

MULTIVARIATE LSTM-BASED FORECASTING FOR RENEWABLE ENERGY: ENHANCING CLIMATE CHANGE MITIGATION

Farshid Kamrani

Department of Mechanical and Aerospace
Engineering
Carleton University
Ottawa, ON, Canada
FarshidKamrani@cmail.carleton.ca

Kristen Schell

Department of Mechanical and Aerospace
Engineering
Carleton University
Ottawa, ON, Canada
KristenSchell@cunet.carleton.ca

ABSTRACT

The increasing integration of renewable energy sources (RESs) into modern power systems presents significant opportunities but also notable challenges, primarily due to the inherent variability of RES generation. Accurate forecasting of RES generation is crucial for maintaining the reliability, stability, and economic efficiency of power system operations. Traditional approaches, such as deterministic methods and stochastic programming, frequently depend on representative scenarios generated through clustering techniques like K-means. However, these methods may fail to fully capture the complex temporal dependencies and non-linear patterns within RES data. This paper introduces a multivariate Long Short-Term Memory (LSTM)-based network designed to forecast RESs generation using their real-world historical data. The proposed model effectively captures long-term dependencies and interactions between different RESs, utilizing historical data from both local and neighboring areas to enhance predictive accuracy. In the case study, we showed that the proposed forecasting approach results in lower CO_2 emissions, and a more reliable supply of electric loads.

1 INTRODUCTION

Renewable energy sources (RESs) have garnered significant attention in recent years due to their environmentally friendly nature and declining costs [1]. RESs, such as solar and wind power, play a critical role in modern power systems [2]. However, their inherent variability poses considerable challenges to the reliable operation of these systems. Accurate energy generation forecasting is essential for balancing power supply and demand and maintaining grid stability [3]. Traditional methods for addressing these challenges often rely on either deterministic approaches [2][3] or stochastic programming [4]. In the literature, K-means clustering has been widely employed to reduce large historical datasets into a smaller set of representative days that capture the uncertainty inherent in RESs. These representative days can serve as scenarios for stochastic programming, enabling more robust decision-making in power system operations [4].

With the advancement of Machine Learning (ML), its applications have expanded across a wide range of disciplines. Authors in [5] proposed a new joint dimensionality reduction (DR) and K-means clustering approach to recover the ‘clustering-friendly’ latent representations and to better cluster the data. Recurrent Neural Networks (RNNs), particularly suited for sequential data, are commonly used to capture latent features and predict future time intervals. These models have proven effective in various time series prediction tasks [6]. The authors in [7] proposed combining recurrent neural networks with a dimension-reducing symbolic representation for time series forecasting. This approach addresses issues like sensitivity to hyperparameters and random weight initialization, while enabling faster training without sacrificing forecasting performance. The authors in [8] proposed a hybrid model combining LSTM and multi-head attention for accurate time series forecasting.

Holland et al. [9] examined why marginal carbon emissions in the U.S. electricity sector have not declined despite increasing renewable energy adoption. The study estimates marginal emissions across different regions and time periods, finding that shifts in the energy mix, particularly the reliance on natural gas rather than renewables for marginal generation, play a significant role. The authors discuss the implications of these findings for climate policy, suggesting that carbon pricing and more targeted policies are needed to ensure emissions reductions. Building on these findings, accurate forecasts of RESs generation are essential for addressing this challenge. Reliable generation forecasts enable better grid management, enhance the integration of renewables, and reduce reliance on carbon-intensive generation.

The authors in [10] proposed a method for multi-decadal wind power forecasting that incorporates turbine locations into Gaussian Process-based downscaling, enabling accurate aggregate wind power predictions from the low-resolution CMIP6 climate model data. The authors in [11] investigated the application of deep learning techniques for forecasting noisy and high-resolution power grid frequency time series, highlighting the effectiveness of the Temporal Fusion Transformer (TFT). In the context of power systems, RESs may be correlated with other types of renewable energies or with RESs at different locations [12].

The primary contribution of this paper is the development of a Multivariate Long Short-Term Memory (M-LSTM)-based network for forecasting RESs generation. The proposed model leverages real-world data of RESs from various locations in Alberta province. Notably, the model predicts RES generation for a specific area by utilizing its historical data as well as data from neighboring areas. We also consider the temporal gap between the day-ahead market closure and the start of the real-time market in the forecasting of RES. This contribution is pivotal for advancing RESs utilization and supporting climate change mitigation efforts.

2 PROPOSED MODEL

We propose a multivariate LSTM network designed to capture latent features and long-term dependencies among various RESs across different regions. By leveraging historical data, the model aims to enhance forecasting accuracy. The following sections provide a detailed explanation of the proposed approach.

2.1 STRUCTURE OF NETWORK

The proposed model employs a two-layer LSTM network. The first layer consists of 64 LSTM units with a Rectified Linear Unit (ReLU) activation function, while the second layer includes 32 LSTM units, also with ReLU activation. The outputs from the second LSTM layer are passed through a Dropout layer to prevent overfitting and then fed into a fully connected Neural Network for final predictions.

2.2 TRAINING THE MULTIVARIATE LSTM

Given a dataset of N samples $x_{i,j}$, where $i = 1, \dots, N$ (time steps) and $j = 1, \dots, F$ (features), we define the inputs and outputs as follows: The first input consists of a sequence of data points, $x_{i,j}$, from $i = 1$ to $i = p$ (the number of look-back steps) for all j . The second input is a shifted sequence from $i = 2$ to $i = p + 1$ for all j . Correspondingly, the first output label is $x_{p+m,J}$, and the second output label is $x_{p+m+1,J}$, where J represents the feature to predict. These outputs represent predictions, with m steps ahead of the inputs. This configuration is used to define multiple samples and their corresponding labels throughout the dataset. Fig 1a illustrates the data samples and labels in the training process.

To ensure effective training, the input data was normalized, addressing potential issues caused by unscaled features. Normalization mitigates the sensitivity of activation functions to input scale and prevents large-scale features from dominating the learning process at the expense of smaller-scale features. The normalized data was fed into the network, which used the Adam optimizer for weight updates and Mean Squared Error (MSE) as the loss function. After training, the network’s outputs were rescaled to their original scale. Also, we incorporated prior knowledge regarding periods when solar PV systems generate no power. Specifically, for predefined dark hours (i.e., nighttime), we

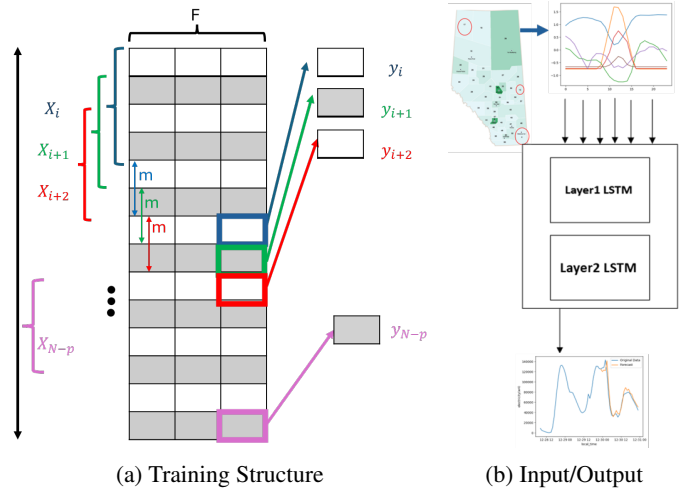


Figure 1: The Proposed M-LSTM

explicitly set the PV generation to zero. This approach eliminates the need for predictions during these hours, ensuring that the model’s outputs remain both accurate and consistent with real-world conditions.

2.3 EXPERIMENTS

We used real renewable energy generation historical data from the Alberta province¹, which includes a full year of data, at one hour time intervals. We considered one type of RES - PV - across three different planning areas in the Alberta province. Our historical dataset consists of 8760 data samples per feature, with three features in total. The model is designed to look back at the past 24 hours of generation data for PV generation across all regions and predict the generation of the desired renewable source for the target time step. Fig 1b illustrates the input and output of our proposed model.

3 RESULTS

To quantify the impact of improved RES forecasting on reducing greenhouse gas emissions, we consider an Economic Dispatch (ED) problem. Three types of synchronous generators and one PV source are available to meet demand. Table2 and Fig3 provide a detailed overview of this system. The forecasted RES generation is utilized in the day-ahead (DA) ED. However, in the real-time (RT) ED, the actual RES generation values become available and are used to adjust the system operation accordingly. The day ahead and real-time ED formulations are provided in the 5.1 and 5.2 respectively.

Based on the DA operation and the RT actual RES generation, the real-time ED must be run to adjust the output of the generators. The discrepancy between the forecasted and actual RES values must be supplied by conventional generators in real-time. These generators are highly flexible, typically powered by natural gas, and contribute to CO₂ emissions. Based on data from [13] and [14], CO₂ emissions for this type of generator fall within a specific range. In this paper, we assumed 202 kg CO₂ per MWh for these generators. If the available generator capacity is insufficient to cover this discrepancy, load shedding may occur, which is costly. Conversely, if the actual RES generation exceeds the forecasted values, it can either help reduce the need for conventional generators and load shedding from the day-ahead operation, or be spilled if it cannot be fully utilized.

We compare three models for forecasting the PV generation in DA market: (1) a K-means approach [4], (2) a monthly average at each hour, and (3) our M-LSTM model. The results show the deviation

¹Available at: <https://www.aeso.ca/market/market-and-system-reporting/data-requests/historical-generation-data>

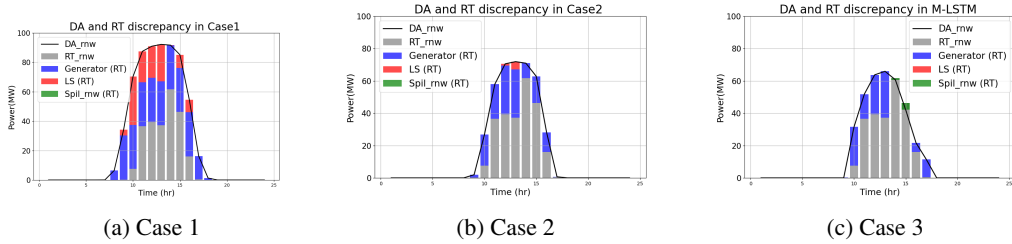


Figure 2: DA and RT discrepancy in three cases

between the DA market and the RT market based on the actual values of RES. Table 1 and Fig. 2 illustrate the results of ED in the DA and RT markets with three different forecast of PV.

	Case1	Case2	M-LSTM
Gas fired unit (MW)	263.7	140.56	108.7
CO ₂ emission (kg)	53,267	28,393	21,957
Load Shedding (MW)	121.7	5.79	0
RES Spillage (MW)	0	0	5
DA+RT Cost (\$)	53,043	44,148	43,490
NMAE	1.57	0.59	0.46

Table 1: Model Comparison and impact on power grid metrics

The dramatic reduction in gas-fired unit usage from 263.7 MW in Case 1 to 108.7 MW with M-LSTM highlights the potential for significant reduction in fossil fuel use in power systems through advanced AI-driven management techniques. This also substantially decreases CO₂ emissions across the different cases, with the M-LSTM approach achieving the lowest emissions at 21,957 kg, compared to 53,267 kg in Case 1 - a remarkable 58% decrease. The M-LSTM method shows superior performance in load management, completely eliminating load shedding (0 MW) compared to 121.7 MW in Case 1 and 5.79 MW in Case 2, indicating its potential for enhancing grid stability and reliability. M-LSTM’s slightly higher spillage reflects its superior forecast accuracy. Other models avoid spillage by overestimating renewable generation, which masks forecast errors. For M-LSTM, spillage only occurs when actual real-time generation surpasses accurate forecasts after meeting all demand and minimizing conventional generation. This minimal spillage demonstrates optimal system balancing, maximizing renewable usage while maintaining grid stability and reducing reliance on conventional power. The approach effectively balances the trade-off between renewable integration and system reliability. Based on the Normalized Mean Absolute Error (NMAE) using Mean Normalization, M-LSTM exhibits the lowest error, indicating the highest prediction accuracy compared to the ground truth. These results underscore the potential of machine learning approaches, particularly M-LSTM, in optimizing power system operations to simultaneously reduce carbon emissions, minimize load shedding, and efficiently integrate renewable energy resources. Additionally, the cost of ED in Case 3 is lower than in the other two cases, suggesting that more accurate forecasting leads to better system performance in terms of both economics and CO₂ emissions.

4 CONCLUSION

The proposed M-LSTM model enhances renewable energy forecasting by capturing complex temporal and spatial dependencies across multiple locations and RES types using historical data. Accurate forecasts lead to more efficient economic dispatch, reduced operational costs, enhanced system reliability, and reducing CO₂ emissions by 58%. This research highlights the critical role of advanced machine learning techniques in improving renewable energy utilization in modern power systems, to mitigate climate change. Results demonstrate that the scaled use of the M-LSTM model across power generation systems could significantly reduce CO₂ emissions.

REFERENCES

- [1] Qusay Hassan, Patrik Viktor, Tariq J Al-Musawi, Bashar Mahmood Ali, Sameer Algburi, Haitham M Alzoubi, Ali Khudhair Al-Jiboory, Aws Zuhair Sameen, Hayder M Salman, and Marek Jaszczur. The renewable energy role in the global energy transformations. *Renewable Energy Focus*, 48:100545, 2024.
- [2] Farshid Kamrani, Sajjad Fattaheian-Dehkordi, Ali Abbaspour, Mahmud Fotuhi-Firuzabad, and Matti Lehtonen. Investigating the impacts of microgrids and gas grid interconnection on power grid flexibility. In *2019 smart grid conference (SGC)*, pages 1–6. IEEE, 2019.
- [3] Farshid Kamrani, Sajjad Fattaheian-Dehkordi, Ali Abbaspour, Mahmud Fotuhi-Firuzabad, and Matti Lehtonen. Flexibility-based operational management of a microgrid considering interaction with gas grid. *IET Generation, Transmission & Distribution*, 15(19):2673–2683, 2021.
- [4] Farshid Kamrani, Sajjad Fattaheian-Dehkordi, Mohammad Gholami, Ali Abbaspour, Mahmud Fotuhi-Firuzabad, and Matti Lehtonen. A two-stage flexibility-oriented stochastic energy management strategy for multi-microgrids considering interaction with gas grid. *IEEE Transactions on Engineering Management*, 70(10):3330–3343, 2021.
- [5] Bo Yang, Xiao Fu, Nicholas D Sidiropoulos, and Mingyi Hong. Towards k-means-friendly spaces: Simultaneous deep learning and clustering. In *international conference on machine learning*, pages 3861–3870. PMLR, 2017.
- [6] Muhammad Waqas and Usa Wannasingha Humphries. A critical review of rnn and lstm variants in hydrological time series predictions. *MethodsX*, page 102946, 2024.
- [7] Steven Elsworth and Stefan Güttel. Time series forecasting using lstm networks: A symbolic approach. *arXiv preprint arXiv:2003.05672*, 2020.
- [8] Hossein Abbasimehr and Reza Paki. Improving time series forecasting using lstm and attention models. *Journal of Ambient Intelligence and Humanized Computing*, 13(1):673–691, 2022.
- [9] Stephen P Holland, Matthew J Kotchen, Erin T Mansur, and Andrew J Yates. Why marginal co2 emissions are not decreasing for us electricity: estimates and implications for climate policy. *Proceedings of the National Academy of Sciences*, 119(8):e2116632119, 2022.
- [10] Nina Effenberger and Nicole Ludwig. Towards turbine-location-aware multi-decadal wind power predictions with cmip6. In *NeurIPS 2024 Workshop on Tackling Climate Change with Machine Learning*, 2024.
- [11] Hadeer El Ashhab, Benjamin Schäfer, and Sebastian Pütz. Feasibility of forecasting highly resolved power grid frequency utilizing temporal fusion transformers. In *NeurIPS 2024 Workshop on Tackling Climate Change with Machine Learning*, 2024.
- [12] Julio Alberto Silva Dias. Long-term hourly scenario generation for correlated wind and solar power combining variational autoencoders with radial basis function kernels. *arXiv preprint arXiv:2306.16427*, 2023.
- [13] Center for Climate and Energy Solutions. Epa regulation of greenhouse gas emissions from new power plants, 2015. Accessed on [20,01,2025].
- [14] EEA Grants. Conversion guidelines. Technical report, EEA Grants, 2022. Accessed on [20,01,2025].

5 APPENDICES

5.1 DAY AHEAD ECONOMIC DISPATCH

$$Obj^{da} = Min\left(\sum_{v,t} C_v \cdot p_{v,t} + \sum_t VOLL \cdot ls_t\right) \quad (1)$$

$$\sum_v p_{v,t} + p_t^{rnw} + ls_t = P_t \quad \forall t \quad (2)$$

$$0 \leq p_{v,t} \leq \overline{P}_v \quad \forall v, t \quad (3)$$

$$-R \leq p_{v,t} - p_{v,t-1} \leq R \quad \forall v, t \neq 1 \quad (4)$$

$$-R \leq p_{v,t} - p_{v,T} \leq R \quad \forall v, t = 1 \quad (5)$$

$$0 \leq p_t^{rnw} \leq \overline{P}^{rnw} \quad \forall t \quad (6)$$

$$0 \leq ls_t \leq P_t \quad \forall t \quad (7)$$

The DA objective function 1 aims to minimize both the generation cost and the load shedding cost. The equation 2 represents the supply-demand balance constraint, ensuring that total generation matches the demand at all times. Other equations impose constraints on the power system components, including upper and lower bounds on the power output of each generator, limits on renewable generation and load shedding, as well as restrictions on the ramping rates of generators. These ramping constraints specifically define how quickly generators can increase (ramp up) or decrease (ramp down) their power output. Collectively, these constraints ensure that the model accurately represents the physical and operational limitations of the power system components.

5.2 REAL TIME ECONOMIC DISPATCH

$$Obj^{rt} = Min\left(\sum_{v,t} C_v \cdot p_{v,t}^{rt} + \sum_t VOLL \cdot ls_t^{rt}\right) \quad (8)$$

$$\sum_v p_{v,t}^{rt} + P_t^{rt,rnw} - p_t^{spill} - p_t^{rnw*} + ls_t^{rt} = P_t \quad \forall t \quad (9)$$

$$0 \leq p_{v,t} + p_{v,t}^{rt} \leq \overline{P}_v \quad \forall v \in V_f, t \quad (10)$$

$$-R \leq (p_{v,t}^* - p_{v,t-1}^*) + p_{v,t}^{rt} - p_{v,t-1}^{rt} \leq R \quad \forall v \in V_f, t \neq 1 \quad (11)$$

$$-R \leq (p_{v,t}^* - p_{v,t-T}^*) + p_{v^{rt},t} - p_{v^{rt},T} \leq R \quad \forall v \in V_f, t = 1 \quad (12)$$

$$0 \leq p_t^{spill} \leq P_t^{rnw,rt} \quad \forall t \quad (13)$$

$$0 \leq ls_t + ls_t^{rt} \leq P_t \quad \forall t \quad (14)$$

These equations are similar to those used in day-ahead ED, but they aim to minimize costs in real-time. Variables marked with an asterisk (*) represent the optimal values determined in the day-ahead market.

5.3 CHARACTERISTIC OF POWER SYSTEM COMPONENTS

Generator	Generation Cost (\$/MW)	Max Generation (MW)	Min Generation (MW)	Availability at DA/RT	Ramp up/down (MW/h)
G1	20	50	0	DA	20
G2	25	50	0	DA	20
G3	30	30	0	DA/RT	30

Table 2: Characteristic of Generators

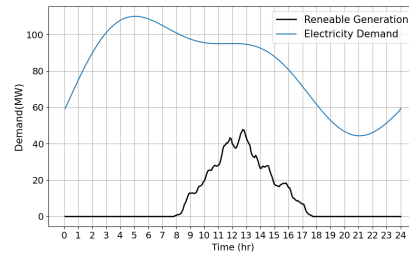


Figure 3: Demand to be supplied, and Actual PV Generation