Multi-Step Wind Speed Forecasting by Secondary Decomposition Algorithm and LSTM

Ari Shawkat Tahir
Computer Science Department, College of Science, University of Zakho, Iraq
ari.tahir@uoz.edu.krd (corresponding author)

Adnan Mohsin Abdulazeez
Energy Department, Technical College of Engineering, Duhok Polytechnic University, Iraq
adnan.mohsin@dpu.edu.krd

Ismael Ali Ali
Computer Science Department, College of Science, University of Zakho, Iraq
ismail.ali@uoz.edu.krd

Received: 2 September 2024 | Revised: 15 September 2024 | Accepted: 4 October 2024

 DOI: https://doi.org/10.48084/etasr.8895

ABSTRACT

Enhancing the reliability of wind speed forecasting is vital for efficient wind power generation. Given the wind's stochastic nature, preprocessing is crucial to obtain a clean wind speed series. This study introduces an innovative wind speed prediction model that integrates Variational Mode Decomposition (VMD), Symplectic Geometry Mode Decomposition (SGMD), and Long Short-Term Memory (LSTM). The model begins with VMD dividing the series into low- and high-frequency parts, then the SGMD further analyzes the high-frequency segment, and LSTM predicts results based on these components. Collaborative use of VMD and SGMD enables thorough decomposition of intricate wind speed data, while LSTM boosts the model's ability to capture patterns and dependencies. This hybrid model addresses the challenges posed by wind power uncertainty, aiming to efficiently integrate wind energy into power systems. The proposed hybrid model was compared to some benchmark models and outperformed them, reducing MAPE by 58% and RMSE by 31% for Dataset 1, and improving MAPE by 14% and RMSE by 36% for Dataset 2. The results confirm the competitive strength of the proposed strategy. Furthermore, the suggested two-stage decomposition technique demonstrates suitability for the examination of nonlinear characteristics in wind speed patterns.

Keywords-wind speed prediction; secondary decomposition; VMD; SGMD; deep learning; LSTM

I. INTRODUCTION

The wind holds significant importance in numerous areas of human existence, particularly in its contribution to renewable energy sources. This natural element not only offers sustainable power solutions but also influences various other aspects of daily life and environmental sustainability [1]. Wind also has a crucial impact on industries such as shipping, where it affects navigation and safety, and agriculture, where it influences crop growth and soil conditions. Additionally, the wind plays a role in natural disasters, as its turbulence intensity can lead to severe weather events, affecting both human activities and the environment [2]. Wind is a key factor that significantly affects agricultural production, influencing various aspects, from pollination and soil erosion to the microclimate of crop fields [3]. Strong winds can pose risks in agricultural settings by hindering pollination and facilitating the spread of pests, such as promoting fungal growth in plants. These adverse effects can lead to significant losses for farmers [4].

In recent decades, a variety of forecasting models have been proposed to enhance the precision of wind speed prediction. In general, these methods can be categorized into four main types: physical methods, which are based on atmospheric data and physical principles; statistical methods, using historical data to predict future trends; artificial intelligence methods, applying advanced algorithms and machine learning; and hybrid methods, which combine elements of the aforementioned approaches for improved accuracy [5, 6]. In hybrid forecasting methods, data preprocessing techniques are often combined with machine learning algorithms to enhance the accuracy of wind speed predictions. For example, in [7], empirical wavelet transform was applied to decompose speed data, LSTM and regularized

Extreme Learning Machine (ELM) were used for prediction, and then the results were reconstructed with the inverse empirical wavelet transform. In [8], wavelet transform was combined with a genetic algorithm-optimized Support Vector Machine (SVM), selecting input variables based on autocorrelation and partial correlation. In [9], the effectiveness of empirical mode decomposition and its variations was evaluated, showing that a combination of the complete ensemble empirical mode decomposition with adaptive noise and SVM yielded the best results. In [10], wavelet packet decomposition and filtering were used, and Elman neural networks and boosting algorithms were employed for forecasting.

The concept of a secondary decomposition algorithm has gained significant attention to reduce non-stationarity in data sequences more effectively. Its objective is to harness the strengths of diverse decomposition techniques to enhance the overall efficacy of the decomposition process. In [11], a multi-step wind speed forecasting model was formulated, combining a secondary decomposition algorithm with an optimized Wavelet Neural Network (WNN). The Ensemble Empirical Mode Decomposition (EEMD) was used to extract the appropriate and detailed components from the original data series. For components with the highest frequency, VMD was employed for further decomposition. Empirical evidence demonstrated that the proposed EEMD-VMD-HSBADAWNN model outperformed other comparative models in terms of accuracy and reliability.

In [12], a hybrid forecasting model was developed, which combined Empirical Mode Decomposition (EMD), Wavelet Packet Decomposition (WPD), Crisscross Optimization (CSO), and ELM. Wind speed data was first broken down into various Intrinsic Mode Functions (IMFs) through EMD. The component with the highest frequency, IMF1, was then further decomposed using WPD. The experimental results indicated that the effectiveness of EMD was enhanced through this secondary decomposition process. In [13], a Secondary Decomposition Algorithm using Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) and VMD was proposed for wind speed prediction. In this approach, the highest frequency component derived from the CEEMDAN decomposition results was re-decomposed using VMD. The proposed prediction framework comprised a Kernel ELM (KELM), which was enhanced with an Improved Hybrid Differential Evolution-Harris Hawks Optimization (IHDEHHO) strategy.

Numerous studies have shown the effectiveness of both decomposition models and machine learning algorithms. It can be inferred that secondary decomposition algorithms, which capitalize on the strengths of various decomposition methods, tend to exhibit superior performance. Moreover, it is important to recognize that commonly utilized decomposition techniques exhibit a mix of advantages and shortcomings [14]. For example, EMD is known for its adaptability, especially in handling nonlinear and non-stationary data series. However, EMD is not without its drawbacks, such as end effects, mode-mixing issues, and over-envelope problems [15]. On the other hand, VMD is effective in accurately extracting low- and medium-frequency signals but struggles with identifying weak high-frequency signals [16]. To address these decomposition challenges in VMD, the SGMD algorithm was used in [17]. SGMD leverages symplectic geometry similarity transformations to break down a signal into a series of independent mode components, demonstrating a strong capacity to mitigate issues such as mode-mixing, sensitivity to user-defined parameters, and lack of noise robustness. The SGMD algorithm has found applications in different areas, such as noise reduction [18], medical diagnosis [14], and fault diagnosis [19]. Numerous studies have shown the effectiveness of SGMD in these fields, which collectively have attested to its superior performance in data decomposition tasks.

This study proposes a secondary decomposition algorithm that combines the strengths of both VMD and SGMD. Initially, VMD is employed to isolate the low-frequency components of the signal. Subsequently, SGMD is applied to further decompose the remaining part of the signal into an appropriate number of components. This preparatory method maximizes the utilization of signal information and obtains several independent and simple decomposition results. Moreover, the LSTM [4] algorithm is subsequently employed to develop a predictor that offers enhanced performance. The LSTM is a deep learning algorithm, which is part of a larger family of machine learning strategies that have been effective against several domains in the time series field over the last years [20].

This study introduces a hybrid model for wind speed forecasting, combining secondary decomposition with deep learning techniques. The model employs a two-tier decomposition approach, initially utilizing VMD to separate the original wind speed series into low- and high-frequency components. This step capitalizes on VMD's proficiency in isolating low-frequency elements, thereby minimizing the initial data's information content to prevent SGMD's potential over-decomposition. Subsequently, SGMD is applied to further break down the high-frequency portion into a series of independent Symplectic Geometry Components (SGCs) and residuals. Then, LSTM is deployed to forecast the components derived from the secondary decomposition process. The aggregate forecast result is then computed by adding the predicted values across all modes. The primary contributions of this study can be summarized as follows:

- Introduces a new method specifically designed for segmenting wind speed series into fully reconstructed modes, enhancing the accuracy and clarity of decomposition.
- Combining the strengths of VMD and SGMD simplifies the complexities involved in wind speed prediction and improves the quality of decomposition beyond what is achievable with traditional single-method approaches.
- Presents a hybrid model for wind speed forecasting that combines VMD, SGMD, and LSTM networks. The model begins with a detailed secondary decomposition process to extract complete and accurate components.
- Uses LSTM for prediction: After decomposition, the LSTM algorithm is employed as a reliable prediction tool, utilizing

the direct forecasting strategy for its accuracy and simplicity in delivering stable forecasting results.

II. METHODOLOGY

A. Variational Mode Decomposition (VMD)

To overcome the difficulties related to noise sensitivity and sampling inherent in EMD, an alternative technique was introduced in [16], known as VMD. This technique can decompose a multi-component signal into several quasi-orthogonal intrinsic mode functions in a non-recursively way [21]. As a method that does not rely on recursion, the goal is to break down a signal into several quasi-orthogonal intrinsic modes, each having a restricted bandwidth. This unique technique of decomposition gives VMD an advantage in terms of resistance to noise and minimizing errors, especially compared to methods that involve recursive calculations [22]. Figure 1 shows the specific steps involved in VMD for a signal $f(t)$ with a certain dimension.

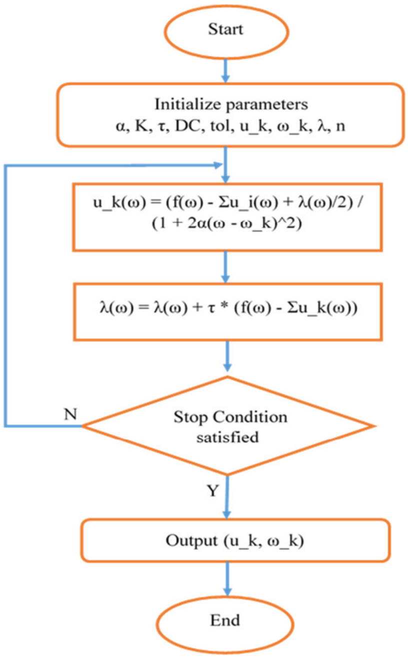

Fig. 1. The flowchart of the VMD algorithm.

B. Symplectic Geometry Mode Decomposition (SGMD)

SGMD is a mathematical technique used in signal processing, particularly for analyzing time series data. It was introduced to address common problems such as the sensitivity to artificial parameters and mode mixing found in decomposition models. SGMD is based on non-linear transformations, making it well-suited for analyzing dynamic systems. The SGMD method can be divided into four main stages: (i) adaptively establishing the embedding dimension of the time series, (ii) employing a symplectic geometry similarity transformation to determine the eigenvalues of the Hamiltonian matrix, (iii) implementing diagonal averaging, and (iv) adaptively reconstructing the components [17]. The main steps of SGMD are as follows.

1) Trajectory Matrix Constructing

For a given time series X of length n, consisting of elements x_1, x_2, ..., x_n, the trajectory matrix X is constructed following the Takens embedding theorem. This matrix X is structured as follows:

$$\mathrm{X} = \begin{bmatrix} x_1 & x_{1+\tau} & \cdots & x_{1+(d-1)\tau} \\ x_2 & x_{1+\tau} & \cdots & x_{2+(d-1)\tau} \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ x_m & x_{m+\tau} & \cdots & x_{m+(d-1)\tau} \end{bmatrix} \quad (1)$$

where d denotes the embedding dimension, τ is the time delay, and $m = n - (d-1)\tau$. The selection of values for d and τ follows specific adaptive rules [22]. The maximum peak frequency f_{max} is calculated from the initial time series Power Spectral Density (PSD). If the normalized frequency is below 10^{-3}, d is chosen to be $n/3$. Conversely, if it exceeds this threshold, d is set to 1.2 times the quotient of the sampling frequency F_s by f_{max} [17].

2) Symplectic Geometry Decomposition (SGD)

To reconstruct the Hamiltonian matrix, an autocorrelation analysis of the trajectory matrix is performed, resulting in the symmetric covariance matrix A, which is expressed as $\mathrm{A} = \mathrm{X}^T\mathrm{X}$. Next, the Hamilton matrix M is formed using the symmetric matrix A, in the structure [17]:

$$\mathrm{M} = \begin{bmatrix} \mathrm{A} & 0 \\ 0 & -\mathrm{A}^{\mathrm{T}} \end{bmatrix} \quad (2)$$

Once the Hamilton matrix M is constructed, its square, denoted as N ($\mathrm{N} = \mathrm{M}^2$), is calculated. Both matrices M and N qualify as Hamilton matrices according to the definition of a Hamilton matrix. Consequently, a symplectic orthogonal matrix Q is derived as per the equation [17]:

$$\mathrm{Q}^{\mathrm{T}}\mathrm{N}\mathrm{Q} = \begin{bmatrix} \mathrm{B} & \mathrm{R} \\ 0 & \mathrm{B}^{\mathrm{T}} \end{bmatrix} \quad (3)$$

Q is an orthogonal symplectic matrix that holds the special properties of symplectic matrices. This is crucial to ensure that the Hamilton matrix's structure is not altered during transformations. Here, B is set as an upper triangular matrix, defined such that $b_{ij} = (i > j + 1)$. By applying the Schmidt orthogonalization technique, this upper triangular matrix B is converted to matrix N, and its eigenvalues are determined as $\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_d$. If A is a real symmetric matrix, then its eigenvalues will be identical to those of B. The eigenvalues of the matrix A are calculated from the properties of the Hamilton matrix [17]:

$$\sigma_i = \sqrt{\lambda_i}\ (i = 1,2,3, \ldots, d) \quad (4)$$

The eigenvalues of the symmetric matrix A are derived based on the characteristics of the Hamilton matrix. The distribution of σ_i is arranged in a descending order, such that, $\sigma_1 > \sigma_2 > \cdot \cdot \cdot > \sigma_d$, with the smaller eigenvalues often considered to represent noise components. Then $\mathrm{S_i} = \mathrm{Q_i^T X^T}$ and $\mathrm{Z_i} = \mathrm{Q_i S_i}$ can be given using the eigenvector $\mathrm{Q_i}$, corresponding to the eigenvalue σ_i of A. The reconstruction matrix Z consists of the initial single component $\mathrm{Z_i}$ ($i = 1, 2, \ldots, d$), defined as follows [17]:

$$\mathrm{Z} = \mathrm{Z}_1 + \mathrm{Z}_2 + \cdots + \mathrm{Z}_\mathrm{d} \quad (5)$$

3) Diagonal Averaging Transformation

Initially, Z is defined as an $m \times d$ matrix. It is essential to reorganize its single components for optimal use. After this reorganization, Z is transformed by diagonal averaging, which forms a set of new time series, all measuring n in length. This results in the formation of d new time series, each with the same n length, and the combined total of these d series effectively restores the original time series.

For each initial single-component matrix Z_I, its elements are defined as z_{ij}, where i ranges from 1 to d and j from 1 to m. Then, d^* is defined as the smaller of m and d, and m^* is the larger of m and d, $n = m + (d-1)\tau$. When m is less than d, setting $z_{ij}^* = z_{ij}$, and if m is greater, then setting $z_{ij}^* = z_{ji}$. This can allow the application of the diagonal averaging transfer to the matrix [17].

$$y_k = \begin{cases} \frac{1}{k}\sum_{p=1}^{k} Z_{p,k-p+1}^* & 1 \le k \le d^* \\ \frac{1}{d^*}\sum_{p=1}^{d^*} Z_{p,k-p+1}^* & d^* < k \le m^* \\ \frac{1}{n-k+1}\sum_{p=k-m^*+1}^{n-m^*+1} Z_{p.k-p+1}^* & m^* < k \le n \end{cases} \quad (6)$$

The matrix Z_i is converted into a sequence of y_i $(y_1, y_2, \dots, y_n)$. Consequently, the reconstruction matrix Z is transformed into a new series of matrix Y, which has a length $d \times n$, by applying diagonal averaging. This process decomposes the original time series into d independent, superimposed components, each with distinct trends and frequency bands. The d single component signals are then obtained through diagonal averaging [17].

$$Y = Y_1 + Y_2 + \cdots + Y_d \quad (7)$$

4) Reconstruction

Through diagonal averaging, d decomposed components are obtained, which might not be entirely independent of each other. Components that share similar frequencies and exhibit high correlation need to be reconstructed into new components. Additionally, some noise elements may be present in these components due to environmental interference [17].

Therefore, the initial component SGC_1 is formed by combining Y_1, the first single component, with its associated components having the same similarity. The matrix left after removing the elements related to SGC_1 is termed G_1. The first column of G_1, along with its corresponding components, are then consolidated to create SGC_2. Following this, any residual signals are grouped into res, leading to the compilation of the final results as follows[17] :

$$S = \sum SGC(n) + res \quad (8)$$

where n is the number of iterations. The Normalized Mean Square Error (NMSE) of the res is calculated, and whenever it becomes less than the given threshold th=1%, the process is finished.

C. Long Short-Term Memory (LSTM) Network

LSTM networks, a subset of Recurrent Neural Networks (RNNs), are designed to address the limitations of traditional RNNs, particularly in processing long sequences of data. LSTMs are distinguished by their ability to remember information for extended periods, making them highly effective for a range of sequential data tasks in deep learning. This model serves as an intricate non-linear component essential for constructing deep neural networks [23].

The core architecture of an LSTM unit consists of a cell state and three gates: the input gate, the forget gate, and the output gate [24]. The cell state acts as a conveyor belt, carrying relevant information throughout sequence processing [25]. The gates in LSTM, controlled by sigmoid neural network layers, decide what information is to be stored, forgotten, or outputted at each step of the sequence. This structure allows LSTMs to mitigate the issues of vanishing and exploding gradients common in traditional RNNs [26]. The forget gate decides which information to discard by processing the previous output and current input through a sigmoid function, producing a value between zero and one. This value determines how much of the previous cell state is forgotten or retained. The input gate selects new information to store by passing the previous output and current input through a sigmoid function, with a tanh layer further refining the amount added to the cell state. The output gate controls the information that is sent out by combining the previous output and the current input, then processing them through a sigmoid function and tanh layer to produce the final output for the current time step [4].

D. The Hybrid VMD-SGMD-LSTM Model

This section presents the development of a hybrid model combining VMD, SGMD, and LSTM, aimed at forecasting wind speed series. The proposed model starts with applying VMD to the original series, separating it into two parts with differing frequency ranges, low and high. The high-frequency portion is then further analyzed using SGMD. Following these decomposition phases, the deep learning strategy of LSTM is implemented to make predictions based on these segmented components. Figure 2 shows the overall architecture of the model. The detailed method for this combined approach is presented in the following steps:

- Step 1: In VMD, the decomposition level is adjusted to two to extract the low-frequency portion from the initial dataset. This extraction results in the identification of the low-frequency component, which is achieved through the primary level of VMD decomposition. Concurrently, the high-frequency element is determined by calculating the difference between the original data series and the extracted low-frequency component. The two parameter keys in VMD are K, which represents the decomposition level, and the penalty factor (alpha), which is set in iterative testing by increasing or decreasing it until it gets the optimal value.
- Step 2: The segment labeled high-frequency, is further broken down into several symplectic geometry components, termed $SGC_1, SGC_2, \dots,$ $SGCm$, and includes a residual component as well. The selection of the decomposition

level m is adaptively determined based on the characteristics of the original data series.

- Step 3: The forecasting of the low-frequency component, the various SGCs, and the residual is accomplished using LSTM. The final forecast result is then calculated by summing all these individual forecasted values.

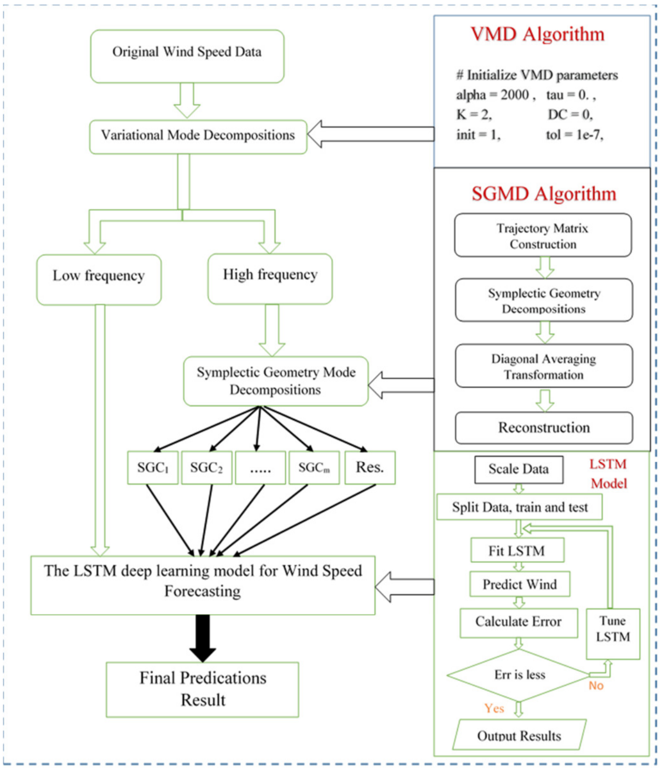

Fig. 2. Flowchart of the proposed VMD-SGMD-LSTM model.

III. CASE STUDY

A. Data Description

To verify the effectiveness of the proposed method, four wind speed datasets were collected from a wind farm in Chengde, China. The two 20-minute datasets come from two seasons, each containing 10 days (720 observations). These datasets were divided into two parts: the first 600 wind speed observations were the training set, and the last 120 were the testing set. The third wind speed dataset is the 10-min dataset for 10 days, and the last dataset is for 10-min wind speed for 90 days. The four wind speed datasets are called Dataset 1, Dataset 2, Dataset 3, and Dataset 4. Figure 3 shows the wind speed for Dataset 4. The LSTM model summary is four layers, 2,816,451 trainable params, 80 epochs, Adam optimizer, and MSE as the loss function.

TABLE I. STATISTICAL METRICS FOR THE DATASETS

Dataset	Maximum	Minimum	Mean	Median	Std.
Dataset 1	12.37	0.36	5.94	5.98	2.26
Dataset 2	21.02	0.34	7.81	7.5	3.97
Dataset 3	9.23	3.14	5.98	6.27	1.59
Dataset 4	30.37	0.10	9.79	9.21	4.977

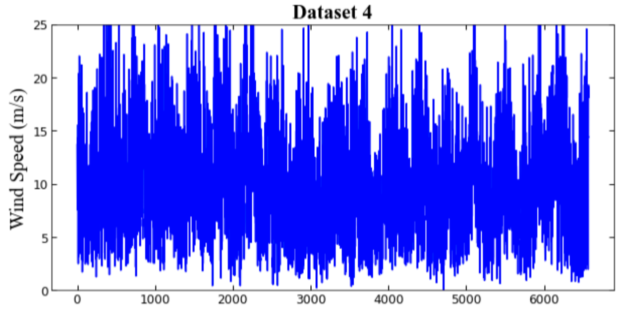

Fig. 3. Dataset 4 of wind speed data.

B. Performance Evaluation Metrics

Three key performance evaluation metrics were employed to assess the effectiveness of various models: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE). The definitions of these metrics are as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (9)$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i-\hat{y}_i}{\hat{y}_i}\right| \quad (10)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (11)$$

where n is the number of data tested, y_i is the value of the i^{th} prediction and $\hat{y}_i$ is the i^{th} true value.

In addition, to further compare the prediction ability of different models, three promoting percentages P_{MAE}, P_{RMSE}, and P_{MAPE}, which respectively denote the promoting percentages of MAE, RMSE, and MAPE, are provided as follows:

$$P_{MAE} = \left|\frac{(MAE_1 - MAE_2)}{MAE_1}\right| \quad (12)$$

$$P_{MAPE} = \left|\frac{(MAPE_1 - MAPE_2)}{MAPE_1}\right| \quad (13)$$

$$P_{RMSE} = \left|\frac{(RMSE_1 - RMSE_2)}{RMSE_1}\right| \quad (14)$$

where MAE_1 and MAE_2 represent the MAE of the initial and the proposed model, respectively. $MAPE_1$ and $MAPE_2$ represent the MAPE of the initial and the proposed model, respectively. $RMSE_1$ represents the RMSE of the initial model and $RMSE_2$ represents the RMSE of the proposed model.

In addition, theoretical and practical analysis was applied to validate and justify the performance of the proposed SD model. SGMD, which relies on symplectic geometry's similarity transformation instead of Euclidean geometry, is well-suited for non-linear analysis due to its non-linear transformation basis. However, during SGMD's traditional reconstruction process, similar components may be separated excessively, leading to over-decomposition. For example, using only SGMD for Dataset 4 produces 182 SGC components and a residual. Such a high number of components escalates computational resource usage and reduces the method's practicality. Therefore, the secondary decomposition method initially uses VMD to isolate the low-frequency elements of the original signal, reducing data complexity. Subsequently, SGMD is applied to the remaining signal, producing a manageable number of components. In the case of Dataset 1, the SDA produces five SGCs and a residual. This secondary decomposition technique effectively harnesses the strengths of both VMD and SGMD to enhance the decomposition quality while also being mindful of computational resource usage. Consequently, it achieves a balance between decomposition effectiveness and computational efficiency.

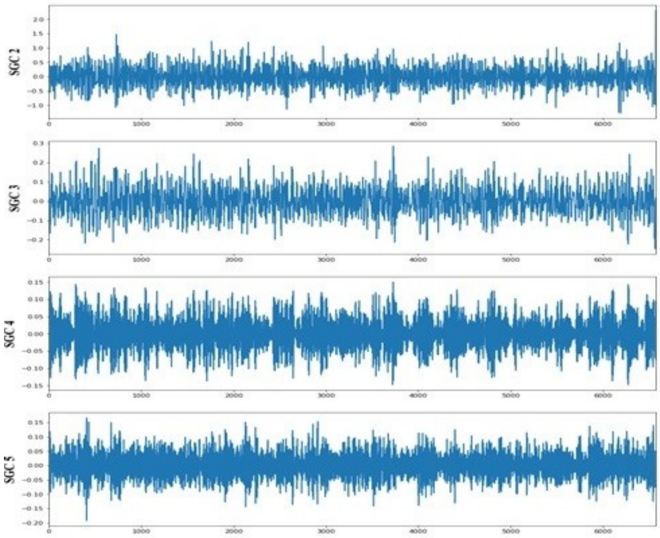

Fig. 4. The decomposition components for Dataset 4 obtained by the secondary decomposition technique.

IV. RESULTS

A. Comparative Analysis of Related Models

The proposed model was compared with two different models, including LSTM, and VMD-LSTM. These models were used for multi-step ahead forecasting, spanning from 1-step to 2-step predictions, allowing for a more thorough evaluation of each model's performance. Figures 5-6 show the forecasting results achieved by the proposed model for 1-2 step ahead predictions, covering Dataset 3 and Dataset 4. These figures show that the model's predictions closely align with the actual wind speed trends, showcasing superior forecasting capabilities in both single-step and multi-step scenarios. Table II presents the error metrics calculated for the various models for the four wind speed datasets. When comparing the proposed model with others, it consistently exhibits the smallest forecasting errors across all three indexes, regardless of the unique characteristics and complexity of the data series. Table III highlights the performance enhancement metrics for 1-2 step forecasting, illustrating the performance improvement of the proposed model. Table III shows that the proposed model achieved a considerable reduction in error rates for the four wind speed datasets. This marked improvement in forecasting accuracy highlights the effectiveness of the proposed model in multi-step wind speed prediction. The wind speed series in Dataset 3 was selected as a representative case.

TABLE II. PERFORMANCE COMPARISON

Dataset	**Models**	**MAPE (%)**		**RMSE(m/s)**		**MAE(m/s)**	
		1step	**2step**	**1step**	**2step**	**1step**	**2step**
Dataset1	LSTM	4.26	4.49	0.36	0.38	0.12	0.2
	VMD-LSTM	3.14	3.84	0.22	0.27	0.09	0.14
	VMD-SGMD-LSTM	0.84	1.16	0.09	0.12	0.06	0.07
Dataset2	LSTM	5.96	6.65	2.14	2.25	0.43	0.45
	VMD-LSTM	4.84	4.89	0.49	0.58	0.37	0.47
	VMD-SGMD-LSTM	3.7	3.96	0.14	0.21	0.11	0.18
Dataset3	LSTM	5.46	5.59	1.89	1.94	0.48	0.5
	VMD-LSTM	4.68	4.76	1.23	1.29	0.36	0.39
	VMD-SGMD-LSTM	3.14	3.18	0.25	0.31	0.23	0.26
Dataset4	LSTM	4.43	4.51	1.93	1.99	0.22	0.27
	VMD-LSTM	3.23	3.29	1.09	1.16	0.3	0.32
	VMD-SGMD-LSTM	1.07	1.28	0.11	0.19	0.06	0.09

TABLE III. PERFORMANCE-PROMOTING PERCENTAGE OF THE PROPOSED MODEL COMPARED TO OTHERS

Dataset	**VMD-SGMD-LSTM vs.**	**P_{MAPE}**		**P_{RMSE}**		**P_{MAE}**	
		1step	**2step**	**1step**	**2step**	**1step**	**2step**
Dataset1	LSTM	80.28	74.16	75	68.42	50	65
	VMD-LSTM	73.25	69.79	59.09	55.56	33.33	50
Dataset2	LSTM	37.92	40.45	93.46	90.67	74.42	60
	VMD-LSTM	23.55	19.02	71.43	63.79	70.27	61.7
Dataset3	LSTM	42.49	43.11	86.77	84.02	52.08	48
	VMD-LSTM	32.91	33.19	79.67	75.97	36.11	33.33
Dataset4	LSTM	75.85	71.62	94.3	90.45	72.73	66.67
	VMD-LSTM	66.87	61.09	89.91	83.62	80	71.88

A more in-depth analysis of the prediction outcomes is presented as follows:

- Comparing the LSTM and VMD-LSTM models with the proposed, the superiority of the latter is clear. For Dataset 1, the MAPE values for 1-2 step ahead forecasting using the proposed method are 0.84 and 1.16, respectively. Compared to the LSTM model, the proposed model significantly improves performance, reducing MAPE, RMSE, and MAE by 80%, 75%, and 50% for 1-step ahead, and 74%, 68%, and 65% for 2-step ahead forecasting. These improvements, including a PMAPE of 80.28% and 74.16% for 1-2 step forecasting, highlight the proposed model's enhanced predictive accuracy.
- The proposed VMD-SGMD-LSTM model outperforms the VMD-LSTM model, especially when forecasting 1-2 steps ahead. For example, in Dataset 3, the VMD-SGMD-LSTM model improved forecasting accuracy by 32.91% and 33.19% in MAPE, 79.67% and 75.97% in RMSE, and 36.11% and 33.33% in MAE compared to the VMD-LSTM model. These results indicate that the VMD-SGMD-LSTM model is significantly more effective in data decomposition for predictions.
- The experimental results show that model accuracy decreases with longer forecasting steps due to reduced data correlation in wind speed predictions. However, the VMD-SGMD-LSTM model shows minimal variability, highlighting its strong robustness in forecasting.
- The use of the VMD algorithm improved accuracy compared to single models, and the VMD-SGMD secondary decomposition further enhanced prediction

accuracy. Overall, the VMD-SGMD-LSTM model proves to be highly effective in interpreting wind speed time and is well-suited for forecasting.

Figures 7-9 display the error metrics for the 1-2-step ahead predictions for these models.

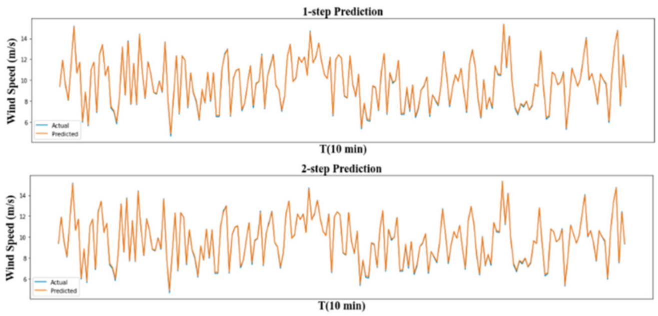

Fig. 5. The 1–2 step ahead forecasting results for Dataset 3 obtained by the proposed model.

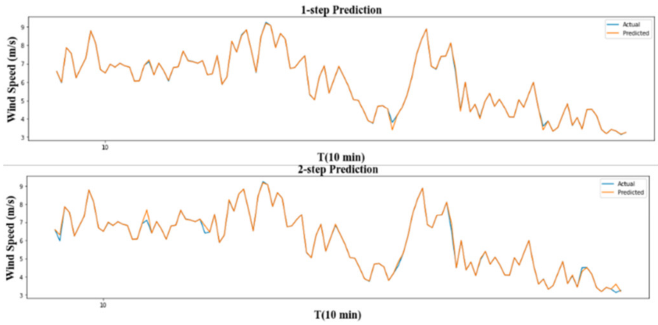

Fig. 6. The 1–2 step ahead forecasting results for Dataset 4 obtained by the proposed model.

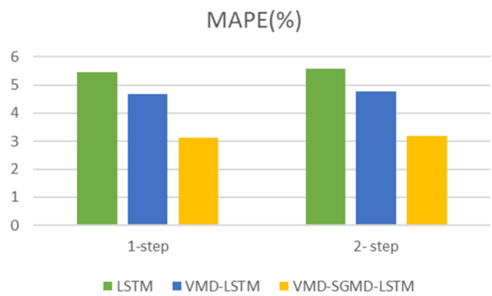

Fig. 7. MAPE values of different prediction steps for Dataset 3.

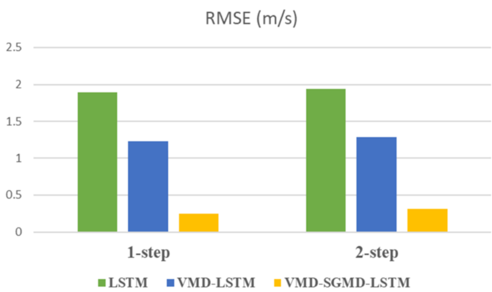

Fig. 8. RMSE values of different prediction steps for Dataset 3.

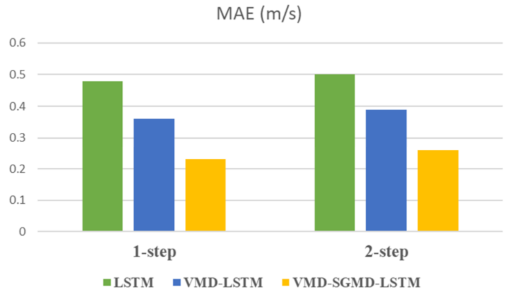

Fig. 9. MAE values of different prediction steps for Dataset 3.

B. Comparative Analysis of Published Models

This section presents an in-depth comparison between the proposed hybrid VMD-SGMD-LSTM model and three classical models: WT-VMD [27], EMD-WPD [12], CEEMDAN-VMD [28], and SSA-VMD [29]. This comparison aimed at thoroughly validating the performance of the proposed model. Secondary decomposition models demonstrate superior performance compared to their single decomposition counterparts, a fact that has been substantiated in the existing literature. In this study, the LSTM method was employed as a standard predictor to maintain the integrity and fairness of the comparative analysis. The effectiveness and consistency of the LSTM model as a predictive tool have been validated through the earlier detailed experiments. It was observed that the quality of the decomposition directly influences the overall predictive performance of the combined model, given the uniformity in the prediction method. Table IV provides error metrics and statistical tests to quantitatively assess the predictive prowess of the various models.

Predictive accuracy was evaluated using three primary indicators: MAPE, RMSE, and MAE. Table IV indicates that the hybrid VMD-SGMD-LSTM model outperformed the other models in multi-step forecasting across all datasets. For instance, considering Dataset 1, the MAPE values for 1-2 step forecasts using the VMD-SGMD-based model were 0.84% and 1.16 %, while those for the CEEMDAN-VMD-based model were 2.73%, and 2.81 %, respectively. These findings highlight the robustness and competitive nature of the decomposition process in the VMD-SGMD model, especially when compared to other secondary decomposition methods.

TABLE IV. COMPARISON RESULTS OF THE VMD-SGMD-LSTM MODEL AND FOUR OTHER ALGORITHMS

Dataset	Models	MAPE (%)		RMSE (m/s)		MAE(m/s)	
		1step	2step	1step	2step	1step	2step
Dataset 1	WT-VMD	3.09	5.45	0.18	0.36	0.15	0.26
	EMD-WPD	2.01	3.68	0.13	0.26	0.10	0.18
	CEEMDAN-VMD	2.73	2.81	0.18	0.16	0.13	0.13
	SSA-VMD	4.23	4.22	0.26	0.26	0.20	0.20
	Proposed model	**0.84**	**1.16**	**0.09**	**0.12**	**0.069**	**0.07**
Dataset 2	WT-VMD	7.16	8.39	0.30	0.42	0.25	0.32
	EMD-WPD	4.29	7.98	0.22	0.38	0.17	0.28
	EEMDAN-VMD	6.83	8.72	0.28	0.32	0.22	0.25
	SSA-VMD	9.33	9.80	0.34	0.36	0.27	0.28
	Proposed model	**3.70**	**5.96**	**0.14**	**0.27**	**0.11**	**0.13**

V. CONCLUSION

Considering the complex characteristics of wind speed, a novel hybrid forecasting system was proposed, designed for the multi-step prediction of wind speed. The model includes a secondary decomposition technique that combines VMD and SGMD for processing complex wind speed data, resulting in thorough and accurate decomposition results. The model starts by using VMD to extract low-frequency and high-frequency components, then the high-frequency is further decomposed into symplectic geometry components (SGC1, SGC2, ..., SGCm) and a residual, then the LSTM is used to forecast these components, and the final result is obtained by summing up the forecasts. Various forecasting models were compared to assess the system's efficacy. The results indicate that the combined strategy of the proposed model outperformed others in terms of accuracy and stability. The SGMD decomposition method effectively addresses the limitations of the VMD model, making it a valuable tool for handling intricate and high-frequency data series. In contrast, the developed secondary decomposition technique produces more robust and detailed results compared to alternative decomposition methods. Compared to WT-VMD, EMD-WPD, and CEEMDAN-VMD, it was found that the proposed forecasting model VMD-SGMD-LSTM performed optimally. Compared to EMD-WPD, the proposed model improved MAE by 31% and 61.1% for 1-2 step forecasting in Dataset 1, and 35.29% and 53.57% for 1-2 step forecasting in Dataset 2. Finally, the empirical study confirms that the proposed model is well-suited for predicting wind speed series over multiple steps.

REFERENCES

[1] A. Lodge and X. H. Yu, "Short term wind speed prediction using artificial neural networks," in *2014 4th IEEE International Conference on Information Science and Technology*, Shenzhen, China, Apr. 2014, pp. 539–542, https://doi.org/10.1109/ICIST.2014.6920535.

[2] F. Li, G. Ren, and J. Lee, "Multi-step wind speed prediction based on turbulence intensity and hybrid deep neural networks," *Energy Conversion and Management*, vol. 186, pp. 306–322, Apr. 2019, https://doi.org/10.1016/j.enconman.2019.02.045.

[3] N. O. Bushara and A. Abraham, "Using Adaptive Neuro-Fuzzy Inference System (ANFIS) to Improve the Long-term Rainfall Forecasting," *Journal of Network and Innovative Computing*, vol. 3, Jan. 2015.

[4] P. P. Prabha, V. Vanitha, and R. Resmi, "Wind Speed Forecasting using Long Short Term Memory Networks," in *2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT)*, Kannur,Kerala, India, Jul. 2019, pp. 1310–1314, https://doi.org/10.1109/ICICICT46008.2019.8993279.

[5] C. Matara, S. Osano, A. O. Yusuf, and E. O. Aketch, "Prediction of Vehicle-induced Air Pollution based on Advanced Machine Learning Models," *Engineering, Technology & Applied Science Research*, vol. 14, no. 1, pp. 12837–12843, Feb. 2024, https://doi.org/10.48084/etasr.6678.

[6] H. Liu and C. Chen, "Data processing strategies in wind energy forecasting models and applications: A comprehensive review," *Applied Energy*, vol. 249, pp. 392–408, Sep. 2019, https://doi.org/10.1016/j.apenergy.2019.04.188.

[7] Y. Li, H. Wu, and H. Liu, "Multi-step wind speed forecasting using EWT decomposition, LSTM principal computing, RELM subordinate computing and IEWT reconstruction," *Energy Conversion and Management*, vol. 167, pp. 203–219, Jul. 2018, https://doi.org/10.1016/j.enconman.2018.04.082.

[8] D. Liu, D. Niu, H. Wang, and L. Fan, "Short-term wind speed forecasting using wavelet transform and support vector machines optimized by genetic algorithm," *Renewable Energy*, vol. 62, pp. 592–597, Feb. 2014, https://doi.org/10.1016/j.renene.2013.08.011.

[9] Y. Ren, P. N. Suganthan, and N. Srikanth, "A Comparative Study of Empirical Mode Decomposition-Based Short-Term Wind Speed Forecasting Methods," *IEEE Transactions on Sustainable Energy*, vol. 6, no. 1, pp. 236–244, Jan. 2015, https://doi.org/10.1109/TSTE.2014.2365580.

[10] Y. Li, H. Shi, F. Han, Z. Duan, and H. Liu, "Smart wind speed forecasting approach using various boosting algorithms, big multi-step forecasting strategy," *Renewable Energy*, vol. 135, pp. 540–553, May 2019, https://doi.org/10.1016/j.renene.2018.12.035.

[11] S. Sun, L. Wei, J. Xu, and Z. Jin, "A New Wind Speed Forecasting Modeling Strategy Using Two-Stage Decomposition, Feature Selection and DAWNN," *Energies*, vol. 12, no. 3, Jan. 2019, Art. no. 334, https://doi.org/10.3390/en12030334.

[12] H. Yin *et al.*, "An effective secondary decomposition approach for wind power forecasting using extreme learning machine trained by crisscross optimization," *Energy Conversion and Management*, vol. 150, pp. 108–121, Oct. 2017, https://doi.org/10.1016/j.enconman.2017.08.014.

[13] W. Fu, K. Zhang, K. Wang, B. Wen, P. Fang, and F. Zou, "A hybrid approach for multi-step wind speed forecasting based on two-layer decomposition, improved hybrid DE-HHO optimization and KELM," *Renewable Energy*, vol. 164, pp. 211–229, Feb. 2021, https://doi.org/10.1016/j.renene.2020.09.078.

[14] H. Pan, Y. Yang, J. Zheng, and J. Cheng, "A noise reduction method of symplectic singular mode decomposition based on Lagrange multiplier," *Mechanical Systems and Signal Processing*, vol. 133, Nov. 2019, Art. no. 106283, https://doi.org/10.1016/j.ymssp.2019.106283.

[15] Y. Jiang, W. Chen, and M. Li, "Symplectic geometry decomposition-based features for automatic epileptic seizure detection," *Computers in Biology and Medicine*, vol. 116, Jan. 2020, Art. no. 103549, https://doi.org/10.1016/j.compbiomed.2019.103549.

[16] K. Dragomiretskiy and D. Zosso, "Variational Mode Decomposition," *IEEE Transactions on Signal Processing*, vol. 62, no. 3, pp. 531–544, Oct. 2014, https://doi.org/10.1109/TSP.2013.2288675.

[17] H. Pan, Y. Yang, X. Li, J. Zheng, and J. Cheng, "Symplectic geometry mode decomposition and its application to rotating machinery compound fault diagnosis," *Mechanical Systems and Signal Processing*, vol. 114, pp. 189–211, Jan. 2019, https://doi.org/10.1016/j.ymssp.2018.05.019.

[18] H. Pan, Y. Yang, J. Zheng, and J. Cheng, "Symplectic transformation based Variational Bayesian Learning and its applications to gear fault diagnosis," *Measurement*, vol. 147, Dec. 2019, Art. no. 106827, https://doi.org/10.1016/j.measurement.2019.07.055.

[19] Z. Cheng and R. Wang, "Enhanced symplectic characteristics mode decomposition method and its application in fault diagnosis of rolling bearing," *Measurement*, vol. 166, Dec. 2020, Art. no. 108108, https://doi.org/10.1016/j.measurement.2020.108108.

[20] K. I. Taher and A. M. Abdulazeez, "Deep Learning Convolutional Neural Network for Speech Recognition: A Review," *International Journal of Science and Business*, vol. 5, no. 3, pp. 1–14, 2021.

[21] Y. Wang, R. Markert, J. Xiang, and W. Zheng, "Research on variational mode decomposition and its application in detecting rub-impact fault of the rotor system," *Mechanical Systems and Signal Processing*, vol. 60–61, pp. 243–251, Aug. 2015, https://doi.org/10.1016/j.ymssp.2015.02.020.

[22] P. Bonizzi, J. M. H. Karel, O. Meste, and R. L. M. Peeters, "Singular Spectrum Decomposition: A New Method for Time Series Decomposition," *Advances in Adaptive Data Analysis*, vol. 06, no. 04, Oct. 2014, Art. no. 1450011, https://doi.org/10.1142/S1793536914500113.

[23] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, https://doi.org/10.1162/neco.1997.9.8.1735.

[24] A. K. Ibrahim, M. M. Hassan, and I. A. Ali, "Smart Homes for Disabled People: A Review Study," *Science Journal of University of Zakho*, vol. 10, no. 4, pp. 213–221, Nov. 2022, https://doi.org/10.25271/sjuoz.2022.10.4.1038.

[25] H. T. Sadeeq, S. Y. Ameen, and A. M. Abdulazeez, "Cancer Diagnosis based on Artificial Intelligence, Machine Learning, and Deep Learning," in *2022 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT)*, Sakheer, Bahrain, Nov. 2022, pp. 656–661, https://doi.org/10.1109/3ICT56508.2022.9990784.

[26] S. V. Belavadi, S. Rajagopal, R. R, and R. Mohan, "Air Quality Forecasting using LSTM RNN and Wireless Sensor Networks," *Procedia Computer Science*, vol. 170, pp. 241–248, Jan. 2020, https://doi.org/10.1016/j.procs.2020.03.036.

[27] H. Liu, Z. Duan, F. Han, and Y. Li, "Big multi-step wind speed forecasting model based on secondary decomposition, ensemble method and error correction algorithm," *Energy Conversion and Management*, vol. 156, pp. 525–541, Jan. 2018, https://doi.org/10.1016/j.enconman.2017.11.049.

[28] T. Peng, J. Zhou, C. Zhang, and Y. Zheng, "Multi-step ahead wind speed forecasting using a hybrid model based on two-stage decomposition technique and AdaBoost-extreme learning machine," *Energy Conversion and Management*, vol. 153, pp. 589–602, Dec. 2017, https://doi.org/10.1016/j.enconman.2017.10.021.

[29] L. Xiang, J. Li, A. Hu, and Y. Zhang, "Deterministic and probabilistic multi-step forecasting for short-term wind speed based on secondary decomposition and a deep learning method," *Energy Conversion and Management*, vol. 220, Sep. 2020, Art. no. 113098, https://doi.org/10.1016/j.enconman.2020.113098.