2023 IEEE Conference on Artificial Intelligence (CAI) | 979-8-3503-3984-0/23/$31.00 ©2023 IEEE | DOI: 10.1109/CAI54212.2023.00141

Blended Temperature Forecasting Model for Thailand Using Multiple Data Sources

Sukrit Jaidee
Energy solution
Electricity Generating Authority of Thailand
Nonthaburi, Thailand
sukrit.jai@egat.co.th

Walanchaporn Boon-nontae
Energy solution
Electricity Generating Authority of Thailand
Nonthaburi, Thailand
Walanchaporn.b@egat.co.th

Weerayut Srithiam
Energy solution
Electricity Generating Authority of Thailand
Nonthaburi, Thailand
weerayut.s@egat.co.th

***Abstract*— With the escalation in the cost of electricity, there has been a noticeable inclination towards the installation of solar photovoltaic (PV) systems in multiple regions across Thailand. The increase in PV installations has led to an electricity demand that fluctuates depending on the prevailing weather conditions, creating challenges in managing and regulating electricity demand. In order to support electricity regulators in managing the fluctuations, it is crucial to implement a solar power forecasting system for individual households. One of the critical variables in forecasting solar power generation, besides solar irradiance, is temperature. This study introduces a temperature prediction system for every geographic location in Thailand at a 10x magnification level, which provided an hourly temperature for each location in the country. The proposed model integrated input data from three open-source platforms, namely Meteostat, Weatherapi, and IBM Weather. Utilizing the capabilities of each input source, the deep learning model was employed. The system, powered by the proposed model, achieved a Mean Squared Error (MSE) of 1.17 °C when compared to the actual data acquired from the Meteorological Department of Thailand.**

***Keywords*— Meteostat, Weatherapi, IBM weather, Deep learning, solar forecasting**

I. Introduction

The growing adoption of solar photovoltaic (PV) systems in residential areas across Thailand has presented new challenges and opportunities for the electricity sector. Given that the efficiency of electricity generation for solar PV systems is highly influenced by weather conditions, Electricity Generating Authority of Thailand (EGAT) requires accurate weather forecasting to manage demand response effectively. This study aimed to develop a deep learning-based temperature prediction system that leveraged data from various open-source platforms such as Meteostat, Weatherapi, and IBM Weather Company. By combining the strengths and addressing the drawbacks of these platforms, the proposed model intended to enhance the accuracy of temperature predictions, enabling better solar forecasting and demand response management.

II. Literature on Temperature Forecasting

This literature review covers current temperature and weather forecasting advancements, focusing on deep learning methods. [1-3] explored hybrid CNN-LSTM models, demonstrating their superior accuracy compared to traditional methods such as ARIMA and STL. The paper [3] also introduces an attention mechanism to further enhance the performance. Paper [4] evaluated various algorithms for temperature forecasting, including Decision Trees, K-Nearest Neighbors, and Support Vector Machines. Paper [5] presented a cascaded LSTM network, improving weather data prediction accuracy. The Papers [6] and [7] investigated deep learning methods, specifically CNNs, LSTMs, and ANNs, for solar power prediction, highlighting their potential benefits for energy management and grid integration. Lastly, papers [8] and [9] compared LSTM-based models, including a spatial feature attention model, with other forecasting methods, demonstrating their effectiveness in various applications, such as energy management and climate studies.

III. Methodology

Current prediction models for time series data include LSTM, GRU, DeepESN, CNN, LSTM-CNN, GRU-CNN, transformer, and even traditional models such as DNN, Extreme Gradient Boosting (XGBoost), and random forest regression. Although the traditional models provide lower accuracy than deep learning models, they take less time to process and use fewer processing resources. Regarding the time constraints of processing 1,100 temperature data points throughout Thailand within an hour, we decided to blend different models to predict temperature. The proposed model combined the following models: Light Gradient Boosting Machine (Light GBM), XGBoost, and Huber. Blending allowed us to extract the advantages and mitigate the disadvantages of each model.

A. Dataset

We explored input data from various open-source sources. The Meteostat provided actual weather station measurements and predicted data (point data). The actual weather station measurements and the predicted data consisted of latitude, longitude, elevation, distance, temperature, dewpoint temperature, relative humidity, wind direction, wind speed, year, month, day, hour, and minute. Additionally, the IBM Weather provided we with precipitation (1 hour), precipitation (6 hours), precipitation (24 hours), pressure change, mean sea-level pressure, relative humidity, temperature, temperature change (24 hours), maximum temperature (24 hours), minimum temperature (24 hours), dewpoint temperature, feels-like temperature, UV index, visibility, wind direction, wind speed, and wind gust. Lastly, the WeatherAPI provided we with temperature, wind, wind (degrees), wind direction, pressure, precipitation, humidity, clouds, feels-like temperature, wind chill, heat index, dewpoint, will it rain, the chance of rain, visibility, and gust.

In this study, we compared the performances of eight fast-processing small models, namely Decision Tree, AdaBoost, Light GBM, Gradient Boosting Regressor (GBR), Least Angle Regression (Lasso Lars), Elastic Net (EN), Stacker, and Blend. The Stacker method operates by taking the predictions of multiple base models and using these predictions as input features for a new model (meta-model). The predictions of this meta-model are then used as the final predictions. This process leverages the strengths of each base model and can learn patterns from the way these models make errors. The Blend method, on the other hand, combines the predictions of

979-8-3503-3984-0/23/$31.00 ©2023 IEEE
DOI 10.1109/CAI54212.2023.00141

multiple models, typically by taking an average, to make the final prediction. Subsequently, we used the temperatures predicted by the most accurate model to calculate the mean squared error (MSE). We then compared this with the MSE of temperatures from all input sources, using the actual data obtained from the Meteorological Department of Thailand as the benchmark. The results confirmed our hypothesis that by leveraging the benefits of each input source and integrating them, we could achieve improved results.

IV. RESULTS

The study revealed that Blend Model yielded the lowest MSE of 1.17°C compared to other models, as shown in Figure 1. When compared the MSE of temperatures at each point in Thailand from each input source, as illustrated in Figures 2 and 3, the Blend Model also yielded the least error across all regions. Additionally, the Blend Model exhibited the lowest MSE value when comparing the MSE values of each input source, as shown in Figure 4., where the MSE value of Station Meteostat (actual measurement data from the Meteostat model), Meteostat Points (predicted data from the Meteostat model), Weather API (data from Weather API sources), and IBM (data from IBM data sources) were 2.67, 2.33, 1.96, and 1.22°C, respectively.

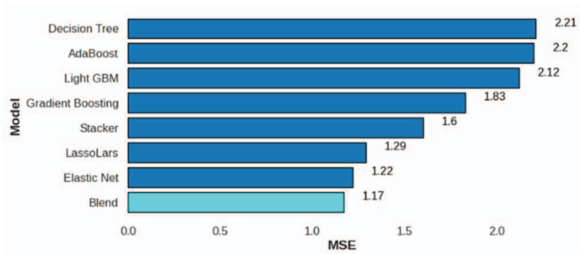

Fig. 1. Compare the MSE values of each model at level 10.

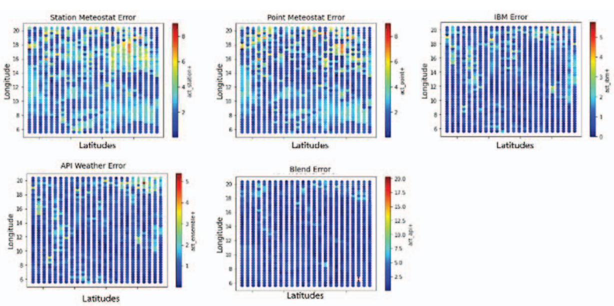

Fig. 2. The MSE value of the temperature for each input source at each point in the country.

V. CONCLUSIONS

In summary, accurate temperature data is crucial for forecasting PV power generation along with solar irradiance. They directly impact the cost-effectiveness of PV installations in different areas. Moreover, the accurate temperature data can be utilized to a geospatial PV power generation forecasting system that meets national demand response management. However, further improvement in the proposed temperature forecasting model is achievable to enhance accuracy and processing speed. When higher spatial resolutions are required, it is essential to optimize prediction accuracy and processing speed to maintain the overall efficiency of the forecasting system.

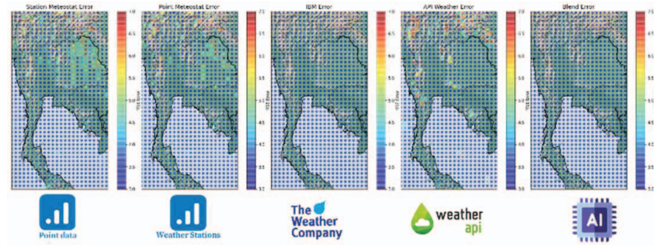

Fig. 3. The MSE value of the temperature for each input source at each point.

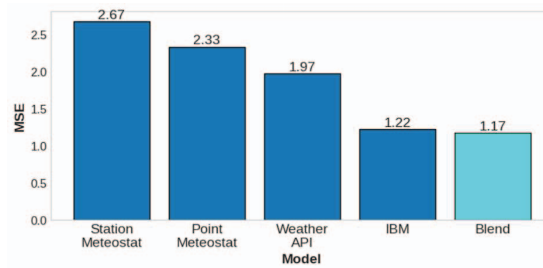

Fig. 4. Compare the MSE values of each input source.

REFERENCES

[1] P. Malini and B. Qureshi, "A Deep Learning Framework for Temperature Forecasting," 2022 7th International Conference on Data Science and Machine Learning Applications (CDMA), Riyadh, Saudi Arabia, 2022, pp. 67-72, doi: 10.1109/CDMA54072.2022.00016.

[2] C. Li, M. Zhao, Y. Liu and F. Xu, "Air Temperature Forecasting using Traditional and Deep Learning Algorithms," 2020 7th International Conference on Information Science and Control Engineering (ICISCE), Changsha, China, 2020, pp. 189-194, doi: 10.1109/ICISCE50968.2020.00049.

[3] X. Zhao, L. Huang and Y. Nie, "Temperature Prediction Based on Integrated Deep Learning and Attention Mechanism," 2021 IEEE 6th International Conference on Computer and Communication Systems (ICCCS), Chengdu, China, 2021, pp. 163-167, doi: 10.1109/ICCCS52626.2021.9449176.

[4] T. Anjali, K. Chandini, K. Anoop and V. L. Lajish, "Temperature Prediction using Machine Learning Approaches," 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), Kannur, India, 2019, pp. 1264-1268, doi: 10.1109/ICICICT46008.2019.8993316.

[5] Z. Al Sadeque and F. M. Bui, "A Deep Learning Approach to Predict Weather Data Using Cascaded LSTM Network," 2020 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), London, ON, Canada, 2020, pp. 1-5, doi: 10.1109/CCECE47787.2020.9255716.

[6] T. M. Al-Jaafreh and A. Al-Odienat, "The Application of Deep Learning Techniques for Solar Power Forecasting," 2022 13th International Conference on Information and Communication Systems (ICICS), Irbid, Jordan, 2022, pp. 214-219, doi: 10.1109/ICICS55353.2022.9811182.

[7] A. El Kounni, H. Radoine, H. Mastouri, H. Bahi and A. Outzourhit, "Solar Power Output Forecasting Using Artificial Neural Network," 2021 9th International Renewable and Sustainable Energy Conference (IRSEC), Morocco, 2021, pp. 1-7, doi: 10.1109/IRSEC53969.2021.9741130.

[8] M. A. R. Suleman and S. Shridevi, "Short-Term Weather Forecasting Using Spatial Feature Attention Based LSTM Model," in IEEE Access, vol. 10, pp. 82456-82468, 2022, doi: 10.1109/ACCESS.2022.3196381.

[9] M. D. Haris, D. Adytia and A. W. Ramadhan, "Air Temperature Forecasting with Long Short-Term Memory and Prophet: A Case Study of Jakarta, Indonesia," 2022 International Conference on Data Science and Its Applications (ICoDSA), Bandung, Indonesia, 2022, pp. 251-256, doi: 10.1109/ICoDSA55874.2022.98628