

Electrical Load Forecasting in Developing Regions: Khost Province Considering Socio-Economic and Environmental Factors

Fazal Hamid Habibi^{1*}, Habiburrahman Shirani², Atiqullah Hamim³

¹ Assistant Professor, Electrical Power Engineering Department, Shaikh Zayed University, Khost, Afghanistan

² Professor, Electrical and Electronic Engineering Department, Kabul University, Kabul, Afghanistan

³ Assistant Professor, Electrical Power Engineering Department, Ghazni Technical University, Ghazni, Afghanistan

ABSTRACT: Accurate long-term electricity load forecasting is essential for strategic power system planning but remains challenging in developing regions like Khost, Afghanistan, due to frequent load shedding and data limitations. This study develops a comprehensive forecasting framework using a Multivariate Long Short-Term Memory (LSTM) model. To mitigate data distortions caused by supply constraints, a Standard Load Profile (SLP) was constructed to represent the underlying electricity demand accurately. The model integrates historical consumption with five key explanatory variables: population growth, Gross Domestic Product GDP development, ambient temperature, humidity, and public holidays. Correlation analysis identified temperature as the primary driver of demand, followed by economic growth. The optimized LSTM architecture, featuring two layers (128 and 64 neurons), achieved a Mean Absolute Percentage Error (MAPE) of 3.67%, demonstrating high predictive accuracy. Furthermore, demand was forecasted for the period 2026–2030 under Base, High-Growth, and Low-Growth scenarios. Monte Carlo simulations were employed to generate probabilistic forecasts (P50, P90, and P95), providing a robust quantification of future uncertainty. The findings indicate a consistent rise in future demand, offering a reliable scientific basis for sustainable energy policy and risk-informed infrastructure investment in Khost Province.

KEYWORDS: Long-Term Electricity Load Forecasting, Multivariate LSTM, Standard Load Profile, Scenario-Based Forecasting, Monte Carlo Simulation, Probabilistic Forecasting

INTRODUCTION

Electricity serves as a fundamental infrastructure driver for economic growth, industrial expansion, and overall social welfare. The increasing demand for electrical energy, propelled by rapid urbanization and demographic shifts, has made power system planning more challenging than ever, necessitating accurate Long-Term Load Forecasting (LTLF) to guide strategic infrastructure investment and energy policy formulation (S. S. Arnob, 2023). Load forecasting is generally classified based on the time horizon into three categories, each serving specific operational and planning objectives: Short-Term Load Forecasting (STLF), ranging from one hour to one week, is essential for daily operational activities like unit commitment and economic dispatch (J. Son, 2022), (Heiko Hahn, 2009); Medium-Term Load Forecasting (MTLF), spanning one week to one year, supports tactical decisions such as maintenance scheduling and fuel procurement (H.M. Al-Hamadi, 2005), (Heiko Hahn, 2009); and Long-Term Load Forecasting (LTLF), which extends beyond one year, provides the foundation for strategic investments in generation plants and transmission networks (Arunesh Kumar Singh, 2012), (Heiko Hahn, 2009).

The evolution of forecasting methodologies is well-documented in the literature. Early research relied heavily on statistical techniques, such as the Particle Swarm Optimization (PSO) and Regression models used in Egypt, which proved more accurate than traditional least error square methods (EL-Naggar, 2010). In North Carolina, researchers utilized Multiple Linear Regression (MLR) with hourly data, demonstrating that high-resolution temporal information significantly improves long-term accuracy compared to monthly methods (T. Hong, Jan. 2014). As the field moved toward intelligent systems, studies in Nigeria achieved 93.1% accuracy using Fuzzy Logic to handle non-linear relationships (Garkida, 2016). Recently, deep learning has emerged as a dominant force; for instance, an LSTM-RNN model in the ISO New England market achieved a MAPE of 6.54% using hourly load data (R. K. Agrawal, 2018). Comparative studies in Turkey found that Multivariate Adaptive Regression Splines (MARS) could achieve up to 99% accuracy (Nalcaci, 2019), while hybrid approaches in West Canada, such as LSTM-GRU frameworks, have outperformed traditional methods by combining the strengths of different architectures (M. Dong and L. Grumbach, A Hybrid Distribution Feeder Long-Term Load Forecasting Method Based on Sequence Prediction, 2020).

Despite these advancements, a significant research gap remains. Most existing models are designed for developed nations with advanced smart-grid infrastructures and high-quality historical data. In developing regions like Afghanistan, and specifically Khost Province, the electricity sector faces unique hurdles, including volatile economic shifts and irregular historical records marred by frequent power outages and supply constraints. Industrialized models often fail to account for the socio-economic and climatic nuances of these areas, leading to inefficient energy management and unstable supply.

To address these critical limitations, this study develops a context-specific Multivariate LSTM forecasting framework tailored for Khost Province. A primary contribution of this research is the development of a Standard Load Profile (SLP) to overcome raw data distortions caused by irregular load shedding, thereby representing underlying demand more accurately. The framework integrates five key explanatory variables: population growth, GDP development, ambient temperature, humidity, and public holidays. By analyzing these drivers, the model provides a reliable scientific basis for future power system expansion and risk-informed infrastructure investment.

METHODOLOGY

Data collection and data preprocessing: Khost Province, located in southeastern Afghanistan, was selected as the geographical focus of this study. The province covers an area of approximately 4,118–4,152 km² and experiences distinct seasonal variations, with higher summer temperatures contributing to increased electricity demand due to cooling requirements [12]. Electricity services in the province are managed by Da Afghanistan Breshna Sherkat (DABS), serving residential, commercial, and industrial consumers [13].

A multi-source data integration approach was adopted to develop a comprehensive multivariate dataset for long-term electricity load forecasting. Historical electricity consumption data for the period 2021–2025 were obtained from the Khost Electricity Directorate (DABS) [13]. Monthly average temperature and relative humidity data were obtained from the NASA POWER database [14]. Demographic and economic variables, including population growth and GDP growth, were obtained from the National Statistics and Information Authority (NSIA), the World Bank, and the International Monetary Fund (IMF) [15], [16]. In addition, official public holiday information was obtained from the Ministry of Labor and Social Affairs to account for calendar-related variations in electricity demand [17]. The variables, units, temporal resolution, and corresponding data sources used in the forecasting model are summarized in Table 1.

Table 1: Description of Variables Used in the Forecasting Model.

Variable	Description	Unit	Temporal Resolution	Data Source
Electricity Load	Standardized electricity demand	MWh	Monthly	DABS
Temperature	Average air temperature	°C	Monthly	NASA POWER
Relative Humidity	Average relative humidity	%	Monthly	NASA POWER
Population Growth	Population growth rate	%	Annual	NSIA / World Bank / IMF
GDP Growth	GDP growth rate	%	Annual	NSIA / World Bank / IMF
Total Holidays	Number of public holidays	Days	Monthly	Ministry of Labor and Social Affairs

Before model development, the collected data were preprocessed to ensure consistency and suitability for multivariate LSTM forecasting. The (SLP) was developed to reduce the influence of irregular load shedding and supply interruptions in the historical electricity records. The SLP was estimated by considering the electricity demand of residential, commercial, and industrial consumers based on standard connected loads and appropriate demand and diversity factors. The estimated sectoral loads were then aggregated to obtain a representative electricity load profile as illustrated in Figure 1, which was subsequently combined with

meteorological, demographic, economic, and calendar-related variables. The resulting dataset was complete and contained no missing observations; therefore, no missing-value imputation was required.

Since the input variables had different units and numerical ranges, Min–Max normalization was applied before model training. This procedure transformed the input variables into a range between 0 and 1, helping to improve numerical stability and facilitate the learning process of the LSTM model. The resulting dataset contained 60 consecutive monthly observations, which were converted into sequential samples using a sliding-window approach with a 12-month time step. To preserve the temporal structure of the data, the observations were divided chronologically into 42 periods (70%) for training and 18 periods (30%) for testing. The training data were used to learn the relationships between electricity demand and the explanatory variables, while the testing data were reserved for evaluating the model's predictive performance and its ability to generalize to previously unseen observations.

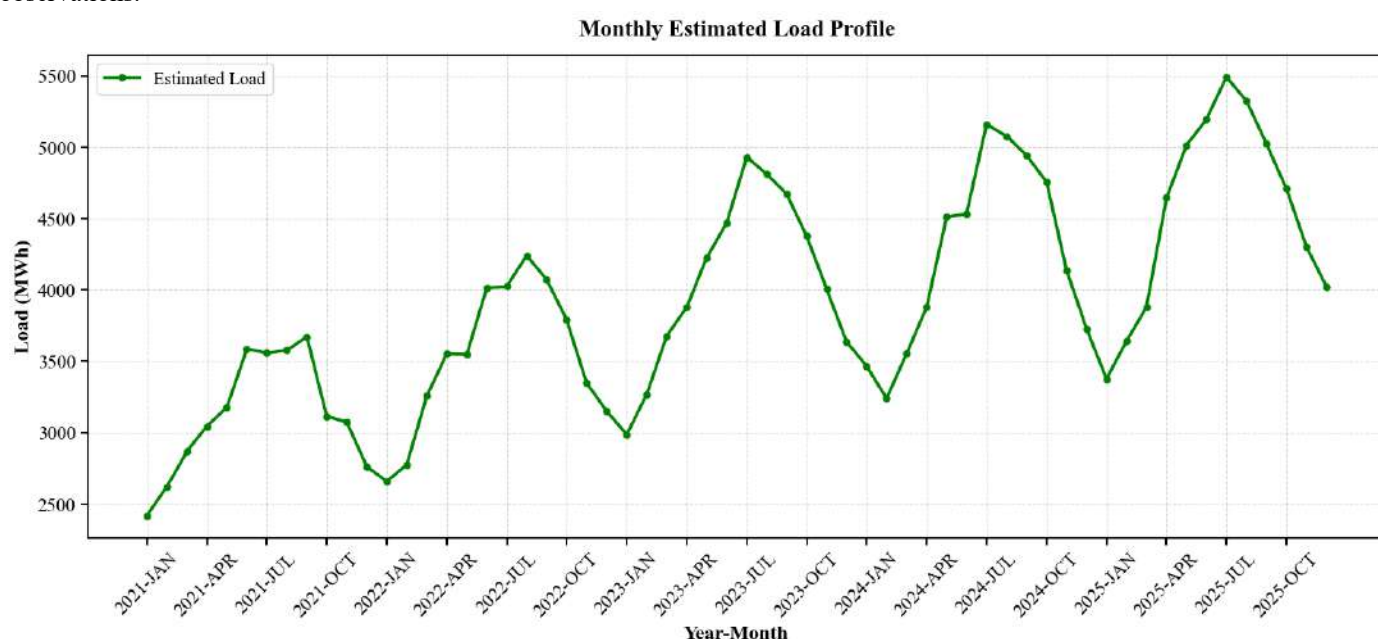


Figure 1. presents the monthly Standard Load Profile (SLP) developed for Khost Province based on the estimated electricity demand data.

Selection of Input Variables and Their Correlation with Target Variable: five key explanatory variables were integrated: population growth, (GDP) development, ambient temperature, relative humidity, and public holidays. These variables were chosen based on their established correlation with load fluctuations, where socioeconomic factors like population and GDP account for long-term growth trends, while climatic factors and calendar effects capture seasonal and short-term operational variations.

As illustrated in figure 2, correlation analysis identifies temperature as the primary driver of electricity demand (0.73), followed by significant positive correlations with GDP growth (0.64) and humidity (0.59). These results statistically validate the selection of these factors as the core input variables for the forecasting model.

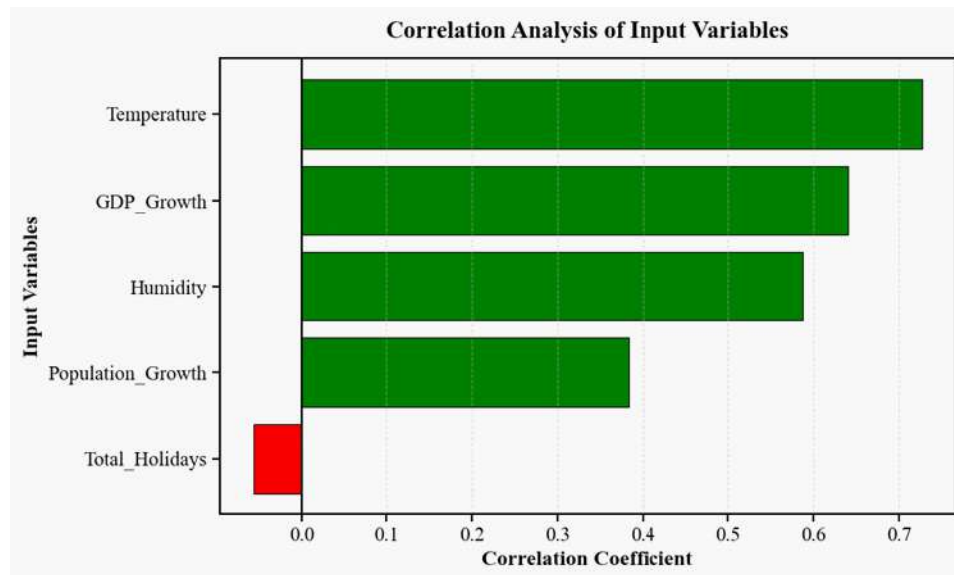


Figure 2. Correlation coefficients between the selected input variables and monthly electricity load demand in Khost Province.

Deep Learning Models for Long-Term Load Forecasting

Recurrent Neural Network (RNN): The (RNN) represents a class of artificial neural networks specifically engineered for processing sequential data by establishing temporal connections between hidden layers. As illustrated in Figure 3, the RNN architecture can be visualized as an unfolded series of feedforward networks where the hidden state at the current time step (h_t) is a function of both the current input (X_t) and the hidden state from the previous step (h_{t-1}). This recurrent mechanism allows the network to maintain a "memory" of historical consumption trends, which is vital for long-term load forecasting where current demand is often influenced by past socioeconomic momentum. However, standard RNNs frequently encounter the vanishing or exploding gradient problem during backpropagation, which hinders their ability to capture long-range dependencies over extended time horizons.

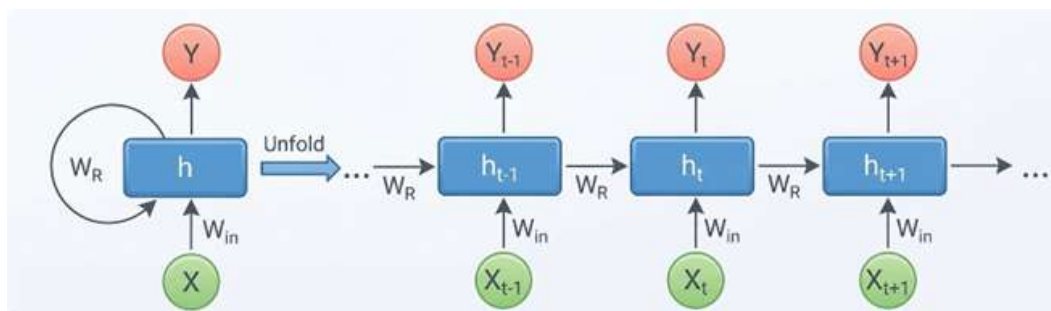


Figure 3. Illustration of an unfolded RNN [17].

Long Short Term Memory (LSTM): The Long Short-Term Memory (LSTM) network overcomes the limitations of conventional RNNs by incorporating a cell state that selectively retains or removes information based on its relevance. As illustrated in Figure 3, an LSTM consists of (S) sequential cells, corresponding to the input time steps, with (C) input features and (D) hidden units. Each cell contains three gates—forget, input, and output—that regulate information flow through weights, biases, and activation functions, enabling the network to capture long-term temporal dependencies.

Forget gate: This gate determines what information from the cell state c_{t-1} (the top horizontal line in Fig. 4 colored in orange) should be thrown away using information from the previous hidden state h_{t-1} and the current input x_t . The current cell input x_t is multiplied by the weight matrix W_f whereas the previous hidden-state h_{t-1} is multiplied by the recurrent weight matrix R_f . The resulting output of these products are added to a bias vector b_f . Finally, a sigmoid function σ_g is activated to get the output vector

f_t that has values varying between 0 and 1. The value “0” means no information from the previous time-step of cell state is allowed to flow (not important information), whereas the value “1” means all previous information of the memory is allowed to flow (extremely important). If the information is partially relevant, the function outputs a value between “0” and “1”. Mathematically, this description can be written as follows:

$$X_t = \sigma_g(W_f x_t + R_f h_{t-1} + b_f) \quad (1)$$

where σ_g denotes the gate’s activation function, and all the other parameters and variables are defined above. If the input variable x_t is a vector of sequences with C features, and each cell has D hidden units, then the weights W_f and R_f are matrices with the dimensions of $D \times C$ and $D \times D$, respectively, whereas the bias b_f is a vector with D elements. As a result, the output of the gate f_t is a vector with D elements.

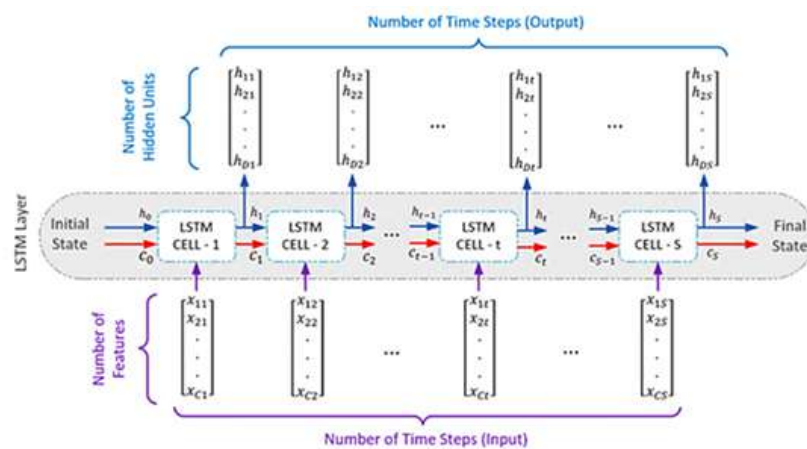


Figure 4. An LSTM layer with multi-inputs and multi-outputs [17].

Update gate: This gate is used to update the cell state or memory that was regulated by the forget gate in the previous step. It is composed of two parts of neural networks: input gate i_t and candidate cell g_t , and are fed with the same inputs used for the forget-gate (h_{t-1} and x_t). However, the weights, biases, and activation functions of i_t and g_t branches are different. For the input gate branch– i_t , we have the input variables x_t and h_{t-1} weighted by the matrices W_i and R_i , respectively and biased with b_i , and finally activated using a sigmoid function σ_g . The same is repeated for the candidate-state branch – g_t using the denoting letter i instead of g , and replacing the sigmoid function σ_g with a tan hyperbolic ($\tanh$ or σ_s) to squishes the data between -1 and 1 . The input branch is used to control the output of the squished data–the candidate state. Finally, the outputs of these two neural networks f_t and g_t are multiplied to produce the output of the update gate. Mathematically, the two networks can be written as follows:

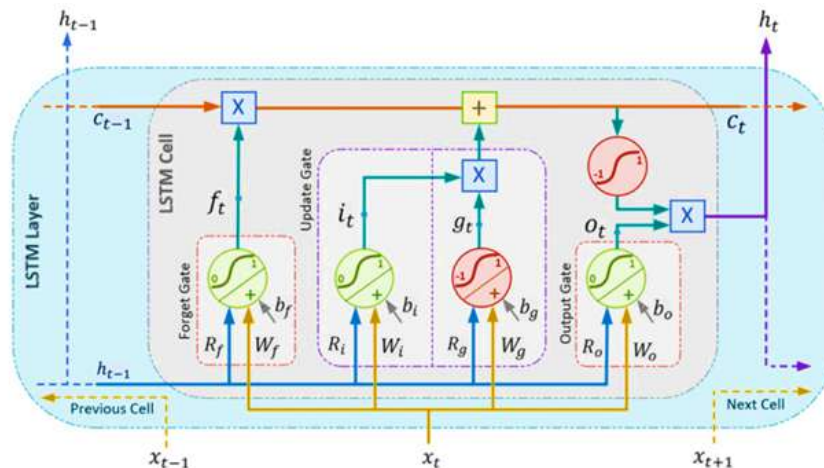


Figure 5. Internal structure of an LSTM cell [18].

$$i_t = \sigma_g(W_i x_t + R_i h_{t-1} + b_i) \quad (2)$$

$$g_t = \sigma_s(W_g x_t + R_g h_{t-1} + b_g) \quad (3)$$

where σ_s denotes the state activation function.

Output gate: This gate is used to compute the current hidden state h_t . We pass a copy of the combined input (x_t and h_{t-1}) to a sigmoid function σ_g after multiplied with the respective weights w_o and R_o , and added to the bias b_o . The resulting output o_t is multiplied with the current cell state c_t after squished to the range $[-1, 1]$ using the tanh function σ_s . Mathematically, the output gate can be described as follows:

$$o_t = \sigma_g(W_o x_t + R_o h_{t-1} + b_o) \quad (4)$$

The equations of the new cell-state c_t and hidden-state h_t are:

$$c_t = f_t * c_{t-1} + i_t * g_t \quad (5)$$

$$h_t = o_t * \sigma_s(c_t) \quad (6)$$

To summarize, the forget gate determines what information from the old memory is relevant to keep and forget the irrelevant ones. The input gate is used to update the relevant memory and generate the current memory used in the next block. The output gate is used to compute the output of the current block and the next hidden-state. We should note that all the gates have the same inputs consisting of three copies of the previous hidden state and the current input combined (the bottom line in Fig. 5). The top of Fig. 5 is the LSTM memory or cell state that is used by the network to learn about the sequence order of input data.

Model Development, Forecasting Scenarios, and Performance Evaluation: Following data preprocessing and model development, the predictive performance of the Multivariate LSTM model was optimized through an iterative hyperparameter tuning process. Different model configurations were systematically evaluated by varying the number of LSTM units, dropout rates, learning rates, and batch sizes over multiple training epochs. The selected configuration was based on the predictive performance obtained during model evaluation.

To investigate potential future electricity demand pathways, a scenario-based forecasting framework was developed for the period 2026–2030. Three scenarios were considered: Base, High Growth, and Low Growth, with their corresponding assumptions summarized in Table 2. The Base Scenario represents the reference trajectory under the assumption that historical demographic and economic trends continue under business-as-usual conditions. The High Growth Scenario represents an accelerated development pathway characterized by rapid electrification and stronger economic expansion, whereas the Low Growth Scenario represents conservative socioeconomic growth and reduced investment activity. These scenarios provide alternative future demand pathways and allow the influence of changes in population and GDP growth on electricity demand to be examined.

Table 2: Assumptions Used for Scenario-Based Electricity Load Forecasting (2026–2030).

Variable	Base Scenario	High-Growth Scenario	Low-Growth Scenario
Population Growth	Historical growth trend maintained ($\approx 2.0\%$)	Higher population growth ($\approx 2.5\%$)	Lower population growth ($\approx 1.5\%$)
GDP Growth	Moderate economic growth ($\approx 1.8\%$)	Accelerated economic growth ($\approx 3.5\%$)	Slow economic growth ($\approx 1.0\%$)
Temperature	Historical monthly averages	Historical monthly averages	Historical monthly averages
Humidity	Historical monthly averages	Historical monthly averages	Historical monthly averages
Public Holidays	Official calendar	Official calendar	Official calendar
Electricity Demand	Expected growth under current conditions	Faster demand growth due to increased economic activity and electrification	Slower demand growth due to reduced economic expansion

In addition to deterministic scenario-based forecasting, a probabilistic forecasting approach was employed to quantify uncertainty in the long-term projections. A Monte Carlo simulation with 1,000 simulation runs was conducted by introducing random variations to selected input variables based on the specified distributions and historical reference values. This procedure generated a range of possible future load trajectories rather than a single deterministic forecast. Quantile forecasts were subsequently

derived at the P50, P90, and P95 levels. The P50 represents the median forecast, while the P90 and P95 provide higher-demand estimates that can be used to assess potential system adequacy and infrastructure requirements under uncertain future conditions. The assumptions used in the probabilistic simulations are summarized in Table 3.

Table 3: Assumptions Used in the Probabilistic Forecasting Analysis.

Variable	Distribution	Mean / Reference Value	Variation Applied
Population Growth	Normal	Base Scenario value	±5%
GDP Growth	Normal	Base Scenario value	±5%
Temperature	Normal	Monthly historical mean	±3%
Humidity	Normal	Monthly historical mean	±3%

Finally, the forecasting performance of the proposed Multivariate LSTM model was evaluated using three standard statistical error metrics: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). These metrics measure the differences between the actual and predicted electricity demand from different perspectives. RMSE gives greater weight to larger errors because of its squared-error formulation, MAE represents the average absolute deviation between observed and predicted values, and MAPE expresses the forecasting error as a percentage, facilitating interpretation across different demand levels. The mathematical formulations of these metrics are presented below:

1. Root Mean Square Error (RMSE):

$$RMSE = \sqrt{\left[\frac{1}{n} \sum_{i=1}^n (y_i - \tilde{y}_i)^2 \right]} \quad (7)$$

where

n Is the total number of observations, y_i The actual electricity load value, $\tilde{y}_i$ Forecasted electricity load value and i Observation index.

2. Mean Absolute Error (MAE):

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \tilde{y}_i| \quad (8)$$

3. Mean Absolute Percentage Error (MAPE):

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{y_i - \tilde{y}_i}{y_i} \right| \quad (9)$$

RESULTS AND DISCUSSION

Performance Evaluation of the Multivariate LSTM Model: The predictive performance of the Multivariate LSTM model was evaluated based on training convergence, hyperparameter optimization, and standard error metrics. The training process showed stable convergence, as indicated by the training and validation loss curves, suggesting that the model was able to learn the underlying patterns in the data without substantial overfitting. A total of eight model configurations were systematically evaluated by varying the LSTM architecture and other hyperparameters. The performance results of the eight configurations are presented in Table 4.

Table 4: Hyperparameter configurations used for the development and optimization of the Multivariate LSTM forecasting models.

Model	LSTM_ Units	Dropout	Learning rate	Batch size	Epoch	MAE[MWh]	RMSE[MWh]	MAPE[%]
Model 1	64_32	0.2	0.001	4	150	345.14	416.21	7.21
Model 2	128_64	0.15	0.0005	4	200	383.67	433.52	8.07
Model 3	128_128	0.1	0.0003	2	250	372.16	426.23	7.7
Model 4	64_32	0.05	0.01	4	250	260.95	302.09	5.38
Model 5	128_64	0.05	0.01	4	250	221.32	292.78	4.62
Model 6	64_32	0.2	0.003	4	250	211.61	285.84	4.62
Model 7	64_32	0.15	0.003	4	250	191.66	230.09	4.32
Model 8	128_64	0.1	0.003	4	250	171.8	236.86	3.67

Among the evaluated configurations, Model 8, which employed a stacked LSTM architecture with 128 and 64 neurons, achieved the lowest prediction errors, with a MAPE of 3.67%, MAE of 171.8, and RMSE of 236.86. These results indicate that the selected architecture provided a good representation of the nonlinear relationships between electricity demand and the incorporated meteorological, demographic, economic, and calendar-related variables. The relatively low MAPE also indicates that the model produced small deviations from the observed electricity demand during the testing period. The training and validation loss curves, together with the prediction performance of the selected model, are presented in Figure 6.

The comparison of the eight configurations further demonstrates that changes in the LSTM architecture and training parameters affected the forecasting performance. The selection of Model 8 was therefore based on its error metrics and convergence behavior rather than on model complexity alone. The corresponding actual-versus-predicted electricity demand plot provides an additional visual assessment of the model's ability to reproduce the observed demand pattern.

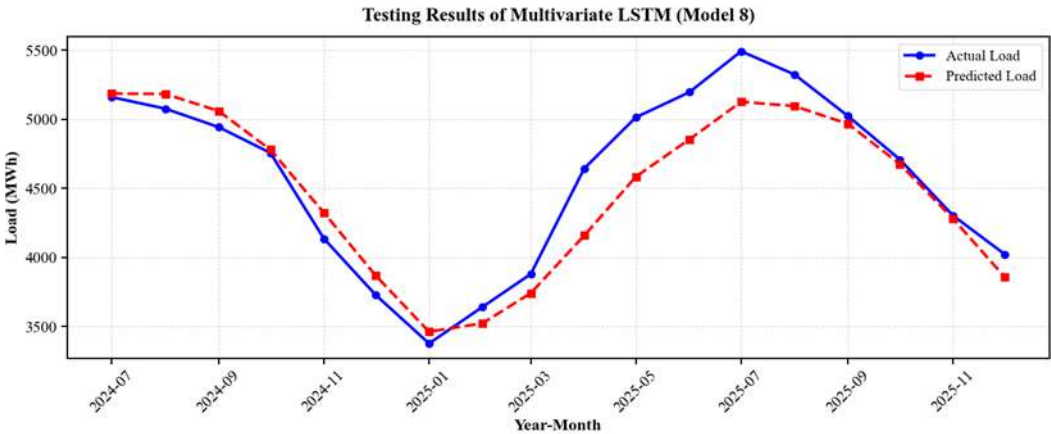


Figure 6. Actual and predicted electricity load profiles for Model 8 during the testing period

Scenario-Based Forecasting Results: The long-term electricity demand for Khost Province was forecasted for the 2026–2030 period under three distinct socioeconomic pathways: Base, High-Growth, and Low-Growth scenarios, as illustrated in Figure 7. Technical analysis indicates that all trajectories maintain a consistent seasonal cycle, characterized by recurring annual peaks during the summer months and troughs during the spring and winter seasons. Under the Base Scenario, which reflects business-as-usual conditions, the peak demand is projected to reach approximately 5,700 MWh by 2030. In contrast, the High-Growth Scenario—assuming accelerated economic activity and rapid electrification—projects a more substantial surge, with peak demand approaching 6,900 MWh by the end of the forecasting horizon. Conversely, the Low-Growth Scenario serves as a conservative lower-bound

estimate, projecting a peak demand of nearly 5,000 MWh by 2030. This divergence toward 2030 highlights the cumulative impact of socioeconomic drivers on the region's long-term energy baseline.

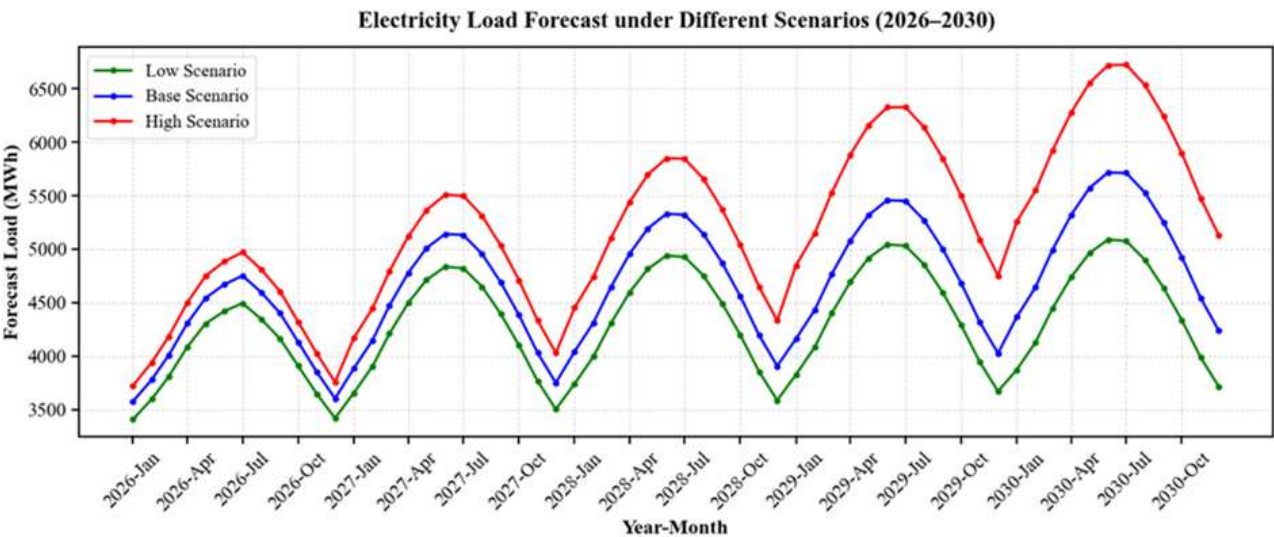


Figure 7. Comparative analysis of scenario-based electricity load forecasts under Low Growth, Base, and High Growth scenarios for Khost Province during the period 2026–2030.

Probabilistic Forecasting Results: The results presented in figure 8 show that the median forecast (P50) follows a consistent upward trajectory, with demand projected to rise from approximately 3,700 MWh in early 2026 to over 5,700 MWh by 2030. This median path remains the most probable reference for future requirements under standard development conditions.

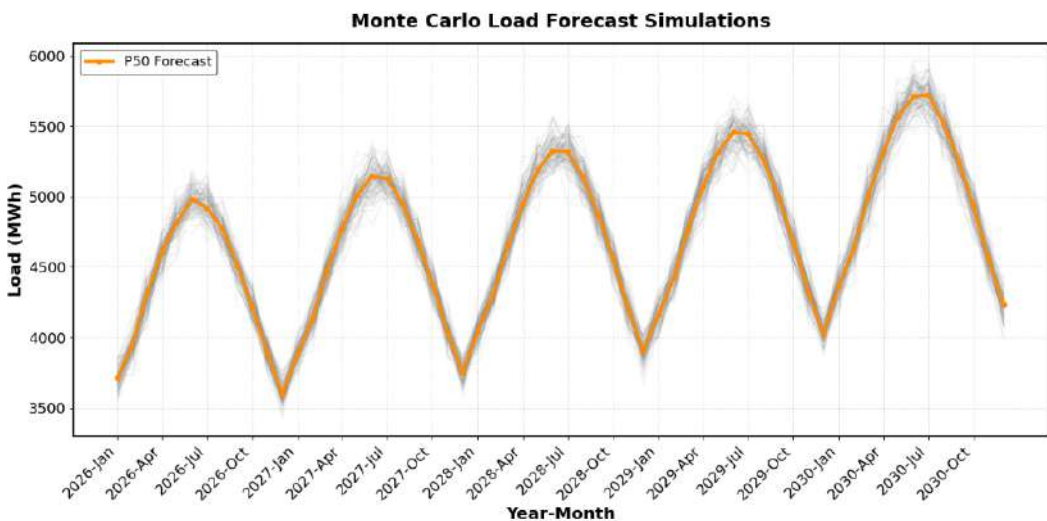


Figure 8. Monte Carlo simulation results for monthly electricity load forecasting in Khost Province (2026–2030), showing multiple simulated demand trajectories and the median (P50) forecast path.

The integration of Quantile Forecasts (P50, P90, and P95) illustrated in Figure 9, provides critical insights for risk-informed infrastructure planning. While the P50 reflects the expected demand, the P90 and P95 quantiles represent conservative upper-bound estimates, where demand levels have only a 10% and 5% probability of being exceeded, respectively.

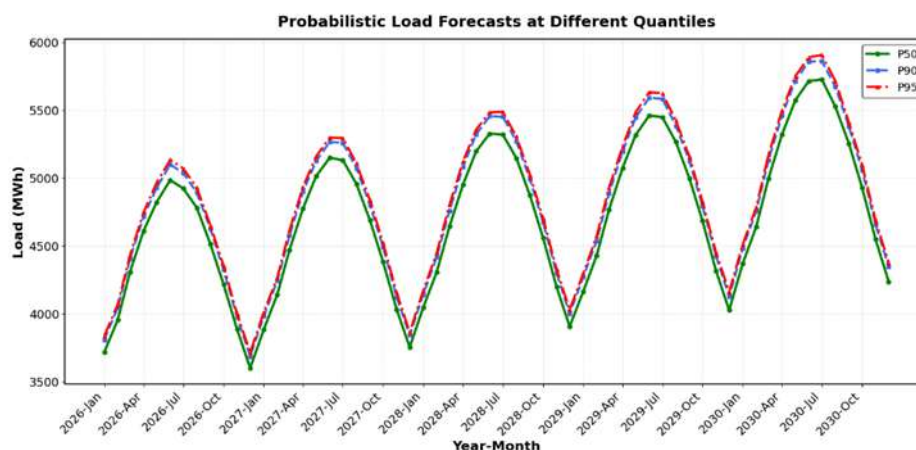


Figure 9. Quantile-based electricity load forecasts (P50, P90, and P95) for Khost Province during the period 2026–2030.

The associated Uncertainty Analysis reveals a progressive expansion of the prediction intervals over the forecasting horizon. The results presented in Figure X, Visualization of the 10%–90% and 5%–95% confidence intervals demonstrates that while short-term projections (2026) exhibit high confidence and narrow dispersion, the range of possible outcomes widens significantly toward 2030.

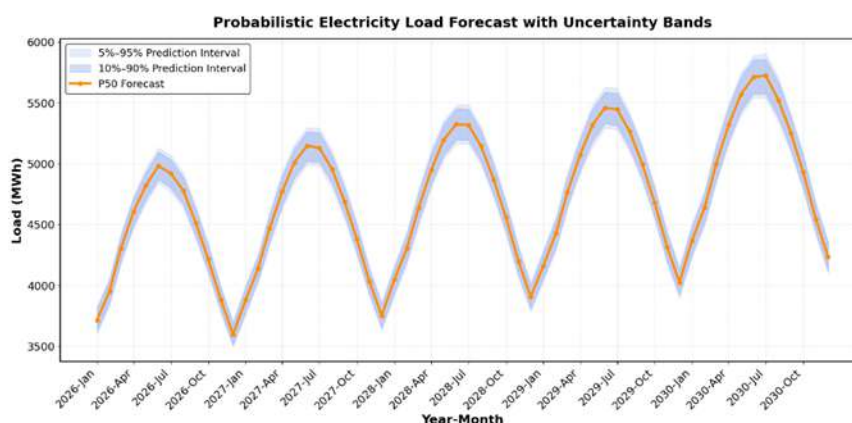


Figure 10. Probabilistic electricity load forecast with 10%–90% and 5%–95% confidence intervals for Khost Province (2026–2030).

CONCLUSION

This study successfully developed a robust, context-specific long-term load forecasting framework for Khost Province, Afghanistan, utilizing a Multivariate Long Short-Term Memory (LSTM) approach. A pivotal contribution was the creation of a Standard Load Profile (SLP), which effectively mitigated historical data distortions caused by frequent load shedding and technical outages, ensuring the model learned true underlying consumption patterns.

The technical evaluation demonstrated the superiority of the deep learning architecture, with the optimized model achieving a Mean Absolute Percentage Error (MAPE) of 3.67%. This high level of precision confirms that the Multivariate LSTM is significantly more effective at capturing non-linear temporal dependencies than traditional statistical methods in data-limited developing regions. Correlation analysis further identified ambient temperature as the most critical driver of demand (0.73 correlation), followed by GDP growth (0.64) and humidity (0.59), revealing a pronounced seasonal cycle where summer cooling requirements place the most substantial pressure on the provincial grid.

The scenario-based projections for 2026–2030 provide a clear roadmap for future energy needs, indicating that peak demand could surge to nearly 6,900 MWh under an accelerated growth pathway. Furthermore, the integration of Monte Carlo

simulations provided a robust quantification of future uncertainty, offering the necessary evidence for risk-informed infrastructure investment. Ultimately, this research provides a scientific foundation for (DABS) to transition toward AI-driven planning, ensuring sustainable power system expansion and resilient energy management in the region.

REFERENCES

1. S. S. Arnob and A. I. M. S. Arefin, "Energy Demand Forecasting and Optimizing Electric Systems for Developing Countries," *IEEE Access*, vol. 11, pp. 39751–39775, 2023, doi: 10.1109/ACCESS.2023.3250110.
2. J. Son and J. C.-M., "Day-Ahead Short-Term Load Forecasting for Holidays Based on Modification of Similar Days' Load Profiles," *IEEE Access*, vol. 10, pp. 17864–17880, 2022, doi: 10.1109/ACCESS.2022.3150344.
3. H. Hahn and S. Meyer-Nieberg, "Electric Load Forecasting Methods: Tools for Decision Making," *European Journal of Operational Research*, pp. 902–907, 2009.
4. H. M. Al-Hamadi and S. S., "Long-Term/Mid-Term Electric Load Forecasting Based on Short-Term Correlation and Annual Growth," *Electric Power Systems Research*, vol. 74, no. 3, pp. 353–361, 2005, doi: 10.1016/j.epsr.2004.10.015.
5. I. S. Arunesh Kumar Singh, "Load Forecasting Techniques and Methodologies: A Review," in *2nd International Conference on Power, Control and Embedded Systems*, pp. 631–641, 2012.
6. M. A. EL-Naggar, "Long Term Electric Load Forecasting Based on Particle Swarm Optimization," *Applied Energy*, pp. 320–326, 2010, doi: 10.1016/j.apenergy.2009.04.024.
7. T. Hong and J. W., "Long Term Probabilistic Load Forecasting and Normalization With Hourly Information," *IEEE Transactions on Smart Grid*, vol. 5, no. 1, pp. 456–462, Jan. 2014, doi: 10.1109/TSG.2013.2274373.
8. D. A. Garkida, "Long-Term Load Forecast Modelling Using a Fuzzy Logic Approach," *Pacific Science Review A: Natural Science and Engineering*, vol. 18, pp. 123–127, 2016, doi: 10.1016/j.psra.2016.09.011.
9. R. K. Agrawal and F. M., "Long Term Load Forecasting With Hourly Predictions Based on Long Short-Term Memory Networks," in *IEEE Texas Power and Energy Conference (TPEC)*, College Station, TX, USA, 2018, doi: 10.1109/TPEC.2018.8312088.
10. G. Ö. Nalcaci, "Long-Term Load Forecasting: Models Based on MARS, ANN and LR Methods," *Central European Journal of Operations Research*, vol. 27, pp. 1033–1049, 2019, doi: 10.1007/s10100-018-0531-1.
11. M. Dong and L. Grumbach, "A Hybrid Distribution Feeder Long-Term Load Forecasting Method Based on Sequence Prediction," *IEEE Transactions on Smart Grid*, vol. 11, pp. 470–482, Jan. 2020, doi: 10.1109/TSG.2019.2924183.
12. National Statistics and Information Authority (NSIA), *Statistical Yearbook 2022–23*. Kabul, Afghanistan, 2023.
13. Da Afghanistan Breshna Sherkat (DABS), "Balanced Growth in Expanding Electricity Networks Nationwide: CEO Message," Jun. 15, 2026. [Online]. Available: <https://dabs.af>.
14. International Monetary Fund (IMF), "IMF DataMapper." [Online]. Available: IMF DataMapper. [Accessed: 2026].
15. National Aeronautics and Space Administration (NASA), "NASA POWER Data Access Viewer: Prediction of Worldwide Energy Resources (POWER)." [Online]. Available: <https://power.larc.nasa.gov/data-access-viewer/>. [Accessed: Mar. 15, 2026].
16. National Statistics and Information Authority (NSIA), *Statistical Yearbook 2022–23*. Kabul, Afghanistan, 2023.
17. Ministry of Labor and Social Affairs (MoLSA), Mar. 15, 2026.
18. Z. A. Hamad, "Deep Learning-Based Load Forecasting Considering Data Reshaping Using MATLAB/Simulink," *International Journal of Energy and Environmental Engineering*, vol. 13, pp. 853–869, 2022, doi: 10.1007/s40095-022-00480-x.
19. M. Dong and L. Grumbach, "A Hybrid Distribution Feeder Long-Term Load Forecasting Method Based on Sequence Prediction," *IEEE Transactions on Smart Grid*, vol. 11, pp. 470–482, 2020, doi: 10.1109/TSG.2019.2924183.

Cite this Article: Habibi, F.H., Shirani, H., Hamim, A. (2026). *Electrical Load Forecasting in Developing Regions: Khost Province Considering Socio-Economic and Environmental Factors. International Journal of Current Science Research and Review*, 9(10), pp. 5456-5466. DOI: <https://doi.org/10.47191/ijcsrr/V9-i10-24>