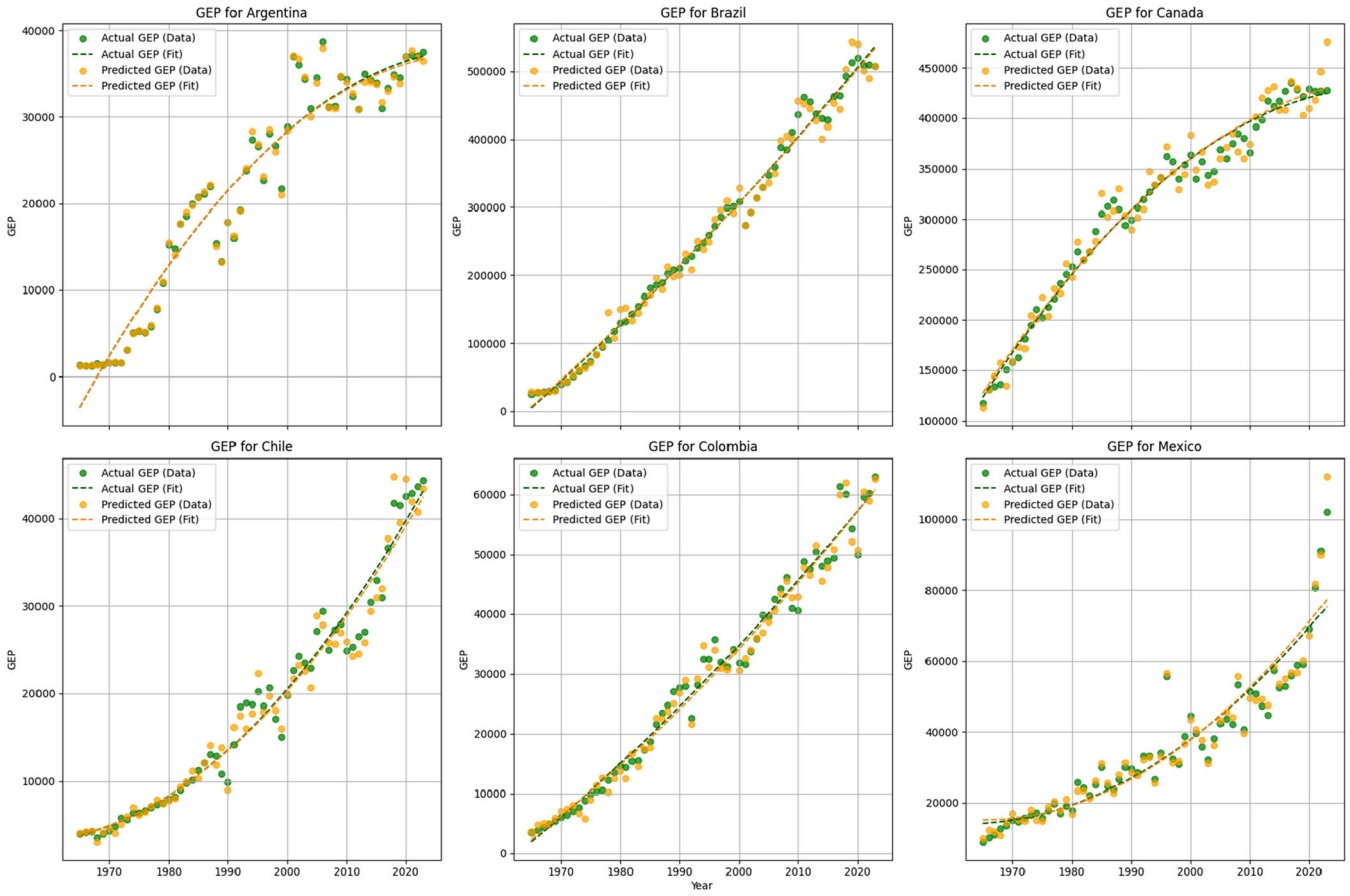

FIGURE 7 | Fitted curves of the predicted values versus the actual values on the test set using TFT. GEP, green electricity production; TFT, temporal fusion transformer.

the predicted value was 446,140, which testifies to the efficiency of the given model in terms of its ability to estimate long-term tendencies. The same situation is observed in Chile, where the model correctly predicts GEP values. For instance, the actual GEP in 2021 was 42,933, which is very close to the predicted GEP of 41,937. The model fits with actual Colombian data, particularly after the year 2000. The actual value of GEP for 2023 was 63,060, while the predicted value was 62,666; hence, the error difference was minimal. Finally, the predicted GEP overlays the actual for Mexico, though there is some disparity in specific years. For instance, the actual GEP in 2018 was 58,780, and the predicted value was 56,798, which is very near the actual value. The fitted curves prove that the TFT model can track the general direction and variations of GEP in all six countries, and the numerical differences are minimal.

4.2 | Interpretability of the TFT on the Data Set

The attention mechanism in the TFT model is one of the most critical components since it contributes to both interpretability and performance improvement. It can dynamically assign different lagged inputs to different importance levels, thus allowing the model to determine which historical values are most relevant to future values. This feature is essential for multivariate time series data and plays a significant role in finding temporal interactions between the variables.

To observe the relevance of different lag orders, the model takes attention allocation through the yearly data from 1965 to 2023 for six countries: Argentina, Brazil, Chile, Colombia, Mexico, and Canada. The model's penetration in Argentina pays the most attention to the third lag (35%), indicating a strong temporal dependence at this interval. Similarly, the third lag (40%) is a concern in Brazil, given its contribution to prediction. It shows that the second and third lags are equally crucial for Chile (30% each), implying important midterm implications. The third lag is the most discussed in Colombia (30%), followed by the second (25%), as in Argentina. The distribution in Mexico is balanced, with increasing importance from the first to the third lag, reaching 30%. In Canada, the third lag reemerges, garnering 35 per cent of the attention and its importance to predictive accuracy is evidenced in Table 7.

Figure 8 shows the fluctuation of attention values for six countries (Argentina, Brazil, Chile, Colombia, Mexico, and Canada) and five lag orders. Every line reflects the dynamics of attention values for a particular country; points are marked if they contain the maximal value. Brazil hits the highest of 0.40 during the third Lag Order, while Argentina, Chile, Colombia, Mexico, and Canada show their maximums at lag orders 2 and 3. The trends also reveal different patterns of attention distribution for each country.

TABLE 7 | Lag orders on the data set.

Lag order	Argentina	Brazil	Chile	Colombia	Mexico	Canada
1	0.1	0.05	0.15	0.2	0.1	0.05
2	0.25	0.2	0.3	0.25	0.2	0.15
3	0.35	0.4	0.3	0.3	0.3	0.35
4	0.2	0.25	0.15	0.15	0.25	0.3
5	0.1	0.1	0.1	0.1	0.15	0.15

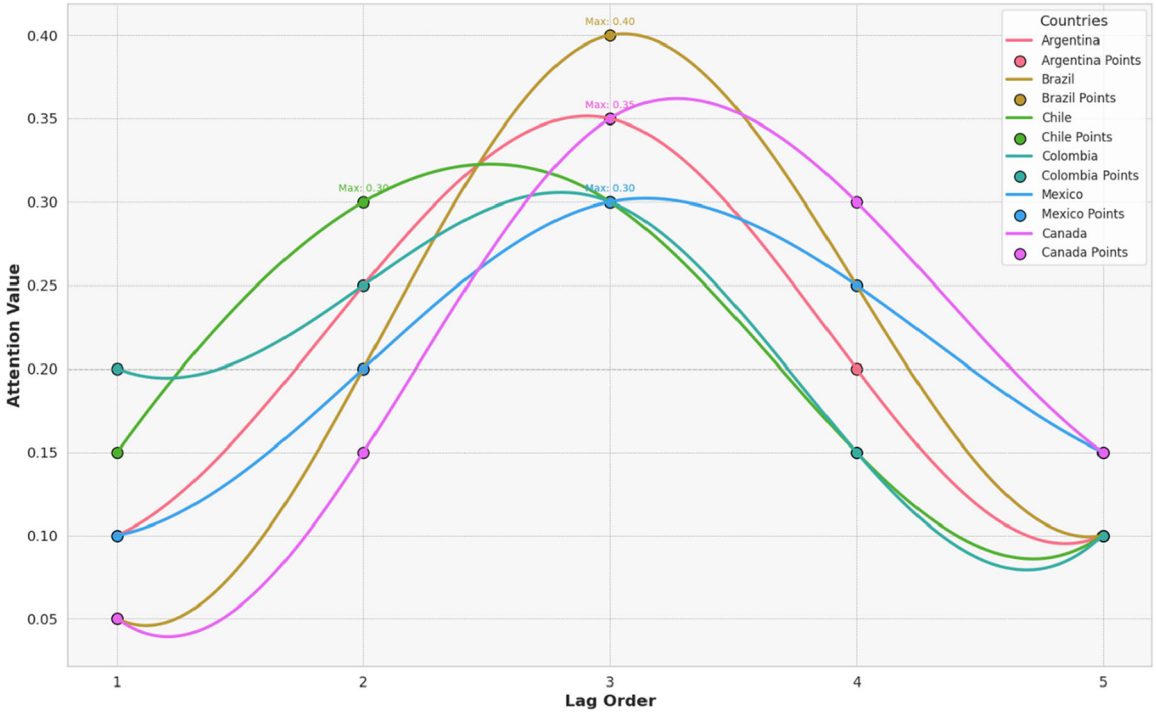

FIGURE 8 | Attention value visualization by lag order.

4.3 | Analyzing the Data Set for Deep Learning Models

This research uses the TFT model because it excels in multivariate time series forecasting, especially in green energy generation. TFT accurately captures short-term and long-term temporal connections, which is essential for annual datasets. Its multivariate capabilities allow various influencing elements, and its interpretability via attention processes helps determine their relevance. TFT also does well at time series forecasting jobs, so it is a good choice for this application.

Instead, we compare TFT against some of the top models in GRU, LSTM, MegaCRN, and DeepAR. Although both models may capture temporal relationships, GRU is computationally efficient, while LSTM is able to model long-term patterns. In particular, DeepAR is a popular probabilistic model for time series forecasting, whereas MegaCRN is a spatiotemporal view with its graph-based approach. The selection of each model is based on its relevance and proven function in time series analysis. However, TFT succeeds in providing interpretability and performance to meet demands from green energy data sets.

Table 8 compares forecasting models (MegaCRN, GRU, DeepAR, LSTM, and TFT) across five Latin American countries and Canada using critical metrics like MSE, RMSE, MAE, and MAPE. TFT shows the best performance for Argentina with an MSE of 0.0007, RMSE of 0.0264, MAE of 0.0261, and MAPE of 2.50%. Other models, such as MegaCRN and GRU, report higher MAPE values of 8.43% and 8.01%, respectively. In Brazil, TFT again leads with an MSE of 0.0023, RMSE of 0.0479, MAE of 0.0963, and a MAPE of 4.30%, while MegaCRN and GRU show MAPE values of 9.11% and 8.97%.

TFT excels in Canada with a MAPE of 1.76%, MSE of 0.0003, and MAE of 0.0112. MegaCRN has 11.37% MAPE, substantially higher. Compared to MegaCRN at 10.31% and GRU at 8.89%, TFT again has the lowest errors in Chile, with a MAPE of 6.30%. TFT continues to perform well in Colombia with an MSE of 0.0033, RMSE of 0.0574, MAE of 0.1019, and MAPE of 5.40%. MegaCRN has 10.01% and GRU 9.23%. TFT is the most accurate model in Mexico, with MSE 0.0082, RMSE 0.0905, MAE 0.1191, and MAPE 7.60%. MegaCRN has the highest MAPE at 15.07%. TFT surpasses other models in MAE and MAPE across all nations, demonstrating its accuracy. In most circumstances,

TABLE 8 | Analyzing forecasting models' performance on a testing set of Latin American Countries and Canada.

Country	Model	MSE	RMSE	MAE	MAPE (%)
Argentina	MegaCRN	0.0091	0.0953	0.0771	8.43
	GRU	0.0090	0.0948	0.0762	8.01
	DeepAR	0.0048	0.0692	0.0566	6.12
	LSTM	0.0023	0.0479	0.0326	4.19
	TFT	**0.0007**	**0.0264**	**0.0261**	**2.50**
Brazil	MegaCRN	0.0109	0.1044	0.1299	9.11
	GRU	0.0101	0.1004	0.1278	8.97
	DeepAR	0.0053	0.0728	0.1187	6.23
	LSTM	0.0034	0.0583	0.1066	5.42
	TFT	**0.0023**	**0.0479**	**0.0963**	**4.30**
Canada	MegaCRN	0.0087	0.0932	0.0889	11.37
	GRU	0.0057	0.0754	0.0601	8.13
	DeepAR	0.0036	0.0600	0.0477	5.90
	LSTM	0.0018	0.0424	0.0232	4.52
	TFT	**0.0003**	**0.0173**	**0.0112**	**1.76**
Chile	MegaCRN	0.0139	0.1178	0.1342	10.31
	GRU	0.0112	0.1058	0.1297	8.89
	DeepAR	0.0097	0.0984	0.1247	8.55
	LSTM	0.0089	0.0943	0.1116	7.35
	TFT	**0.0046**	**0.0678**	**0.1098**	**6.30**
Colombia	MegaCRN	0.0109	0.1044	0.1268	10.01
	GRU	0.0097	0.0984	0.1207	9.23
	DeepAR	0.0088	0.0938	0.1146	8.94
	LSTM	0.0047	0.0685	0.1095	6.29
	TFT	**0.0033**	**0.0574**	**0.1019**	**5.40**
Mexico	MegaCRN	0.0189	0.1374	0.1602	15.07
	GRU	0.0143	0.1195	0.1405	13.21
	DeepAR	0.0109	0.1044	0.1338	10.58
	LSTM	0.0099	0.0994	0.1281	8.29
	TFT	**0.0082**	**0.0905**	**0.1211**	**7.60**

Abbreviations: DeepAR, deep autoregression; GRU, gated recurrent unit; LSTM, long short-term memory; MAE, mean absolute error; MAPE, mean absolute percentage error; MegaCRN, meta graph-based convolutional recurrent network; MSE, mean square error; RMSE, root mean square error; TFT, temporal fusion transformer.

LSTM outperforms MegaCRN and GRU, which have more excellent error metrics and are less dependable.

Figure 9 compares MegaCRN, GRU, DeepAR, LSTM, and TFT deep learning models across five Latin American countries (Argentina, Brazil, Chile, Colombia, and Mexico) and Canada using MSE, RMSE, MAE, and MAPE. Bar graphs in the top half compare MSE and RMSE, while line graphs in the bottom half compare MAE and MAPE. The TFT model has the lowest MSE, RMSE, MAE, and MAPE in most nations. LSTM performs well in Argentina, Brazil, and Canada. MegaCRN and GRU have better error rates across all measures. Mexico's MAPE is the highest among the models, suggesting bigger predicting mistakes. This thorough depiction shows that the TFT model outperforms the other models in predicting accuracy across all measures.

4.4 | Study of Ablation on the Data Set of Latin American Countries and Canada

In this ablation study, we examine the effects of training settings and changes to the TFT model on its performance. This process enables an understanding of the impact of different contexts on the model's ability to predict GEP fitness and, thereby, a sense of how much of a task it is to write models for this kind of dynamics. This subsection compares a few epochs, optimization algorithms, batch sizes, and learning rates. Further, we analyze the partial efficiency of each factor in the presented TFT model (Figure 10).

Table 9 compares the diverse model aspects of the Latin American countries and Canada data set. It compares the

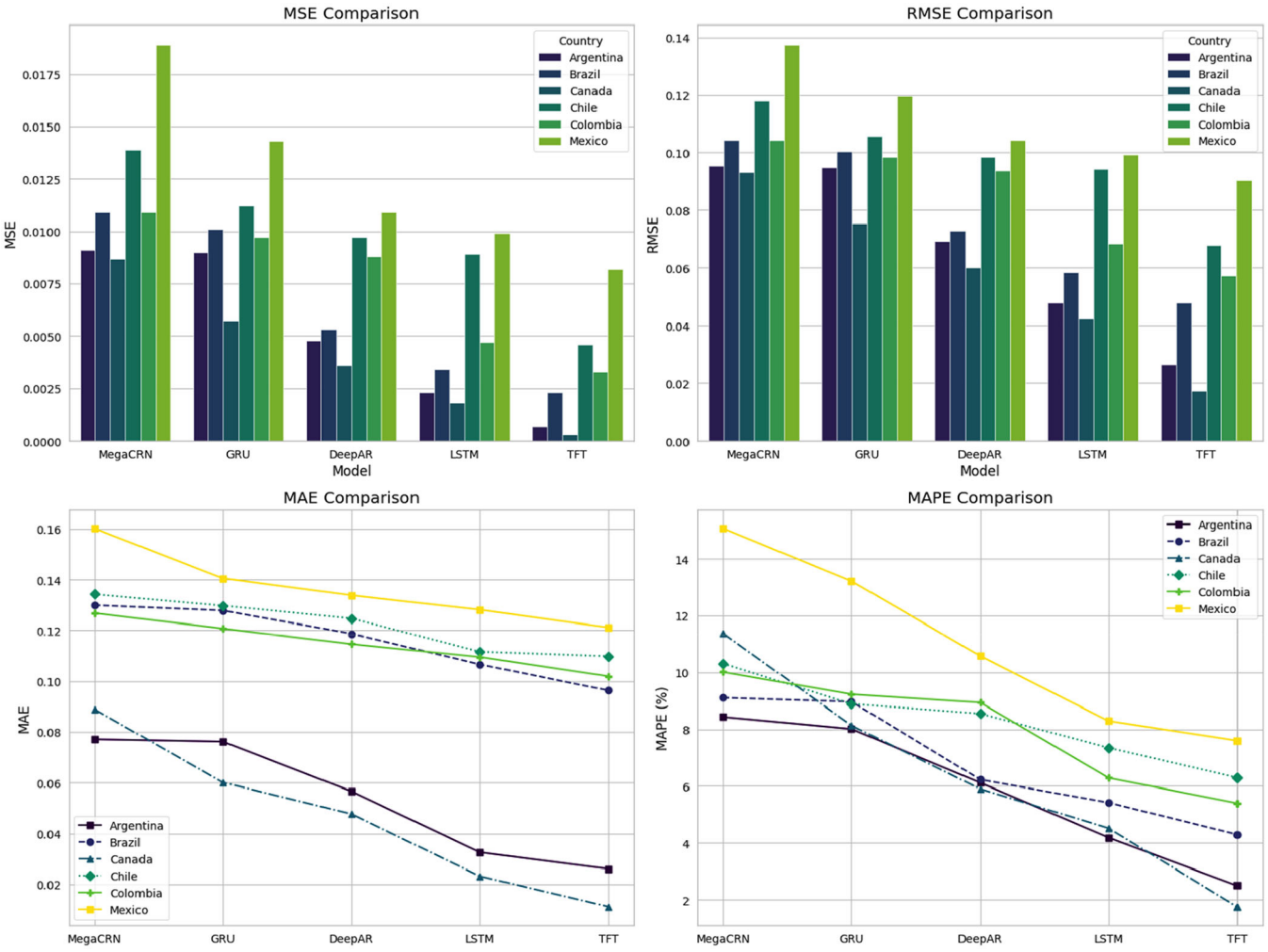

FIGURE 9 | Analyzing forecasting models' performance on a testing set. DeepAR, deep autoregression; GRU, gated recurrent unit; LSTM, long short-term memory; MAE, mean absolute error; MAPE, mean absolute percentage error; MegaCRN, meta graph-based convolutional recurrent network; MSE, mean square error; RMSE, root mean square error; TFT, temporal fusion transformer.

performance of Full TFT, temporal self-attention, GRN, and LSTM models across four error metrics: MSE, RMSE, MAE, and MAPE. The full TFT model has the lowest error rates, including the MSE of 0.0003 and a MAPE of 1.76%. On the other hand, the LSTM model gives the highest error values for each of the indices considered, with an MSE of 0.0008 and a MAPE of 7.60%.

Table 10 describes the experiment of training the TFT model using hyperparameters of epochs, optimizers, batch size, and learning rate. The lead index identifies the performance of five epoch settings ranging from 50 to 250 with optimizers: Adam, SGD, RMSprop, and AdamW. Of these configurations, the Adam optimizer with 100 epochs, a batch size of 16 and a learning rate of 0.001 gave the best results. MSE is also the lowest at 0.0003, with RMSE at 0.0173, MAE at 0.0112 and MAPE at 1.76%, thus confirming that this setup is the best for the data set of Latin American countries and Canada.

The ablation study conducted on the data set of Latin American countries and Canada indicates that the parameters used for the TFT model in Table 4 demonstrate the best practical performance. It is evident compared to the variations observed in Table 9 and Table 10. Specifically, the parameters used—100 epochs, the Adam optimizer, a batch size of 16, and a learning rate of 0.001—show the best results. Some aspects, like the number of attention heads, sequence length, and learning rate, can be adjusted to improve the TFT model's applicability, and some methods should be used for regularization to improve the model's generalization. Information from the real environment, including weather or policy changes, can be included for better accuracy and scalability using distributed computing for large amounts of data.

4.5 | Comparison of TFT Model Performance with Other Existing Studies

Most investigations are now conducted separately on solar, wind, and photovoltaic (PV) electricity. It does not look at how one can fit strategies related to total renewable energy sources. The contribution of our study is that we consider total forms of energy. Table 11 compares our data set with other datasets. Nevertheless, no other research has made use of this data set. This comparison demonstrates variations in the volume of data covered and functionalities. Straight comparisons are only

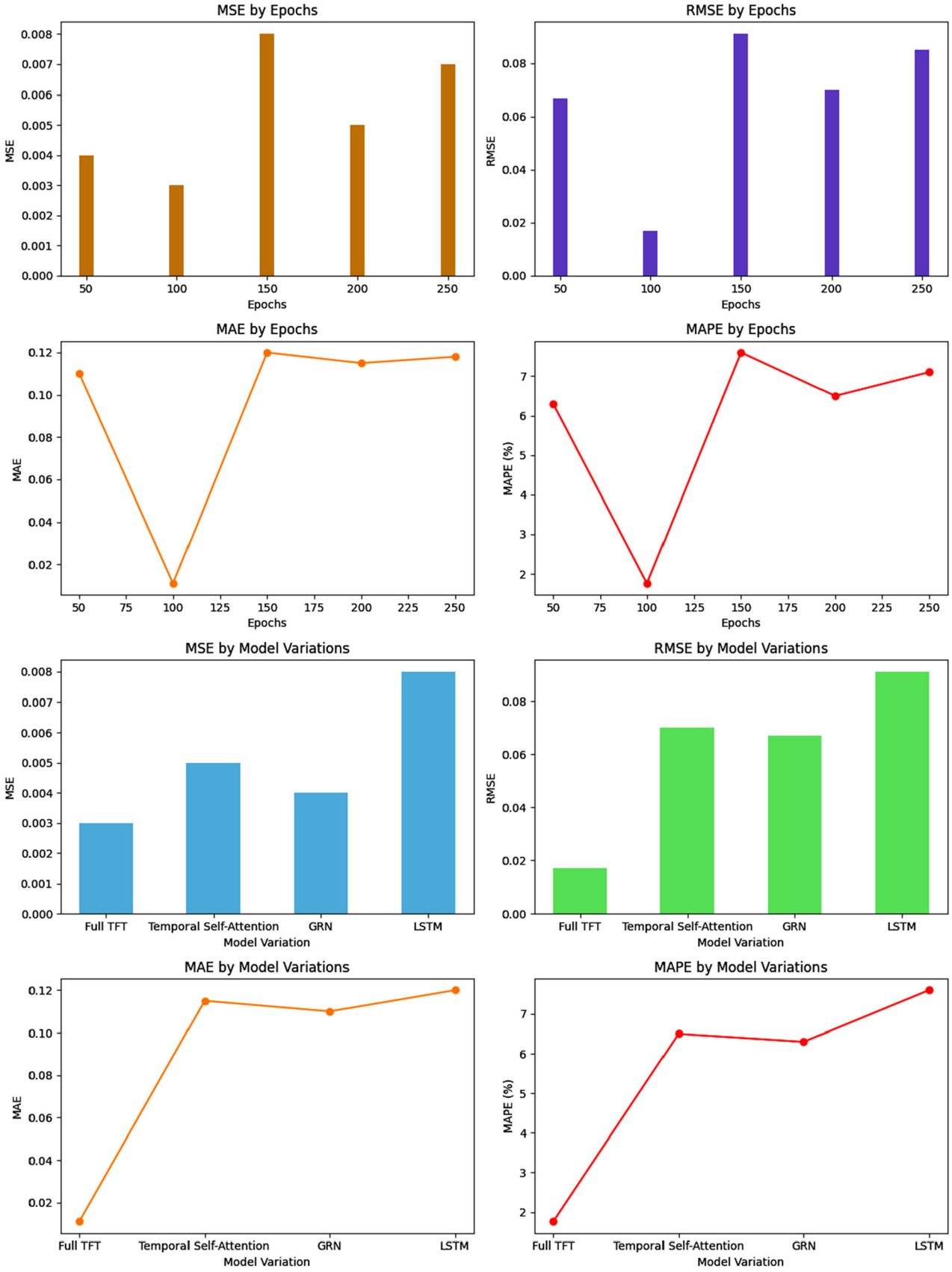

FIGURE 10 | Ablation study on the data set of Latin American countries and Canada. GRN, gating of residual network; LSTM, long short-term memory; MAE, mean absolute error; MAPE, mean absolute percentage error; MSE, mean square error; RMSE, root mean square error; TFT, temporal fusion transformer.

feasible if testing procedures and samples vary. In contrast, Table 8 presents the experimental results of other state-of-the-art models applied to the same data set of five Latin American countries and Canada, allowing for a more relevant comparison. However, the CNN-LSTM-Transformer model demonstrated moderate performance with the MAE (0.393) and RMSE (0.344) [20]. The Ensemble method exhibited higher error values with the MSE (3.9972), RMSE (1.9993), and MAE (0.8306) [22]. The CNN model performed better than the Ensemble method, with a reported MSE of 1.36 and RMSE of 0.47 [23]. Similarly, the LSTM model achieved an RMSE of 0.47 [24], and the LSTM-GRU combination has the MAE of 0.89 and the RMSE of 0.75 [27]. LightGBM, with the MAE of 1.1072 and RMSE of 1.7479, showed higher errors than other models [29]. The KNN and XGBoost combination yielded moderate results with the MAE of 0.578 and RMSE of 1.056 [30], while the GP-Ensemble method achieved better performance with the MSE (0.0631), RMSE (0.2383), and MAE (0.1754) [31]. However, the proposed TFT model outperformed all others with the lowest MSE (0.0003), MAE (0.0112), and RMSE (0.0173), indicating its superior accuracy and potential in energy forecasting.

TABLE 9 | Model component variations on the data set of Latin American countries and Canada.

Model variation	MSE	RMSE	MAE	MAPE
Full TFT	0.0003	0.0173	0.0112	1.76%
Temporal self-attention	0.0005	0.0224	0.1156	6.50%
GRN	0.0004	0.0200	0.1125	6.30%
LSTM	0.0008	0.0283	0.1211	7.60%

Abbreviations: GRN, gating of residual network; LSTM, long short-term memory; MAE, mean absolute error; MAPE, mean absolute percentage error; MSE, mean square error; RMSE, root mean square error; TFT, temporal fusion transformer.

5 | Conclusions and Future Work

This research addresses significant voids in green energy forecasting research: long-term forecast accuracy, simultaneous prediction of multiple energy types, and outdated datasets. Due to these limitations, there is no chance to build an accurate and stable prognosis, which is crucial for sustainable energy management. To this end, the primary objectives of this research were to enhance the overall forecasting precision of green electricity through the TFT model, enhance general long-term green electricity forecasting without overemphasizing errors, and lastly, utilize an upgraded data set to provide an accurate estimate of anticipated green electricity generation up to year 2040.

These objectives were achieved using the TFT, a temporal and co-dependencies multivariate time series model with advanced mechanisms. The paper employs data from 1965 to 2023 to generate accurate estimates on green electricity generation for five Latin American nations and Canada. According to the results, the proposed TFT model performs better than other models, such as MegaCRN, GRU, DeepAR, and LSTM. For instance, the TFT efficiently fetched an MSE of 0.0003, RMSE of

TABLE 10 | Impact of hyperparameters on the data set of Latin American countries and Canada.

Epochs	Optimizer	Batch size	Learning rate	MSE	RMSE	MAE	MAPE
50	Adam	16	0.001	0.0004	0.0200	0.1120	6.30%
100	Adam	16	0.001	0.0003	0.0173	0.0112	1.76%
150	SGD	32	0.0005	0.0008	0.0283	0.1243	7.60%
200	RMSprop	16	0.001	0.0009	0.0300	0.1158	6.50%
250	AdamW	16	0.0003	0.0007	0.0265	0.1181	7.10%

Abbreviations: MAE, mean absolute error; MAPE, mean absolute percentage error; MSE, mean square error; RMSE, root mean square error.

TABLE 11 | Analysis with prior studies.

Reference, year	Methodology	MSE	MAE	RMSE
[20], 2023	CNN-LSTM-Transformer	---	0.393	0.344
[22], 2023	Ensemble method	3.9972	0.8306	1.9993
[23], 2024	CNN	1.36	---	---
[24], 2024	LSTM	---	---	0.47
[27], 2024	LSTM-GRU		0.89	0.75
[29], 2024	LightGBM	---	1.1072	1.7479
[30], 2024	KNN and XGBoost	---	0.578	---
[31], 2024	GP-Ensemble	0.0631	0.1754	0.2383
Proposed model	**TFT model**	**0.0003**	**0.0112**	**0.0173**

Abbreviations: CNN, convolutional neural network; GRU, gated recurrent unit; LSTM, long short-term memory; MAE, mean absolute error; MAPE, mean absolute percentage error; MSE, mean square error; RMSE, root mean square error; TFT, temporal fusion transformer.

0.0173, and MAE of 0.0112 and offered 1.76% MAPE for Canada. As can be seen, the development of green energy production in all regions significantly corresponds to the SDGs, and the fulfilled calculations provide practical recommendations for energy policy and planning.

As already stated, the findings of the study were positive. However, many things could have been done better. It does not explore other factors that can influence the predictions, such as the weather and changes resulting from policies, to make the predictions even more appropriate. Thus, the research scope is narrow because it includes only five countries from Latin America and Canada. Despite this, the TFT model is still possible, and if accompanying real-time data and datasets of better detail are put in, it will make better predictions. The study also considers basic mistake measures and could also consider what these measures imply in the real world. Therefore, this model has not been applied in other energy fields or areas and cannot be used in as many places as desired.

The researchers should extend future studies to enhance forecasts' accuracy by including other variables, including weather conditions, policy changes, and economic factors. The model could be more helpful by increasing geographical coverage and bringing aspects from different regions with different energy systems. Using actual time data and high-resolution data in the model will improve the dynamic response and the accuracy of the outcome while including the quantification of uncertainty, which will provide more reliability in decision-making. Expanding its application to other sectors of the energy system, such as grid balancing and energy storage, would further prove the effectiveness of the TFT model. Future research should also extend to the enhanced measure of prediction errors, emphasizing the real-life impact of such mistakes. Future research should also include a mixture of TFT with other methods to optimize the use of the technique. Addressing these aspects will ensure the model remains spirited and essential in the emergent green energy environment.

Author Contributions

Muhammad Shoaib Saleem: conceptualization, methodology, supervision, software, writing–original draft. **Javed Rashid:** investigation, formal analysis, resources, supervision, writing–review and editing. **Sajjad Ahmad:** data curation, formal analysis, investigation, resources, validation, writing–original draft. **Ali M. Al-Shaery:** data curation, investigation, software, validation, writing–review and editing. **Saad Althobaiti:** methodology, investigation, software, resources, writing–review and editing. **Muhammad Faheem:** data curation, investigation, software, validation, writing–review and editing.

Acknowledgments

The authors thank their affiliated universities and institutes for providing research facilities. The work of Muhammad Faheem is funded by the VTT Technical Research Centre of Finland. The authors extend their appreciation to Taif University, Saudi Arabia, for supporting this work through project number (TU-DSPP-2024-66).

Conflicts of Interest

The authors have no conflicts of interest.

Data Availability Statement

The data will be available upon request to the corresponding author.

Code Availability

The code will be available upon request to the corresponding author.

References

1. L. Chen, Y. Hu, R. Wang, et al., "Green Building Practices to Integrate Renewable Energy in the Construction Sector: a Review," *Environmental Chemistry Letters* 22, no. 2 (2024): 751–784.

2. "Irena, World Energy Transitions Outlook," 2022, https://www.irena.org/Digital-Report/World-Energy-Transitions-Outlook-2022.

3. S. Saeed and T. Siraj, "Global Renewable Energy Infrastructure: Pathways to Carbon Neutrality and Sustainability," *Solar Energy and Sustainable Development Journal* 13, no. 2 (2024): 183–203.

4. A. Silveira, "Minister of Mines and Energy," accessed October 12, 2024, https://www.gov.br/mme/pt-br/assuntos/secretarias/sntep/publicacoes./resenha-energetica-brasileira/resenhas/brazilian-energy-review-2023.

5. "Let's Talk Science, Generating Electricity: Hydroelectric Power," accessed October 12, 2024, https://letstalkscience.ca/educational-resources/stem-explained/generating-electricity-hydroelectric-power.

6. C. Cacciuttolo, M. Navarrete, and E. Atencio, "Renewable Wind Energy Implementation in South America: A Comprehensive Review and Sustainable Prospects," *Sustainability* 16, no. 14 (2024): 6082.

7. N. M. Rivera, J. C. Ruiz-Tagle, and E. Spiller, "The Health Benefits of Solar Power Generation: Evidence From Chile," *Journal of Environmental Economics and Management* 126 (2024): 102999.

8. S. B. Botero, C. M. G. Mazo, and F. J. M. Arboleda, "Power Generation Mix in Colombia Including Wind Power: Markowitz Portfolio Efficient Frontier Analysis With Machine Learning," *Journal of Open Innovation: Technology, Market, and Complexity* 10 (2024): 100402.

9. D. Toledo-Vázquez, R. J. Romero, G. Hernández-Luna, J. Cerezo, and M. Montiel-González, "Projections for the 2050 Scenario of the Mexican Electrical System," *Energies* 17, no. 17 (2024): 4326.

10. A. Ullah, H. Nobanee, S. Ullah, and H. Iftikhar, "Renewable Energy Transition and Regional Integration: Energizing the Pathway to Sustainable Development," *Energy Policy* 193 (2024): 114270.

11. I. D. Mienye, T. G. Swart, and G. Obaido, "Recurrent Neural Networks: A Comprehensive Review of Architectures, Variants, and Applications," *Information* 15, no. 9 (2024): 517.

12. F. M. A. Mazen, Y. Shaker, and R. A. Abul Seoud, "Forecasting of Solar Power Using GRU–Temporal Fusion Transformer Model and DILATE Loss Function," *Energies* 16, no. 24 (2023): 8105.

13. J. Li, W. Chen, Z. Zhou, J. Yang, and D. Zeng, "DeepAR-Attention Probabilistic Prediction for Stock Price Series," *Neural Computing and Applications* 36 (2024): 1–18.

14. G. Jin, Y. Liang, Y. Fang, et al., "Spatio-Temporal Graph Neural Networks for Predictive Learning in Urban Computing: A Survey," *IEEE Transactions on Knowledge and Data Engineering* 36 (2023): 5388–5408.

15. M. Kiasari, M. Ghaffari, and H. Aly, "A Comprehensive Review of the Current Status of Smart Grid Technologies for Renewable Energies Integration and Future Trends: The Role of Machine Learning and Energy Storage Systems," *Energies* 17, no. 16 (2024): 4128.

16. M. Awais, R. Mahum, H. Zhang, et al., "Short-Term Photovoltaic Energy Generation for Solar Powered High Efficiency Irrigation Systems Using LSTM With Spatio-Temporal Attention Mechanism," *Scientific Reports* 14, no. 1 (2024): 10042.

17. S. Lv, L. Wang, and S. Wang, "A Hybrid Neural Network Model for Short-Term Wind Speed Forecasting," *Energies* 16, no. 4 (2023): 1841.

18. S. Kaewarsa and V. Kongpaseuth, "An Energy Prediction Approach Using Bi-Directional Long Short-Term Memory for a Hydropower Plant in Laos," *Electrical Engineering* 106, no. 3 (2024): 2609–2625.

19. X. Wang, H. Wang, B. Bhandari, and L. Cheng, "AI-Empowered Methods for Smart Energy Consumption: A Review of Load Forecasting, Anomaly Detection and Demand Response," *International Journal of Precision Engineering and Manufacturing-Green Technology* 11, no. 3 (2024): 963–993.

20. E. M. Al-Ali, Y. Hajji, Y. Said, et al., "Solar Energy Production Forecasting Based on a Hybrid CNN-LSTM-Transformer Model," *Mathematics* 11, no. 3 (2023): 676.

21. T. Anu Shalini and B. Sri Revathi, "Hybrid Power Generation Forecasting Using CNN Based BILSTM Method for Renewable Energy Systems," *Automatika* 64, no. 1 (2023): 127–144.

22. A. Mystakidis, E. Ntozi, K. Afentoulis, et al., "Energy Generation Forecasting: Elevating Performance With Machine and Deep Learning," *Computing* 105, no. 8 (2023): 1623–1645.

23. A. Nikulins, K. Sudars, E. Edelmers, et al., "Deep Learning for Wind and Solar Energy Forecasting in Hydrogen Production," *Energies* 17, no. 5 (2024): 1053.

24. M. Pasandideh, M. Taylor, S. R. Tito, M. Atkins, and M. Apperley, "Predicting Steam Turbine Power Generation: A Comparison of Long Short-Term Memory and Willans Line Model," *Energies* 17, no. 2 (2024): 352.

25. D. Venkateswaran and Y. Cho, "Efficient Solar Power Generation Forecasting for Greenhouses: A Hybrid Deep Learning Approach," *Alexandria Engineering Journal* 91 (2024): 222–236.

26. A. P. Wibawa, A. F. Fadhilla, A. K. A. I. Paramarta, et al., "Bidirectional Long Short-Term Memory (Bi-LSTM) Hourly Energy Forecasting," in *E3S Web of Conferences* (Les Ulis: EDP Sciences, 2024).

27. A. L. F. Marques, M. J. Teixeira, F. V. De Almeida, and P. L. P. Corrêa, "Neural Networks Forecast Models Comparison for the Solar Energy Generation in Amazon Basin," *IEEE Access* 12 (2024): 17915–17925.

28. M. Abubakar, Y. Che, M. Faheem, M. S. Bhutta, and A. Q. Mudasar, "Intelligent Modeling and Optimization of Solar Plant Production Integration in the Smart Grid Using Machine Learning Models," *Advanced Energy and Sustainability Research* 5, no. 4 (2024): 2300160.

29. D. B. Unsal, A. Aksoz, S. Oyucu, J. M. Guerrero, and M. Guler, "A Comparative Study of AI Methods on Renewable Energy Prediction for Smart Grids: Case of Turkey," *Sustainability* 16, no. 7 (2024): 2894.

30. S. Hasan, I. U. Hossain, N. Hasan, I. B. Sakib, A. Hasan, and T. U. Amin, "Forecasting and Predictive Analysis of Source-Wise Power Generation Along With Economic Aspects for Developed Countries," *Energy Conversion and Management: X* 22 (2024): 100558.

31. T. Yan, J. Rashid, M. S. Saleem, S. Ahmad, and M. Faheem, "A Hybrid Deep Learning Approach for Green Energy Forecasting in Asian Countries," *Computers, Materials and Continua* 81 (2024): 2685–2708.

32. B. Wu, L. Wang, and Y. R. Zeng, "Interpretable Wind Speed Prediction With Multivariate Time Series and Temporal Fusion Transformers," *Energy* 252 (2022): 123990.

33. B. Wu and L. Wang, "Two-Stage Decomposition and Temporal Fusion Transformers for Interpretable Wind Speed Forecasting," *Energy* 288 (2024): 129728.

34. B. Lim, S. Ö. Arık, N. Loeff, and T. Pfister, "Temporal Fusion Transformers for Interpretable Multi-Horizon Time Series Forecasting," *International Journal of Forecasting* 37, no. 4 (2021): 1748–1764.

35. B. Wu, S. Yu, L. Peng, and L. Wang, "Interpretable Wind Speed Forecasting With Meteorological Feature Exploring and Two-Stage Decomposition," *Energy* 294 (2024): 130782.

36. B. Hossainds, "Dataset of Green Electricity Production 1965–2021," accessed October 8, 2024, https://www.kaggle.com/datasets/belayethossainds/renewable-energy-world-wide-19652022?select=03+modern-renewable-prod.csv.

37. "Ember, Dataset of Green Electricity Production 2022–2023," accessed October 8, 2024, https://ember-energy.org/data/electricity-data-explorer/.

38. F. T. Lima and V. M. A. Souza, "A Large Comparison of Normalization Methods on Time Series," *Big Data Research* 34 (2023): 100407.

39. Q. H. Nguyen, H. B. Ly, L. S. Ho, et al., "Influence of Data Splitting on Performance of Machine Learning Models in Prediction of Shear Strength of Soil," *Mathematical Problems in Engineering* 2021, no. 1 (2021): 1–15.

40. B. Ayhan, E. P. Vargo, and H. Tang, "On the Exploration of Temporal Fusion Transformers for Anomaly Detection With Multivariate Aviation Time-Series Data," *Aerospace* 11, no. 8 (2024): 646.

41. P. Zeng, G. Hu, X. Zhou, S. Li, P. Liu, and S. Liu, "Muformer: A Long Sequence Time-Series Forecasting Model Based on Modified Multi-Head Attention," *Knowledge-Based Systems* 254 (2022): 109584.

42. D. Li, Y. Tan, Y. Zhang, S. Miao, and S. He, "Probabilistic Forecasting Method for Mid-Term Hourly Load Time Series Based on an Improved Temporal Fusion Transformer Model," *International Journal of Electrical Power & Energy Systems* 146 (2023): 108743.

43. P. C. Huy, N. Q. Minh, N. D. Tien, and T. T. Q. Anh, "Short-Term Electricity Load Forecasting Based on Temporal Fusion Transformer Model," *IEEE Access* 10 (2022): 106296–106304.

44. S. F. Stefenon, L. O. Seman, L. S. A. da Silva, V. C. Mariani, and L. dos Santos Coelho, "Hypertuned Temporal Fusion Transformer for Multi-Horizon Time Series Forecasting of Dam Level in Hydroelectric Power Plants," *International Journal of Electrical Power & Energy Systems* 157 (2024): 109876.

45. S. I. Alma'asfa, F. Y. Fraige, M. S. Abdul Aziz, C. Y. Khor, and L. A. Al-Khatib, "Evaluating the Performance of the Anwaralardh Photovoltaic Power Generation Plant in Jordan: Comparative Analysis Using Artificial Neural Networks and Multiple Linear Regression Modeling," *International Journal of Renewable Energy Development* 13, no. 4 (2024): 608–617.

46. "Google Colaboratory, Welcome to Colaboratory," accessed October 8, 2024, https://colab.research.google.com/notebooks/intro.ipynb?utm_source%20=%20scs-index.

47. M. Faheem, M. A. Al-Khasawneh, A. A. Khan, and S. H. H. Madni, "Cyberattack Patterns in Blockchain-Based Communication Networks for Distributed Renewable Energy Systems: A Study on Big Datasets," *Data in Brief* 53, no. 5 (2024a): 110212, https://doi.org/10.1016/j.dib.2024.11021250.66.

48. B. Raza, Y. J. Kumar, A. K. Malik, A. Anjum, and M. Faheem, "Performance Prediction and Adaptation for Database Management System Workload Using Case-Based Reasoning Approach," *Information Systems* 76, no. 5 (2018): 46–58, https://doi.org/10.1016/j.is.2018.04.00551.67.

49. M. Faheem and M. A. Al-Khasawneh, "Multilayer Cyberattacks Identification and Classification Using Machine Learning in Internet of Blockchain (IoBC)-Based Energy Networks," *Data in Brief* 54, no. 5 (2024): 110461, https://doi.org/10.1016/j.dib.2024.110461.68.

50. M. Faheem, B. Raza, M. S. Bhutta, and S. H. H. Madni, "A Blockchain-Based Resilient and Secure Framework for Events Monitoring and Control in Distributed Renewable Energy Systems," *IET Blockchain* 4 (2024): 1–15, https://doi.org/10.1049/blc2.12081.69.

51. A. A. Khan, R. K. Madendran, U. Thirunavukkarasu, and M. Faheem, "D 2 PAM: Epileptic Seizures Prediction Using Adversarial Deep Dual Patch Attention Mechanism," *CAAI Transactions on Intelligence Technology* 8, no. 3 (2023): 755–769, https://doi.org/10.1049/cit2.12261.70.

52. M. Faheem, M. Umar, R. A. Butt, B. Raza, M. A. Ngadi, and V. C. Gungor, "Software Defined Communication Framework for Smart Gridto Meet Energy Demands in Smart Cities," in *7th International Istanbul Smart Grids and Cities Congress and Fair (ICSG)* (Istanbul, Turkey: ICSG, IEEE, 2019), 51–55.

53. M. A. S. Al-Khasawneh, M. Faheem, E. A. Aldhahri, A. Alzahrani, and A. A. Alarood, "A Mapreduce Based Approach for Secure Batch Satellite Image Encryption," *IEEE Access* 11 (2023): 62865–62878.

54. Y. Chen, M. S. Bhutta, M. Abubakar, et al., "Evaluation of Machine Learning Models for Smart Grid Parameters: Performance Analysis of ARIMA and Bi-LSTM," *Sustainability* 15, no. 11 (2023): 8555.