A Hybrid Model Of Decomposition, Extended Kalman Filter And Autoregressive - Long Short - Term Memory Network For Hourly Day Ahead Wind Speed Forecasting

Nguyen Thi Hoai Thu[1*], Phan Quoc Bao[2], and Pham Nang Van[1]

[1] *Power Grid and Renewable Energy Lab., School of Electrical and Electronic Engineering, Hanoi University of Science and Technology, Vietnam*
[2] *School of Informatics, Computing and Cyber security, Northern Arizona University, USA*
[*] *Corresponding author. E-mail: thu.nguyenthihoai@hust.edu.vn*

Received: June 30, 2023; Accepted: November 17, 2023

This research work introduced an innovative forecasting approach that combined the Autoregressive – Long Short-Term Memory (AR-LSTM) neural network with decomposition techniques and the Extended Kalman Filter (EKF) to predict hourly day-ahead wind speed. The process started with data pre-processing, followed by decomposition into three distinct components: trend, seasonal, and residual, using the Seasonal and Trend decomposition using Loess (STL) filter. The forecasting process was designed to handle each of these decomposed components independently. The trend and seasonal components were forecasted using the AR model, utilizing historical patterns and temporal dependencies. On the other hand, the residuals were predicted by a Long Short-Term Memory network, optimized through the application of the Extended Kalman Filter to improve the filtering process. Predictions from these individual components were then combined to generate the final wind speed forecast. To validate the proposed method, it was applied to real-world wind speed datasets from both Hanoi and Tokyo. The model's performance was systematically compared with alternative methods. The results consistently demonstrated the superiority of the proposed approach over the three alternative methods, as evidenced by the Mean Absolute Error (MAE), Mean Squared Error (MSE), and Mean Absolute Percentage Error (MAPE) metrics. Impressively low values of MAE and MSE, along with an impressive MAPE value, were achieved, namely 6.79% and 10.3%, for hourly day-ahead wind speed prediction in Hanoi and Tokyo, respectively. These findings underscore the robustness and effectiveness of the proposed model in delivering highly accurate wind speed predictions for both geographical locations.

Keywords: Wind speed forecast; Hybrid model; Decomposition; Autoregressive – Long short-term memory; Extended Kalman Filter

http://dx.doi.org/10.6180/jase.202409_27(9).0004

1. Introduction

Due to the rapid economic development and the rise of the living standard, the energy demand is more and more increasing. However, the major issues of the traditional fossil fuel-based energy that the world is facing are the depletion of fossil fuels and their environmental adverse impacts. These lead to the wide use of renewable energy [1, 2]. Among which, wind energy has received extensive global attention and rapidly developed in recent years thanks to its advantages of abundant availability, competitive cost, low carbon footprints [3]. However, with the increasing wind power penetration, its drawbacks of intermittent, randomness, and volatility will make challenging

to the safety and stability of the power system [4, 5]. The high stochasticity and low predictability of wind nature can cause serious issues such as voltage and frequency fluctuations, harmonics, etc. [6]. Forecasting the future wind power accurately is one of promising solutions for these challenges [7]. Especially day ahead predictions play an important role for system operations. Wind power can be predicted either directly using the historical wind power data and meteorological data or undirectedly by forecasting wind speed (WS) first, and then using a wind power curve to produce wind power forecasts [8, 9]. The latter option is necessary when the historical data of wind power is incomplete or unavailable.

In literature, many WS forecasting approaches have been proposed and implemented. Based on the modelling theory, forecasting models can be divided into four types: physical models, traditional statistical models, artificial intelligence (AI)-based models, and hybrid models [5]. Physical models, for example, numerical weather prediction (NWP) usually take into account the physical considerations of various meteorological factors to forecast WS by using the mathematical model of the atmosphere [10]. This approach was shown to be suitable for medium-term and long-term WS forecasting [11]. Statistical models, which include autoregressive (AR) model [12, 13], autoregressive moving average (ARMA) [12], autoregressive integrated moving average (ARIMA) [13], usually characterize the linear fluctuation of WS. These models are effective in very short-term and short-term WS forecasting. Thanks to the advances in computer science, many AI-based models have also been applied in predicting WS, such as MLP Neural Network [14], RBF Neural network [15], artificial neural network (ANN) [16]. The application of 1 single model basically did not show good performance in prediction. Then a combination of several methods was considered for higher performance. In [17], an ensembled model of four different types of ANNs were made with certain combination weights to forecast WS. Shi. et al. proposed two hybrid models (ARIMA-ANN and ARIMA-SVM) to simultaneously extract linear and nonlinear fluctuations in WS, and their performance was shown to be better than that of single forecasting models [18].

Moreover, decomposition techniques were proven to improve the forecasting accuracy [[5, 19, 20]. Wang et al. [19] proposed a hybrid model of SSA (Singular Spectrum Analysis) and a Laguerre neural network for wind speed prediction. The wind speed series was decomposed using SSA into low and high-frequency series, and these series are forecasted by a Laguerre neural network. In [20], Zhang et al. developed forecasting model based on a wavelet method. The results showed that the model's performance had improved meaningfully.

Although numerous studies have been conducted in WS forecasting methods, this issue is still challenging because of the highly intermittent nature of wind speed. In this paper, a hybrid model of statistical and AI-based method combined with data pre-processing and decomposition technique for day-ahead WS forecasting was proposed to improve the forecast accuracy. First, the data was pre-processed by the Spectral Residual (SR) method and decomposed into trend, seasonal and residual series using Seasonal and Trend decomposition using Loess filter (STL) method. Then, the trend and seasonal components were predicted using AR models while the residual series was filtered by Extended Kalman Filter (EKF) before predicted using Long Short-Term Memory (LSTM) neural network. Finally, the obtained results were combined to receive the forecasting of wind power. The proposed model was applied to 2 real WS series collected in Hanoi (Vietnam) and Tokyo (Japan).

2. Proposed forecasting model

As mentioned above, using one single forecasting model was not enough to accurately predict the time series, especially data with many fluctuations like WS. Various hybrid models have been developed and have also shown their strengths in forecasting. In this paper, we proposed a novel hybrid model of AR and LSTM neural network. Moreover, because the time series of WS always have outliers/anomaly points, data need to be preprocessed before being decomposed. The data was split into linear and nonlinear components. The linear components included trend and seasonal series that would be predicted by a simple but highly accurate model of AR. Meanwhile, the nonlinear component was the residuals which were extremely fluctuated and would be filtered using EKF to improve the results. The filtered series was then forecasted by using LSTM network. The proposed model (STL-EKF-AR-LSTM) is shown in Fig. 1. The details of the models were presented in the following subsection.

2.1. Data Pre-processing

The raw data of WS needs to be processed before predicting. This process includes detecting outliers/anomaly/missing values. WS is a time series with high uncertainties, which leads to the presence of outliers/anomaly in the historical data set. The main reasons for these values are due to storage and communication errors [21, 22]. Therefore, it is crucial to detect and remove the outliers/anomaly. In this study, the Spectral Residual (SR) method, a simple yet

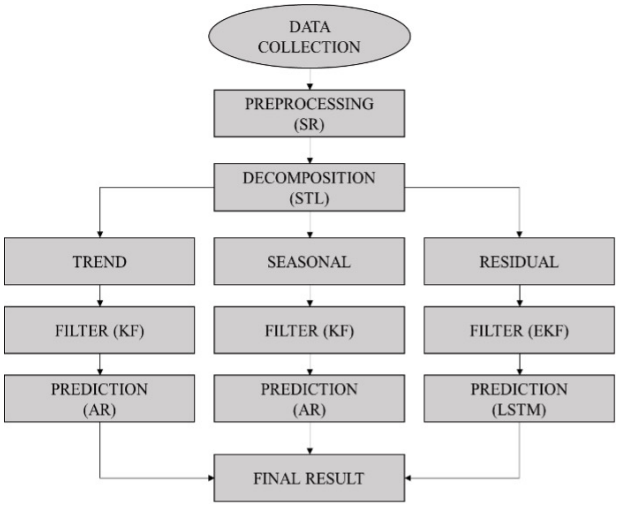

Fig. 1. Structure of the proposed forecasting model (STL-EKF-AR-LSTM).

powerful approach based on Fast Fourier Transform (FFT) was applied [22]. The SR approach is unsupervised and has been proved to be efficient and effective in visual saliency detection applications. The algorithm consists of 3 steps [23]:

- Fourier Transform: Temporal signals are transformed into the frequency domain to analyze frequency structures.
- Log Amplitude Diagram: Computation of a log amplitude diagram using the logarithmic values of frequency component amplitudes, focusing on smaller values and reducing the influence of larger ones.
- Residual Spectrum Identification: Comparison of the log amplitude diagram with a threshold to identify residual spectra, indicating significant differences between actual and predicted signals. The threshold selection is crucial and can be determined using statistical measures such as standard deviation or percentage of residuals.

2.2. Decomposition

Time series data is usually decomposed into 4 following components: Trend, Seasonality, Cyclic, and Irregular remainder. There are a lot of decomposition methods such as classical decomposition, X11 decomposition, Seasonal and Trend decomposition using Loess filter (STL) decomposition [24, 25], etc. Among the mentioned methodologies, classical decomposition and X11 decomposition exhibit a relatively simplistic approach, better suited for datasets showcasing evident and uncomplicated patterns. In contrast, the Seasonal and Trend decomposition using Loess filter (STL) method is renowned for its versatility in capturing intricate and non-linear trends, as well as seasonality commonly encountered in real-world time series data. Hence, it was deemed fitting for our study, where an effective methodology was imperative to adeptly model the intricate patterns inherent in wind speed data. As a result, STL was selected due to its robustness and efficiency. In this study, we employed the STL method to decompose WS data into three essential components: Trend, Seasonality, and Residuals. These components constituted the foundational input for the subsequent forecasting steps.

2.3. Extended Kalman Filter

Kalman Filter is a Linear-Gaussian State Space Model, which is named after Rudolf E. Kalman who introduced the algorithm through its famous research in 1960. Kalman Filter is expanded from the development of Gauss' Original Development of Least Squares to estimate unknown parameters of a model. To respond to nonlinear issues, the new Kalman Filter has been developed with the name Extended Kalman Filter (EKF). In order to use the EKF to estimate the inside status of a process only gives a sequence of noise observations, the state transition and observation model do not need to be linear functions of state [26]:

$$x_k = f\left(x_{k-1}, u_k\right) + w_k \tag{1}$$

$$z_k = h\left(x_k\right) + v_k \tag{2}$$

Where, u_k is the control vector, w_k and v_k are the noises of process and observation which are assumed to be zero-mean multivariate Gaussian noises with covariance Q_k and R_k.

The algorithm of Kalman Filter including 2 stages can be illustrated as following:

Time Update: x_n^-, priori state estimate and P_n^-, priori estimate error covariance at step n based on the previous time step values are:

$$x_n^- = Fx_{n-1} \tag{3}$$

$$P_n^- = FP_{n-1}F^T + GQG^T \tag{4}$$

Measurement Update: when the new observation value y_n is known, the estimate of x_n at time n will be calculated by:

$$x_n = x_n^- + K_n\left(y_n - Fx_n^-\right) \tag{5}$$

The Kalman Filter gain K_n and the estimate covariance P_n at n :

$$K_n x_n = P_n^- H^T\left(HP_n^- H^T + R_0\right)^{-1} \tag{6}$$

$$P_n = \left(I - K_n H\right) P_n^- \tag{7}$$

Where F and H are state transition matrix and measurement matrix respectively, I is identity matrix.

2.4. Autoregressive model

Autoregressive (AR) model is a model that estimate the future value of a time series by using a linear combination of past values of the variable. The term autoregression indicates that it is a regression of the variable against itself. An autoregressive model of order p, AR(p), can be written as

$$Y_t = \mu + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + \varepsilon_t \quad (8)$$

where Y_t is the value of Y at t, $\phi_1, \ldots, \phi_p$ denote the parameters, ε_t is the variance of the error term.

The order p was selected based on the autocorrelation (ACF) and partial autocorrelation (PACF) of the WS series with its lags, meanwhile the other parameters were estimated using the maximum likelihood [2]. The equation of ACF and PACF was shown in (9), (10), respectively.

$$ACF_k = \frac{\sum_{t=k+1}^{n} (y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^{n} (y_t - \bar{y})^2} \quad (9)$$

$$\text{PACF}_k = \frac{\text{Cov}(y_t, y_{t-k} \| y_{t-1}, y_{t-2}, \ldots, y_{t-k+1})}{\sqrt{\text{Var}(y_t \| y_{t-1}, y_{t-2}, \ldots, y_{t-k+1}) \cdot \text{Var}(y_{t-k} \| y_{t-1}, y_{t-2}, \ldots, y_{t-k+1})}} \quad (10)$$

Where, y_t and y_{t-k} are the values of the time series at time t and $t-k$, respectively.

In this paper, we used the AR(1) model to forecast the linear components such as Trend and Seasonal string.

2.5. Long short-term memory model

LSTM network, which is one sort of the recurrent neural network (RNN), has the advantage in the relatively long-term memory of valuable information [27–30]. This makes it broadly used in time series prediction.

Fig. 2 shows the structure of a LSTM cell. It has 3 gates: input, forget and output gate. The input gate controls whether the new state of the calculation can be updated to the memory unit. The forget gate controls the forgotten information in the previous memory unit. The output gate regulates the extent to which the current output depends on the current memory unit [5, 28].

Forget gate: The forget gate in a LSTM network decides what information from the previous cell state should be discarded or remembered. It uses the current input and the previous hidden state to determine how much of the previous cell state to keep.

$$f_t = \sigma \cdot \left(W_f \cdot [h_{t-1}, x_t] + b_f \right) \quad (11)$$

Input gate: The input gate controls the update of the cell state with new information. It decides which portions of the new input and the previous hidden state should be added to the cell state to update its information.

$$i_t = \sigma \cdot (W_i \cdot [h_{t-1}, x_t] + b_i) \quad (12)$$

Output gate: The output gate selects relevant portions of the cell state to be the actual output of the LSTM cell. It filters and transforms the cell state before producing the output, ensuring that the relevant information is utilized for predictions or downstream tasks.

$$o_t = \sigma \cdot (W_o \cdot [h_{t-1}, x_t] + b_o) \quad (13)$$

$$\tilde{C}_t = \tanh \cdot (W_C \cdot [h_{t-1}, x_t] + b_c) \quad (14)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (15)$$

Where $f_t, i_t, \tilde{C}_t, o_t$ are the output values of the forget gate, input gate, update signal, and output gate, respectively. $W_{i,f,o,c}$ are the weight matrix of each layer. $b_{f,i,o}$ are the bias of each gate and b_y is the output bias. σ is a nonlinear activation function.

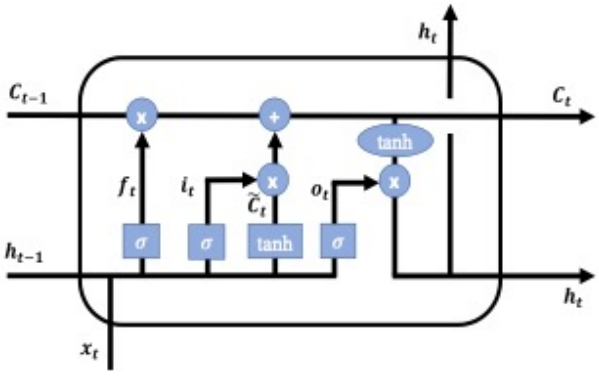

Fig. 2. The structure of a LSTM cell.

2.6. Evaluation metrics

This study employed various evaluation criteria to assess the effectiveness of the forecasting system. Specifically, it focused on point forecasting criteria, which evaluated the accuracy of individual predictions. The utilized evaluation metrics included the mean squared error (MSE), which measured the average difference between projected and observed values. Additionally, the mean absolute error (MAE) was also applied to measure the average difference between projected and actual values. Furthermore, the mean absolute percentage error (MAPE) was computed to represent the average percentage difference between projected and actual values. These metrics served as reference points for evaluating the accuracy and reliability of the

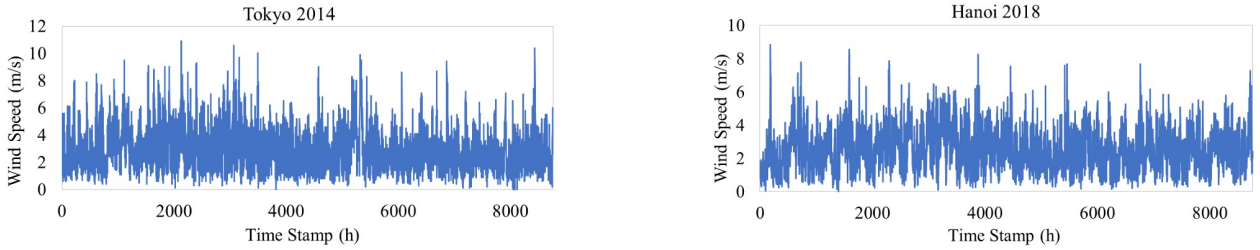

Fig. 3. The wind speed time series in Tokyo (2014) and Hanoi (2018).

forecasting system. The corresponding formulas for these criteria were presented in equations (16), (17), and (18).

$$MSE = \frac{1}{n} * \sum \left(y_{\text{true}} - y_{\text{pred}} \right)^2 \tag{16}$$

$$MAE = \frac{1}{n} * \sum \left| y_{\text{true}} - y_{\text{pred}} \right| \tag{17}$$

$$MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{y_{\text{true}} - y_{\text{pred}}}{y_{\text{true}}} \right| \tag{18}$$

When y_{true} is the true values, y_{pred} is the forecasting values, n is the number of sample data.

3. Result and discussion

3.1. Data samples and preprocessing

The WS data used in the study has been collected from the original website of the Vietnamese and Japanese Meteorological Agency, data sampled by hour for 1 year (2018). Fig. 3 shows the change of wind speed in Tokyo and Hanoi over time. In Tokyo, WS ranges from 0 to 12 m/s, whereas in Hanoi, it typically stays within the 0-9 m/s range.

As mentioned above, the input data needs to be processed, removed unusual points by SR method and decomposed into 3 different components. After using SR, we have been able to eliminate anomalies. We proposed using EKF to filter out the noise components and using the STL model to decompose the new data into 3 parts as shown in Fig. 4. The trend and seasonal components have a simpler form than the remainder series. Moreover, based on the residual plots, it is evident that the oscillations in Tokyo for the year 2014 fluctuated more than that of Hanoi in 2018. Specifically, the residual of WS in Tokyo exhibited a range of oscillations from 0 to 5, whereas that in Hanoi ranged from 0 to 2.5.

3.2. Model parameters estimation

The ACF and PACF of trend and seasonal components of the WS series at two locations were calculated. Figure 5 illustrated the ACF and PACF of the trend component of

Fig. 4. The result of STL decomposition model for wind speed in Tokyo and Hanoi.

the Hanoi WS dataset. From the results, it can be observed that the present value mainly depended on the 1-lag value. Consequently, the order of AR in this case was selected as 1. Similarly, the order of AR models for the seasonal series and for Tokyo's series was determined. The lag settings used in this study pertain to the time intervals between data samples within the time series. This means that when computing the auto-correlation function (ACF), the level of correlation between a current value and past values at time intervals specified by the chosen lag was determined. By identifying these lags, the correlation between data values in the time series could be analyzed, and significant correlation patterns during the analysis and modeling process could be identified.

This study implemented a neural network model using Python, employing the popular LSTM architecture to handle the nonlinearity in the Residual component. The model was structured with multiple layers, each configured with

specific parameters:

- The first layer was an LSTM layer with 128 units and a dropout rate of 0.2: LSTM (128), Dropout = 0.2
- The second layer was another LSTM layer with 64 units and a dropout rate of 0.2: LSTM (64), Dropout = 0.2
- The final layer was a Dense layer with 24 units: Dense = 24.

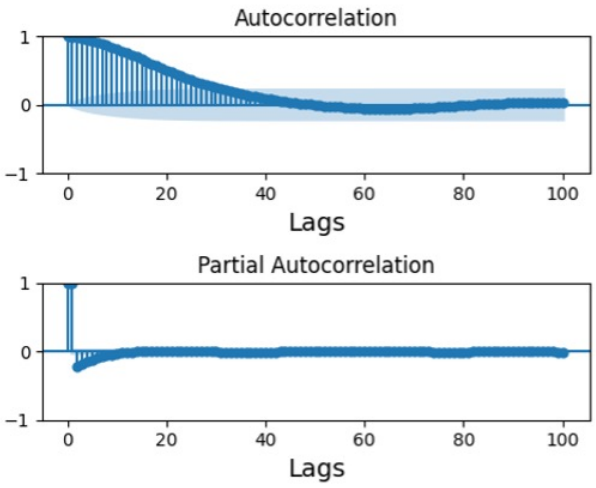

Fig. 5. ACF and PACF of trend component of WS series in Hanoi (2018).

3.3. Forecasting result and comparison with other methods

In order to demonstrate the effectiveness of the proposed model, 3 models of ARMA(1,1) model, LSTM model and the combination of decomposition and AR-LSTM model without EKF were used to predict the WS and compare with the proposed model. Fig. 6 presents the forecasting value of WS using 4 models and the actual value of WS in 2 locations. The proposed model would obtain better performance in forecasting. The difference in errors between the AR-LSTM model and the proposed model is because of the participation of the EKF. EKF filtered the noise from the residual components, making them less fluctuated and stochastic, thus being better trained and resulting in better performance in forecasting.

To evaluate the accuracy of the proposed method, we estimated the criteria of MAPE (%), MSE and MAE of the prediction and compared it with other methods. Table 1 showed the comparison of evaluation metrics when forecasting the WS in Tokyo and Hanoi using 4 mentioned methods. In the context of data from Hanoi, the analysis of comparing MAPE values has revealed that the "ARMA(1,1)" method achieves the highest MAPE value of 24.39%. Immediately following, the "LSTM" model attains a MAPE of 22.88%, followed by the "AR-LSTM" approach with a MAPE value of 14.72%. This analysis sheds light on the capabilities of the AR-LSTM combined model in enhancing performance compared to individual methods. Notably, the proposed "STL-EKF-AR-LSTM" model achieves the lowest MAPE value, at only 6.79%. This significant reduction compared to alternative methods underscores the model's prowess in minimizing prediction errors, further accentuating the importance and rationale for method integration. The two remaining error metrics exhibit similar trends to MAPE.

For the Tokyo dataset, the comparison of MAPE values reveals that the LSTM method obtains the highest MAPE value at 34%, followed by the ARMA(1,1) method at 27.93%, and the AR-LSTM method with a MAPE value of 15.4%. This comparative analysis elucidates the notable efficacy of the combined AR and LSTM methods. Furthermore, the proposed model continues to outperform across all three-error metrics, achieving the lowest MAPE value of 10.3%. The comparison between the two geographical locations highlights the consistent prominence of the proposed model (STL-EKF-AR-LSTM) as the most accurate forecasting method for all error metrics in both Hanoi and Tokyo. The distinct divergence in error values among methods emphasizes the substantial improvements the proposed model brings when compared to alternative approaches. The AR-LSTM combined model stands out markedly in enhancing prediction accuracy compared to the other models. The MAPE analysis demonstrates that AR-LSTM provides better performance than both individual ARMA(1,1) and LSTM models. The amalgamation of LSTM's modeling capability and AR's predictive strength significantly enhances data wind estimation. Moreover, the proposed model (STL-EKF-AR-LSTM) maintains its prominence over the AR-LSTM model. The integration of Extended Kalman Filter (EKF) and AR-LSTM benefits data separation and information update at each time step, simultaneously reducing the influence of noise and state space transformations in the prediction model. This emphasizes that the proposed model not only harnesses the advantages of AR-LSTM but also optimizes the integration of data separation and information filtering, consequently improving accuracy and confidence in wind prediction.

Nevertheless, in comparison between the 2 locations, the total forecasting results of Hanoi's WS were better than Tokyo's WS. This can be explained by their geography.

Table 1. Evaluation metrics of forecasting WS in Hanoi and Tokyo.

	Forecasting method	MAPE (%)	MSE	MAE
Hanoi	ARMA(1,1)	24.39	0.9	0.62
	LSTM	22.88	0.5	0.52
	AR-LSTM	14.72	0.44	0.34
	Proposed model (STL-EKF-AR-LSTM)	6.79	0.05	0.17
Tokyo	ARMA(1,1)	27.93	0.4	0.5
	LSTM	34	0.7	0.7
	AR-LSTM	15.4	0.3	0.4
	Proposed model (STL-EKF-AR-LSTM)	10.3	0.1	0.2

Tokyo is a coastal city, so it will suffer a lot of influence from the sea and pressures, leading to the unstable and strong fluctuation in WS, then the error in the forecasting will be quite large. Unlike Tokyo, Hanoi is on the mainland, so the wind data will be more stable, and the forecasting accuracy is also higher. These values are not only much smaller than those of other methods but also considerably good value when compared with other studies on hourly day ahead WS forecasting in reference.

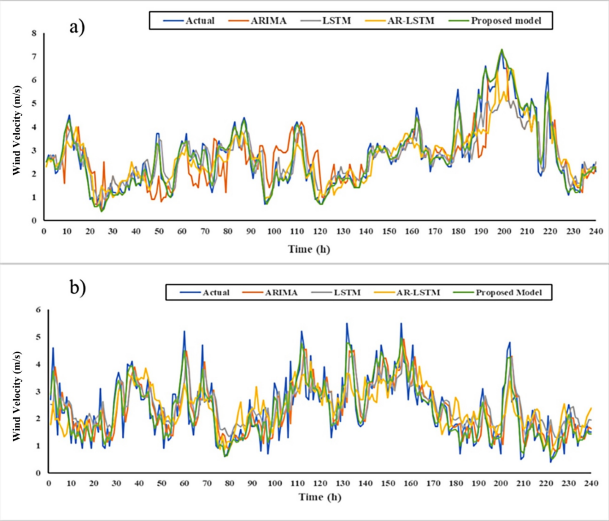

Fig. 6. The forecasting result of (a) Tokyo and (b) Hanoi.

4. Conclusion

This paper proposed a high-accuracy hybrid model of STL-EKF-AR-LSTM for day ahead WS forecasting. The model included 4 stages: data pre-processing, decomposition, forecasting and composition. After the WS data is decomposed by STL, the AR model was used to predict the linear components while LSTM model was used to predict the nonlinear component. Additionally, in order to improve the forecasting performance, EKF was adopted to filter the noise out of residual series before inputting it into the LSTM model. The proposed method was applied to WS in Tokyo and Hanoi. Calculation of the evaluation metrics such as MAPE, MAE and MSE was carried out. The results show that the proposed method obtained the lowest error in forecasting when compared with 3 different methods. Moreover, Hanoi WS has better forecasting performance than Tokyo WS because its WS is more stable than Tokyo's.

References

[1] Y. Wang, R. Zou, F. Liu, L. Zhang, and Q. Liu, (2021) *"A Review of Wind Speed and Wind Power Forecasting with Deep Neural Networks"* **Applied Energy *304***: 117766. DOI: 10.1016/j.apenergy.2021.117766.

[2] N. T. H. Thu, P. N. Van, N. V. N. Nam, and P. H. Minh, (2022) *"Forecasting Wind Speed Using A Hybrid Model Of Convolutional Neural Network And Long-Short Term Memory With Boruta Algorithm-Based Feature Selection"* **Journal of Applied Science and Engineering *26***(8): 1053–1060. DOI: 10.6180/jase.202308_26(8).0001.

[3] A. Gupta, A. Kumar, and K. Boopathi, (2021) *"Intra-day Wind Power Forecasting Employing Feedback Mechanism"* **Electric Power Systems Research *201***: 107518. DOI: 10.1016/j.epsr.2021.107518.

[4] Z. Zhen, G. Qiu, S. Mei, F. Wang, X. Zhang, R. Yin, Y. Li, G. J. Osório, M. Shafie-khah, and J. P. Catalão, (2022) *"An Ultra-Short-Term Wind Speed Forecasting Model Based on Time Scale Recognition and Dynamic Adaptive Modeling"* **International Journal of Electrical Power & Energy Systems *135***: 107502. DOI: 10.1016/j.ijepes.2021.107502.

[5] N. T. H. Thu, P. N. Van, and P. Q. Bao. "Multi-Step Ahead Wind Speed Forecasting Based on a Bi-LSTM Network Combined with Decomposition Technique". In: *Computational Intelligence Methods for Green Technology and Sustainable Development*. Ed. by Y.-P. Huang, W.-J. Wang, H. A. Quoc, H.-G. Le, and H.-N. Quach. Cham: Springer International Publishing, 2023, 569–580.

[6] Y.-K. Wu, J.-J. Zeng, G.-L. Lu, S.-W. Chau, and Y.-C. Chiang, (2020) *"Development of an Equivalent Wind Farm Model for Frequency Regulation"* **IEEE Transactions on Industry Applications** ***56***(3): 2360–2374. DOI: 10.1109/TIA.2020.2974418.

[7] Y. Wang, Q. Hu, L. Li, A. M. Foley, and D. Srinivasan, (2019) *"Approaches to Wind Power Curve Modeling: A Review and Discussion"* **Renewable and Sustainable Energy Reviews** ***116***: 109422. DOI: 10.1016/j.rser.2019.109422.

[8] T. H. T. Nguyen and Q. B. Phan, (2022) *"Hourly Day Ahead Wind Speed Forecasting Based on a Hybrid Model of EEMD, CNN-Bi-LSTM Embedded with GA Optimization"* **Energy Reports** ***8***: 53–60. DOI: 10.1016/j.egyr.2022.05.110.

[9] J. Yan and T. Ouyang, (2019) *"Advanced Wind Power Prediction Based on Data-Driven Error Correction"* **Energy Conversion and Management** ***180***: 302–311. DOI: 10.1016/j.enconman.2018.10.108.

[10] J. Jung and R. P. Broadwater, (2014) *"Current Status and Future Advances for Wind Speed and Power Forecasting"* **Renewable and Sustainable Energy Reviews** ***31***: 762–777. DOI: 10.1016/j.rser.2013.12.054.

[11] N. E. Huang and Z. Wu, (2008) *"A Review on Hilbert-Huang Transform: Method and Its Applications to Geophysical Studies"* **Reviews of Geophysics** ***46***(2): DOI: 10.1029/2007RG000228.

[12] J. Torres, A. García, M. De Blas, and A. De Francisco, (2005) *"Forecast of Hourly Average Wind Speed with ARMA Models in Navarre (Spain)"* **Solar Energy** ***79***(1): 65–77. DOI: 10.1016/j.solener.2004.09.013.

[13] K. Yunus, T. Thiringer, and P. Chen, (2016) *"ARIMA-Based Frequency-Decomposed Modeling of Wind Speed Time Series"* **IEEE Transactions on Power Systems** ***31***(4): 2546–2556. DOI: 10.1109/TPWRS.2015.2468586.

[14] R. C. Deo, M. A. Ghorbani, S. Samadianfard, T. Maraseni, M. Bilgili, and M. Biazar, (2018) *"Multi-Layer Perceptron Hybrid Model Integrated with the Firefly Optimizer Algorithm for Windspeed Prediction of Target Site Using a Limited Set of Neighboring Reference Station Data"* **Renewable Energy** ***116***: 309–323. DOI: 10.1016/j.renene.2017.09.078.

[15] Y. Zhang, C. Zhang, Y. Zhao, and S. Gao, (2018) *"Wind speed prediction with RBF neural network based on PCA and ICA"* **Journal of Electrical Engineering** ***69***(2): 148–155. DOI: 10.2478/jee-2018-0018.

[16] E. Cadenas and W. Rivera, (2009) *"Short Term Wind Speed Forecasting in La Venta, Oaxaca, México, Using Artificial Neural Networks"* **Renewable Energy** ***34***(1): 274–278. DOI: 10.1016/j.renene.2008.03.014.

[17] J. Song, J. Wang, and H. Lu, (2018) *"A Novel Combined Model Based on Advanced Optimization Algorithm for Short-Term Wind Speed Forecasting"* **Applied Energy** ***215***: 643–658. DOI: 10.1016/j.apenergy.2018.02.070.

[18] J. Shi, J. Guo, and S. Zheng, (2012) *"Evaluation of Hybrid Forecasting Approaches for Wind Speed and Power Generation Time Series"* **Renewable and Sustainable Energy Reviews** ***16***(5): 3471–3480. DOI: 10.1016/j.rser.2012.02.044.

[19] C. Wang, H. Zhang, and P. Ma, (2020) *"Wind Power Forecasting Based on Singular Spectrum Analysis and a New Hybrid Laguerre Neural Network"* **Applied Energy** ***259***: 114139. DOI: 10.1016/j.apenergy.2019.114139.

[20] K. Zhang, R. Gençay, and M. Ege Yazgan, (2017) *"Application of Wavelet Decomposition in Time-Series Forecasting"* **Economics Letters** ***158***: 41–46. DOI: 10.1016/j.econlet.2017.06.010.

[21] Y. Zhao, L. Ye, W. Wang, H. Sun, Y. Ju, and Y. Tang, (2018) *"Data-Driven Correction Approach to Refine Power Curve of Wind Farm Under Wind Curtailment"* **IEEE Transactions on Sustainable Energy** ***9***(1): 95–105. DOI: 10.1109/TSTE.2017.2717021.

[22] X. Hou and L. Zhang. "Saliency Detection: A Spectral Residual Approach". In: *2007 IEEE Conference on Computer Vision and Pattern Recognition*. 2007 IEEE Conference on Computer Vision and Pattern Recognition. 2007, 1–8. DOI: 10.1109/CVPR.2007.383267.

[23] H. Ren, B. Xu, Y. Wang, C. Yi, C. Huang, X. Kou, T. Xing, M. Yang, J. Tong, and Q. Zhang. "Time-Series Anomaly Detection Service at Microsoft". In: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. Comment: KDD 2019. 25, 2019, 3009–3017. DOI: 10.1145/3292500.3330680. arXiv: 1906.03821 [cs, stat].

[24] *6.6 STL Decomposition | Forecasting: Principles and Practice (2nd Ed)*. Accessed: Jun. 22, 2022.

[25] O. Trull, J. C. García-Díaz, and A. Peiró-Signes, (2022) *"Multiple Seasonal STL Decomposition with Discrete-Interval Moving Seasonalities"* **Applied Mathematics and Computation** ***433***: 127398. DOI: 10.1016/j.amc.2022.127398.

[26] C.-N. Ko and C.-M. Lee, (2013) *"Short-Term Load Forecasting Using SVR (Support Vector Regression)-Based Radial Basis Function Neural Network with Dual Extended Kalman Filter"* **Energy *49***: 413–422. DOI: 10.1016/j.energy.2012.11.015.

[27] N. T. Hoai Thu, P. N. Van, P. Q. Bao, N. V. Nhat Nam, P. H. Minh, and T. N. Quang. "Short-Term Forecasting of Solar Radiation Using a Hybrid Model of CNN-LSTM Integrated with EEMD". In: *2022 6th International Conference on Green Technology and Sustainable Development (GTSD)*. 2022 6th International Conference on Green Technology and Sustainable Development (GTSD). 2022, 854–859. DOI: 10.1109/GTSD54989.2022.9988761.

[28] J. Qu, Z. Qian, and Y. Pei, (2021) *"Day-Ahead Hourly Photovoltaic Power Forecasting Using Attention-Based CNN-LSTM Neural Network Embedded with Multiple Relevant and Target Variables Prediction Pattern"* **Energy *232***: 120996. DOI: 10.1016/j.energy.2021.120996.

[29] N. N. V. Nhat, D. N. Huu, and T. N. T. Hoai, (2023) *"Evaluating the EEMD-LSTM Model for Short-Term Forecasting of Industrial Power Load: A Case Study in Vietnam"* **International Journal of Renewable Energy Development *12***(5): 881–890. DOI: 10.14710/ijred.2023.55078.

[30] T. H. T. Nguyen, Q. B. Phan, V. N. N. Nguyen, and H. M. Pham. "Day-Ahead Electricity Load Forecasting Based on Hybrid Model of EEMD and Bidirectional LSTM". In: *The 5th International Conference on Future Networks & Distributed Systems*. ICFNDS 2021. New York, NY, USA: Association for Computing Machinery, 13, 2022, 31–41. DOI: 10.1145/3508072.3508079.