Acta Geodyn. Geomater., Vol. 20, No. 4 (212), 139–153, 2023
DOI: 10.13168/AGG.2023.0013
journal homepage: https://www.irsm.cas.cz/acta

ORIGINAL PAPER

AN IMPROVED DFVMD-LSTM HYBRID TIME SERIES MODEL FOR PREDICTING TOTAL SOLAR IRRADIANCE

Hongkang CHEN [1, 2], Tieding LU [1, 2], Xiaoxing HE [3] *, Xiwen SUN [1, 2], Jiahui HUANG [3], Jie WANG [3] and Shengbo YANG [3]

[1)] *School of Surveying and Geoinformation Engineering, East China University of Technology, Nanchang, 341000, China*
[2)] *Key Laboratory of Mine Environmental Monitoring and Improving around Poyang Lake of Ministry of Natural Resources, East China University of Technology, Nanchang 330013, China*
[3)] *School of Civil and Surveying and Mapping Engineering, Jiangxi University of Science and Technology, Ganzhou, 341000, China*

**Corresponding author's e-mail:* xxh@jxust.edu.cn

ARTICLE INFO

Article history:
Received 9 September 2023
Accepted 20 October 2023
Available online 26 October 2023

Keywords:
Total Solar Irradiance
Deep Learning
Time Series Prediction
VMD
LSTM

ABSTRACT

The prediction of total solar irradiance (TSI) time series holds significant importance in the study of solar activity and the assessment of solar energy resources. The variational mode decomposition-based long short-term memory (VMD-LSTM) model is a hybrid deep learning model that demonstrates high prediction accuracy on long-term time series. To address the information leakage issue faced by hybrid models based on VMD and other data preprocessing methods, this study proposes a prediction method for hybrid deep learning models called dual-fusion variational mode decomposition (DFVMD), which modifies the VMD decomposition approach. The DFVMD-LSTM model utilizes multiple TSI datasets as model features, and multisource datasets and multiple model comparison experiments are employed to verify the applicability and robustness of the model. The experimental results show that the DFVDM-LSTM model significantly reduces the periodic TSI prediction deviation introduced by the LSTM model. Furthermore, regardless of the training period or prediction horizon, the DFVMD-LSTM model exhibits an average root mean square error (*RMSE*) reduction of 14.79 % and an average mean absolute error (*MAE*) decrease of 21.50 %, demonstrating the superior predictive performance and improved reliability of the DFVDM-LSTM method.

1. INTRODUCTION

Solar radiation is the primary external energy source and the fundamental driver of the Earth's climate (Beer et al., 2007). Total solar irradiance (TSI), defined as the total amount of electromagnetic radiation energy derived from all solar spectral bands that reaches the Earth per unit time, is a crucial parameter in solar energy utilization and global meteorological research (Almorox et al., 2021; Mishchenko et al., 2007). As global energy demand continues to increase and environmental concerns become more pressing, solar energy, as a renewable, abundant, and clean energy source, is increasingly being adopted by human societies (Wan et al., 2015; Shatat et al., 2013). TSI time series forecasting has significant implications for solar activity research, solar energy resource assessment, meteorology, and environmental science (Herrera et al., 2015).

Currently, research on solar irradiance mainly focuses on predicting solar irradiance at the Earth's surface. Zang et al. (2020) (employed different hybrid deep learning models to predict solar irradiance in Texas, USA, and demonstrated that the hybrid long short-term memory (LSTM) convolutional neural network (CNN) model yielded higher prediction accuracy than the other tested models. Majumder et al. (2018) fused variational mode decomposition (VMD) with the extreme learning machine (ELM) algorithm to predict solar irradiance in India, and it was found that the hybrid model achieved significantly improved accuracy over the original model. Gao et al. (2020) combined complete ensemble empirical mode decomposition-based adaptive noise (CEEMDAN) with the hybrid CNN-LSTM model to predict hourly solar irradiance, and the results showed that the hybrid model outperformed the CNN-LSTM model in terms of prediction accuracy. These studies provide feasible prediction methods for solar irradiance time series and confirm that hybrid models have higher prediction accuracy than their initial models. However, the solar irradiance reaching the Earth's surface can be affected by the Earth's rotation and the atmosphere, resulting in inaccurate TSI data observed at ground-based platforms (Huang, 2008). To reduce the interference of environmental factors in ground-based monitoring, this paper selects fused long-term continuous TSI time series data from multiple satellite platforms for prediction purposes (Montillet et al., 2022; Dudok et al., 2017; Gueymard et al., 2018; Ball et al., 2012).

VMD is a signal processing method that decomposes nonstationary signals into multiple intrinsic mode functions (IMFs) and a residue (Wang

Cite this article as: Chen H, Lu T, He X, Sun X, Huang J, Wang J, Yang S: An improved DFVMD-LSTM hybrid time series model for predicting total solar irradiance. Acta Geodyn. Geomater., 20, No. 4 (212), 139–153, 2023. DOI: 10.13168/AGG.2023.0013

et al., 2015). LSTM is a recurrent neural network that is widely used for processing sequence data in natural language processing, speech recognition, time series prediction, and other fields (Hochreiter and Schmidhuber, 1997). Recently, many researchers have combined VMD with LSTM, which has resulted in high prediction accuracy for cases involving time series data, such as wind speed and photovoltaic power prediction (Han et al., 2019; Sun et al., 2019; Lv et al., 2021; Shi et al., 2018; Zhou et al., 2022; Chen et al., 2023). However, most researchers typically decompose their testing and training datasets together using VMD, which can result in information leakage and inflated prediction results. Furthermore, the IMFs obtained through VMD have low correlations with the original data, making them unsuitable for research fields with high correlations.

Based on these issues, this paper proposes a dual-fused VMD-LSTM (DFVMD-LSTM) model that avoids information leakage by modifying the VMD method. The dual-fusion approach extracts valuable information from the original signal, and the LSTM stores the state of each time step in multiple memory cells to achieve high-precision TSI prediction. Experiments are conducted by dividing two TSI time series into different training and prediction periods, and different datasets are used as features to compare the prediction accuracies of the LSTM, VMD-LSTM, and DFVMD-LSTM models. The results demonstrate the prediction accuracy improvement and the applicability and robustness of the DFVMD-LSTM model under different datasets.

The structure of this paper is as follows: In Section 2, ANN, VMD and LSTM algorithm principles and accuracy evaluation indexes are introduced, and DFVMD-LSTM model principles and specific processes are elaborated. In Section 3, the TSI data and data preprocessing strategies are introduced, and the prediction accuracy of DFVMD-LSTM model under different *K* values is analyzed. Section 4 focuses on the prediction results and accuracy of ANN model, LSTM model and DFVMD-LSTM mixed model, and analyzes the prediction results of the DFVMD-LSTM model under different accuracy evaluation indexes. The conclusion and prospect are given in Section 5.

2. MODELS AND EVALUATION INDICES

2.1. VMD

VMD is a completely non-recursive method for solving modal variation problem (Dragomiretskiyk and Zosso, 2014; ur Rehman and Aftab, 2019). The core principle of this method is to decompose a multicomponent signal into multiple single-component amplitude-modulated and frequency-modulated signals, thereby solving the endpoint effect and modal component overlap problems that exist in empirical mode decomposition (EMD). The main idea is to construct and solve a variational problem. The specific decomposition process is as follows (Wang et al., 2022; Zhang et al., 2017).

1. For each modal component, its corresponding analytic signal is computed through Hilbert transformation, and a one-sided frequency spectrum is accordingly obtained.

$$\left[\delta(t)+\frac{j}{\pi t}\right]*\mu_K(t) \tag{1}$$

In the above equation, $j^2=-1$, and δ is the Dirac distribution.

2. To estimate the center frequencies of individual modes and modulate their spectral components to the corresponding fundamental frequency bands, an exponential term is incorporated into the mode shape.

$$\left[\left(\delta(t)+\frac{j}{\pi t}\right)*\mu_K(t)\right]e^{-j\omega_K t} \tag{2}$$

3. The bandwidth of ω_K can be estimated by the smoothness of the H1 demodulated signal. This leads to a constrained variational problem.

$$\begin{cases}\min\limits_{\{\mu_K\},\{\omega_K\}}\left\{\sum\limits_K\left\|d_t\left[\left(\delta(t)+\frac{j}{\pi t}\right)*u_K(t)\right]e^{-j\omega_K t}\right\|_2^2\right\}\\ s.t.\sum\limits_K\mu_K=f\end{cases} \tag{3}$$

In the above equation, f represents the original signal, $\{\mu_K\}$ represents the decomposed modal functions, and $\{\omega_K\}$ represents the corresponding central frequency of each mode.

4. Based on this, we introduce a quadratic penalty factor α and a Lagrange multiplier λ_t to transform the problem into a non-constrained variational problem. The extended Lagrange expression is given below.

$$\begin{aligned}L(\{\mu_K\},\{\omega_K\},\lambda)=\alpha\sum_K\Big\|\partial_t\Big[\Big(\delta(t)+\frac{j}{\pi t}\Big)\\ *\mu_K(t)\Big]e^{-j\omega_K t}\Big\|_2^2\\ +\Big\|f(t)-\sum_K\mu_K(t)\Big\|_2^2\\ +\langle\lambda(t),f(t)-\sum_K\mu_K(t)\ \rangle\end{aligned} \tag{4}$$

In this equation, α is the quadratic penalty factor, and λ_t is the Lagrange multiplier. The unconstrained variational problem is then solved using the alternating direction method of multipliers, and the "saddle point" of the extended Lagrange expression, which is the optimal solution of the constrained variational model in Equation (3), is sought by alternately updating ${\mu_K}^{n+1}$, ${\omega_K}^{n+1}$, and λ^{n+1}.

2.2. ANN

An artificial neural network (ANN) is a complex network structure that abstracts, simplifies, and

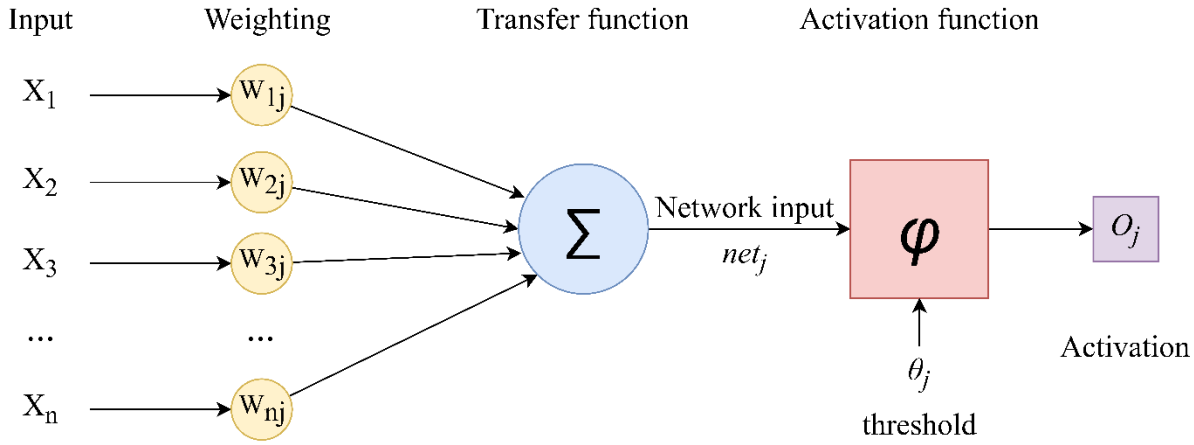

Fig. 1 Basic structure of an ANN.

simulates the neural network of the human brain. It is one of the main tools in machine learning (Zupan, 1994). An ANN is a nonlinear, versatile tool that is stable in applications such as classification and time series prediction. As a result, ANNs are widely used in fields that require mathematical and statistical analysis, such as statistics, metrology, medicine, and economics. As the most used prediction model in solar irradiance time series forecasting, an ANN serves as the baseline model for conducting a comparative analysis with the model proposed in this study (Pazikadin et al., 2020). This ANN is composed of interconnected layers of neurons or nodes. Information is transmitted between the neurons through weighted connections. In this study, the TSI time series is utilized as the training feature input to the ANN. The ANN then performs forward propagation to calculate the output values. Subsequently, the loss is computed, and backpropagation is employed to calculate the loss function. The Adam optimization algorithm is utilized to update the parameters. Afterward, the optimal model parameters are selected based on the validation set, and the final predictions are conducted. The neural structure of the ANN is illustrated in Figure 1.

As shown in Figure 1, X_1, X_2, ..., X_n represent the input values, while W_{1j}, W_{2j}, ..., W_{nj} correspond to their respective weights. The input signals are summed up in the unit Σ to obtain the weighted sum u_j.

$$u_j = \sum_{k=1}^{n} w_{kj} x_k \tag{5}$$

Subsequently, the weighted sum, after thresholding in θ_j, is input into the non-linear activation function φ to obtain the output activation state O_j.

$$v_j = net_j = u_j - \theta_j \tag{6}$$

$$O_j = \varphi(v_j) \tag{7}$$

2.3. LSTM

LSTM is an optimized variant of recurrent neural networks (RNN) that can better store and access historical data through memory modules, effectively avoiding the "gradient vanishing" and "gradient explosion" problems that often occur in RNNs (Sherstinsky, 2020; Yu et al., 2019). Compared to traditional methods, LSTM has strong advantages in modeling and predicting long time series and is widely used in time series prediction and fault detection tasks (Sagheer and Kotb, 2019; Yadav et al., 2020). The LSTM network structure in this paper consists of an input layer, 128 hidden layers, and an output layer. Each hidden layer is composed of memory blocks, as shown in Figure 2. Each memory block is composed of three gate structures: a forget gate, an input gate, and an output gate.

As depicted in the Figure2, LSTM handles high-temporal data concerning sea surface elevation and the previous moment's hidden state output via the operation of three gates. The core procedure can be summarized as follows:

1. Within the LSTM framework, the decision of whether to discard or preserve information pertaining to X_t and h_{t-1} is made by the forget gate, marked as f_t. This choice is regulated by the activation function σ associated with the forget gate.

$$f_t = \sigma(W_f \cdot [h_{t-1}, X_t] + b_f) \tag{8}$$

In the equations, W signifies the weight matrices, while b corresponds to the biases. f_t is a vector with elements ranging from 0 to 1, where each element in the vector signifies the extent of information preservation in the cell state C_{t-1}. A value of 0 indicates no preservation, whereas a value of 1 indicates complete preservation.

2. The cell state undergoes an update process facilitated by the input gate. This involves passing two components, X_t and h_{t-1}, through an

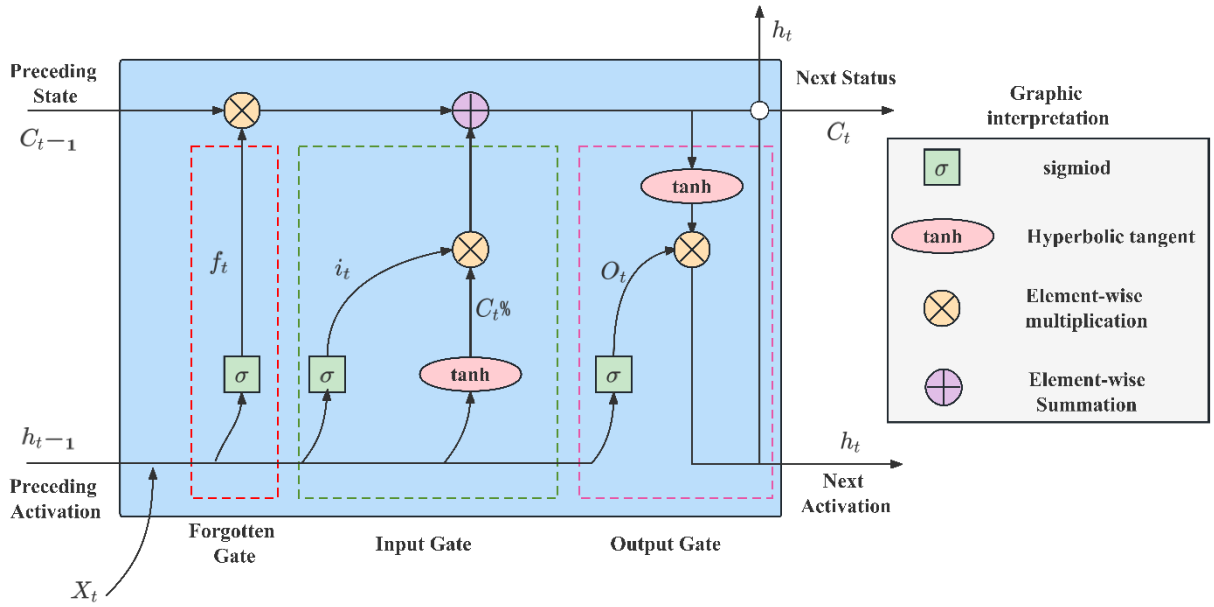

Fig. 2 Basic structure of LSTM.

activation function σ to ascertain the information update. X_t and h_{t-1} are further processed through a hyperbolic tangent (tanh) function to generate a new candidate value vector, C'_t (where C'_t is a vector with values in the range of -1 to 1), and the output from the tanh operation is then scaled by the multiplication factor σ.

$$i_t = \sigma(W_i \cdot [h_{t-1}, X_t] + b_i) \tag{9}$$

$$C'_t = tanh(W_c \cdot [h_{t-1}, X_t] + b_c) \tag{10}$$

3. The cell state from the preceding layer is subjected to an element-wise multiplication with the forget vector, followed by an element-wise addition to the output of the input gate. This process yields the updated cell state.

$$C_t = f_t * C_{t-1} + i_t * C'_t \tag{11}$$

In the equations, the component $f_t * C_{t-1}$ determines the extent to which information in the previous memory cell state C_{t-1} is forgotten, while $i_t * C'_t$ determines the extent to which the information in C'_t is added to the new memory cell state C_t.

4. By means of the output gate O_t, the value of the subsequent hidden state h_t is established, and this hidden state incorporates information from prior inputs.

$$O_t = \sigma(W_O \cdot [h_{t-1}, X_t] + b_O) \tag{12}$$

$$h_t = O_t * tanh(C_t) \tag{13}$$

2.4. DUAL-FUSION VARIATIONAL MODE DECOMPOSITION-BASED LONG SHORT-TERM MEMORY NETWORK (DFVMD-LSTM) MODEL

As a classic hybrid model, the VMD-LSTM model has been widely used in time series prediction cases, such as photovoltaic prediction and wind speed prediction (Han et al., 2019; Sun et al., 2019; Lv et al., 2021; Shi et al., 2018; Zhou et al., 2022; Liao et al., 2021; Jin et al., 2020). However, most researchers currently directly decompose the training set, validation set, and test set together through VMD in hybrid algorithms and then use the decomposed modal components as model inputs for prediction, adding up the predicted results of each component to obtain the final prediction result. However, this approach has two shortcomings (Wang et al., 2016; Qian et al., 2014).

(1) When using traditional hybrid deep learning algorithms for decomposition, all data, including the relevant data to be predicted, are decomposed. However, when VMD decomposes the entire dataset, it can result in test set decompositions being influenced by the training set data, leading to a certain degree of information leakage. (2) Traditional hybrid deep learning methods for time series prediction typically involve autoregressive prediction of the IMFs obtained after VMD decomposition. However, the IMFs and residuals obtained after VMD decomposition often have weaker correlations with the original data. This may result in poorer prediction performance for some less applicable models, which is not conducive to multi-feature forecasting.

If the training and validation sets are separated for VMD, it is necessary to establish N prediction models (N = number of validations sets) due to the sliding nature of prediction to address the information leakage issue. However, this approach significantly increases the computational complexity of the model, as the gradient is steepened (Wang et al., 2016; Qian et al., 2014).

The proposed DFVMD-LSTM model solves information leakage and the other problems that exist in the VMD-LSTM model, and the fused time series have high correlations with the original time series.

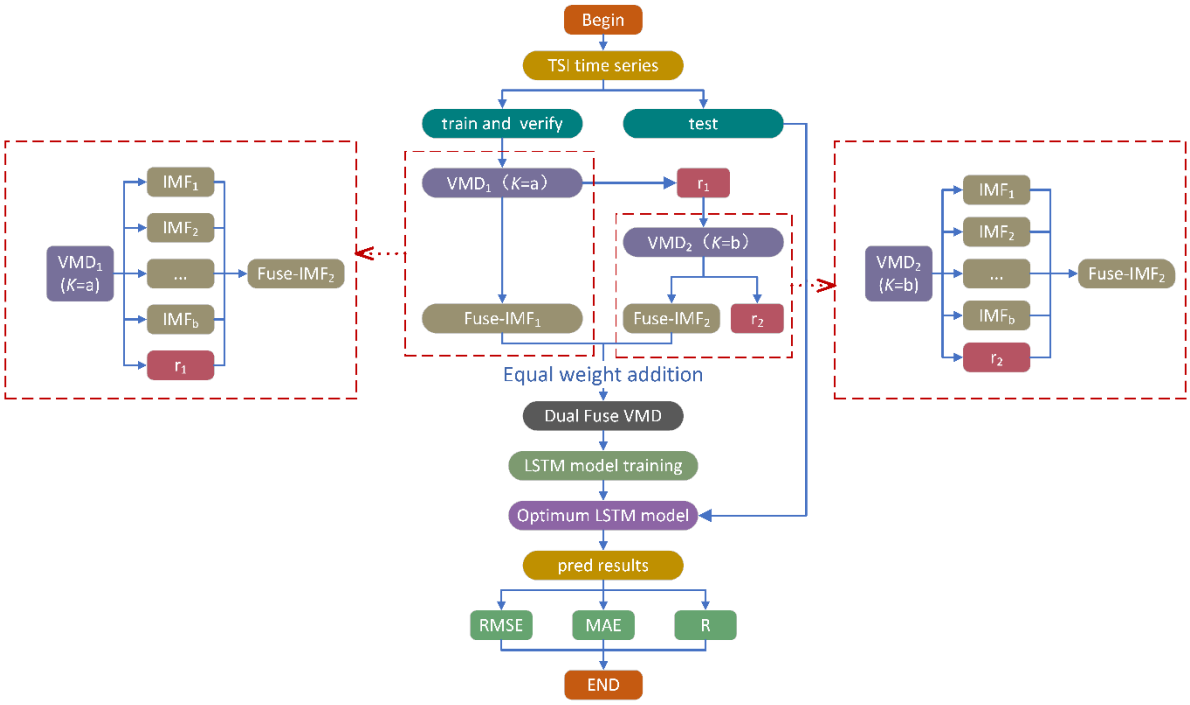

Fig. 3 Hybrid DFVMD-LSTM model prediction process.

The specific process is shown in Figure 3.

DFVMD-LSTM model prediction process:

Step 1: The input TSI time series is linearly interpolated to fill in missing values. The resulting time series is divided into a training set, a validation set, and a test set for experimentation.

Step 2: The daily TSI data in the training and validation sets are decomposed into different IMFs and a residual term r1 using VMD with different *K* values.

Step 3: The decomposed IMFs are added together to obtain a fused IMF, Fuse-IMF$_1$. The residual term r$_1$ is then decomposed again using VMD with different *K* values to obtain a new set of IMFs and a residual term r$_2$.

Step 4: The residual r$_1$ obtained after performing decomposition is added to the individual mode components to obtain Fuse-IMF$_2$. Then, Fuse-IMF1 and Fuse-IMF$_2$ are added together to obtain the dual-fuse VMD results.

Step 5: The dual fuse VMD results are used as the model feature values and input into the LSTM model for training. The optimal LSTM model is obtained through iterative optimization.

Step 6: The optimal LSTM model is selected to predict the data in the test set. The predicted results obtained from the LSTM model represent the predictions of the DFVMD-LSTM model. The predicted results are then evaluated for accuracy.

2.5. PRECISION EVALUATION INDICES

This study primarily utilizes three performance evaluation metrics to assess the accuracy of the models: Root Mean Squared Error (RMSE) (Chai and Draxler, 2014), Mean Absolute Error (MAE) (Willmott and Matsuura, 2005), and the Pearson correlation coefficient (*R*) (Adler and Parmryd, 2010). The *RMSE*, *MAE* and *R* are calculated using the following formulae, respectively.

(1) Root Mean Squared Error (*RMSE*)

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{14}$$

(2) Mean Absolute Error (*MAE*)

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|(y_i - \hat{y}_i)| \tag{15}$$

According to the above formula, we consider the actual values of TSI denoted as y_i and the predicted values from various models denoted as $\hat{y}_i$, where n represents the number of TSI data points. For *RMSE* and *MAE*, smaller values indicate higher prediction accuracy of the model, while larger values indicate lower accuracy. The applicability of these metrics may decrease in time series analysis.

(3) Pearson correlation coefficient (*R*)

$$R_{Y\hat{Y}} = \frac{E(Y\hat{Y}) - E(Y)E(\hat{Y})}{\sigma_Y \sigma_{\hat{Y}}} \tag{16}$$

In the above equation, Y represents the original time series, while $\hat{Y}$ represents the time series resulting from applying the DFVMD method or the predicted time series. The Pearson correlation coefficient has a range of -1 to 1.A positive value of $R_{Y\hat{Y}}$ implies a direct correlation between the two temporal sequences, whereas a negative value of $R_{Y\hat{Y}}$ signifies an inverse relationship.

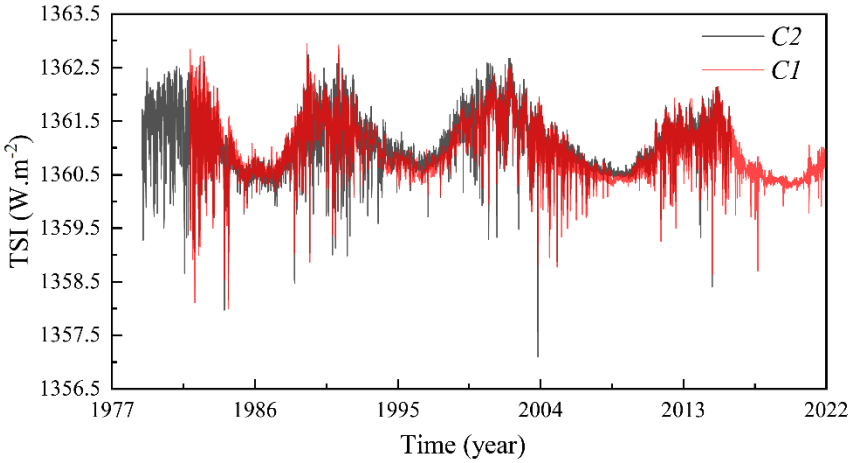

Fig. 4 *C1* and *C2* data distributions.

To more intuitively demonstrate the degree of improvement exhibited by the model, this paper determines the magnitude of the improvement achieved for each accuracy evaluation index. The calculation formula is as follows (Li et al., 2023):

$$I_{y\hat{y}} = \frac{y-\hat{y}}{y} \tag{17}$$

In the above equation, y and $\hat{y}$ represent the three accuracy evaluation indicators: the *RMSE*, the *MAE*, and *R*. y represents the accuracy evaluation indicators of the initial model prediction results, while $\hat{y}$ represents the accuracy evaluation indicators of the mixed model prediction results. The larger $I_{y\hat{y}}$ is, the greater the improvement yielded by the mixed model in the corresponding evaluation indicator, and vice versa.

3. DATA AND EXPERIMENTS

3.1. DATA PREPROCESSING

This study uses two sets of TSI data, one from the Physikalisch-Meteorologisches Observatorium Davos/World Radiation Center (https://www.pmodwrc.ch/en/research-development/solar-physics/tsi-composite/) and another from the International Space Science Institution (http://www.issibern.ch/teams/solarirradiance) (Montillet et al., 2022; Dudok et al., 2017). Since the data are obtained from a mixture of multiple sources, they are referred to as Composite 1 (*C1*) and Composite 2 (*C2*) in this study. *C1* includes 14,601 daily data points from 1981 to 2021, with a missing data rate of 0.068 %. *C2* includes 13,559 daily data points from 1978 to 2015, with no missing data. Since the data missing rate is small, linear interpolation is used in this study to handle missing values. The processed data is shown in Figure 4.

As shown in Figure 4, the TSI data used in this study exhibit significant fluctuation characteristics with clear amplitudes and periods. Based on this, deep learning models can capture the underlying patterns of these fluctuations, thereby achieving better prediction results.

3.2. DFVMD-LSTM MODEL ANALYSIS

In theory, the VMD can capture all frequency components present in different datasets and decompose them into various mode components. The residual term should ideally contain only random signals without any oscillatory characteristics. However, in practical applications, due to limitations in the VMD algorithm or the characteristics of the signal itself, VMD may not fully decompose the data, and the residual term may still retain certain oscillatory features. (Li et al., 2017; Li et al., 2022). Unlike the EMD method, VMD allows the number of IMFs to be fixed for decomposing the original data, making it particularly important to determine a suitable number of IMFs, which is denoted as *K*. In theory, any positive integer can be selected as *K*. However, when *K* is too large, the generation of virtual components will cause the sum of the energies of each component to be too high, and a large *K* value will result in gradient ascent for the model load. Therefore, in practice, *K* should not be too large. When *K* is too small, too much important information in the original signal will be filtered out, resulting in signal under decomposition and affecting the prediction accuracy of the model.

In reference (Han et al., 2019), experiments were conducted with VMD using *K* values of 3 and 5, and it was found that excessive decomposition layers could lead to certain issues. In reference (Sun et al., 2019), VMD was performed on wind speed time series with *K* set to 5, followed by subsequent experiments. In reference (Lv et al., 2021), VMD experiments were conducted using *K* values ranging from 5 to 12, and the optimal parameter was determined to be 7 through the center frequency method.

Considering that the setting of the crucial parameter *K* in VMD can significantly affect the data decomposition results, this paper conducts the

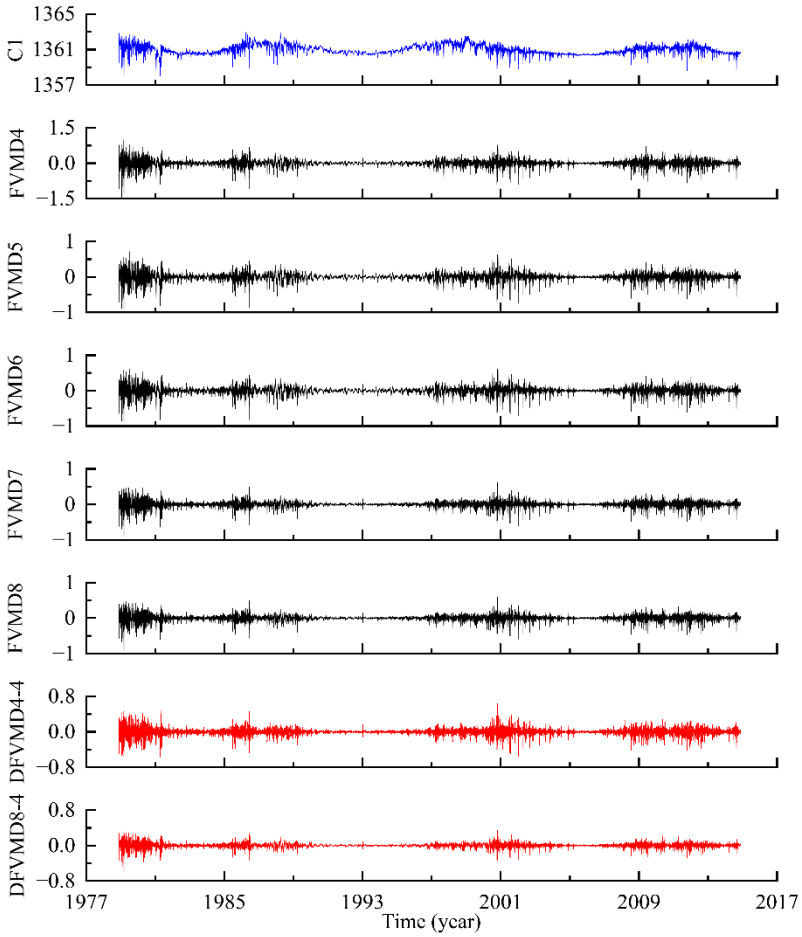

Fig. 5 Residual Terms of FVMD and DFVMD with Different *K* Values (The blue curve represents TSI time series *C1*. The black curves depict the residual terms obtained from FVMD with different values of *K*, where FVMD4 indicates VMD decomposition with *K* set to 4. The red curves represent the residual terms obtained from DFVMD with different values of *K*, where DFVMD8-4 indicates the first VMD decomposition with *K* set to 8 and the second VMD decomposition with *K* set to 4. The units in the figure are all $W.m^{-2}$).

decomposition of the *C1* dataset using different values of *K*. The aim is to investigate and discuss the oscillatory characteristics present in the residual terms under different *K* values for both the Fusion Variational Mode Decomposition (FVMD) and the DFVMD methods. The oscillatory features in the residual terms obtained by decomposing the *C1* dataset with different K values are shown in Figure 5.

The figure shows that "FVMD" refers to the single fusion variational mode decomposition, which means that after a single VMD decomposition, the residual term is removed, and the IMFs are directly summed with equal weights to obtain the FVMD result. As shown in Figure 5, it is evident that different values of *K* used in VMD decomposition exhibit certain oscillatory trends, and these trends are consistent with the original time series. For FVMD, the residual values show little variation with increasing *K*, with only a slight decrease observed at the peaks. In the case of DFVMD4-4, it exhibits a certain level of reduction in numerical values compared to FVMD, but it still retains strong oscillatory characteristics. On the other hand, DFVMD8-4 shows a significant decrease in numerical values, and the oscillatory features are noticeably weakened when compared to FVMD. DFVMD can reduce the impact of inadequate VMD decomposition to a certain extent. These observations indicate that different values of *K* have a considerable impact on DFVMD, making the selection of an appropriate *K* value crucial for achieving meaningful results with DFVMD.

Due to the absence of a direct relationship between the residual terms and the prediction accuracy of the model, and the fact that smaller residual terms may indicate the presence of more noise in the data, this study aims to conduct an in-depth investigation of the influence of the *K* value in the DFVMD-LSTM model on the predictive results. To achieve this, the DFVMD-LSTM model is compared and evaluated with the FVMD-LSTM model using different values of *K* for predicting the TSI time series *C1*. The

Table 1 Prediction accuracy of the DFVMD-LSTM model under different *K* values.

Model	R_1	*I* (%)	*RMSE* ($W.m^{-2}$)	*I* (%)	*MAE* ($W.m^{-2}$)	*I* (%)
LSTM	0.9644	/	0.0581	/	0.0496	/
FVMD4-LSTM	0.9709	0.67	0.0401	30.98	0.0288	41.96
FVMD5-LSTM	0.9652	0.08	0.0417	28.20	0.0293	40.92
FVMD6-LSTM	0.9678	0.35	0.0402	30.89	0.0287	42.17
FVMD7-LSTM	0.9680	0.37	0.0400	31.23	0.0284	42.78
FVMD8-LSTM	0.9664	0.21	0.0442	23.96	0.0338	31.93
DFVMD4-4-LSTM	0.9698	0.56	0.0383	34.05	0.0273	44.87
DFVMD4-5-LSTM	0.9720	0.79	0.0374	35.63	0.0268	45.92
DFVMD4-6-LSTM	0.9682	0.39	0.0410	29.37	0.0304	38.73
DFVMD4-7-LSTM	0.9684	0.41	0.0401	30.94	0.0291	41.39
DFVMD4-8-LSTM	0.9686	0.44	0.0400	31.18	0.0291	41.26
DFVMD5-4-LSTM	**0.9722**	**0.81**	**0.0372**	**35.93**	**0.0267**	**46.25**
DFVMD6-4-LSTM	0.9696	0.54	0.0392	32.48	0.0283	42.86
DFVMD7-4-LSTM	0.9694	0.52	0.0400	31.10	0.0292	41.09
DFVMD8-4-LSTM	0.9669	0.26	0.0419	27.94	0.0308	37.86

predictive accuracy of the models under different *K* values is then compared and assessed.

In this study, the *C1* dataset is divided into a training set from 1981 to 2013, a validation set from 2013 to 2017, and a test set from 2017 to 2021 for experimentation. The experiments involve grid searching for the *K* value, with the FVMD-LSTM model using *K* values ranging from 4 to 8 for VMD. For the DFVMD-LSTM model, two sets of experiments were conducted. In the first set, the *K* value for the first VMD decomposition was fixed at 4, and the *K* value for the second VMD decomposition was varied from 4 to 8. In the second set, the *K* value for the second VMD decomposition was fixed at 4, while the *K* value for the first VMD decomposition was varied from 4 to 8. The experimental results, including accuracy metrics and improvement rates, are presented in Table 1.

In Table 1, FVMD4-LSTM represents the FVMD-LSTM model with TSI data decomposed using VMD with *K*=4 for prediction. DFVMD4-n-LSTM (n=4~8) represents the DFVMD-LSTM model with the first VMD of TSI data using *K*=4 and the second VMD of the residual r using *K*=n for prediction. DFVMDm-4-LSTM (m=4~8) represents the DFVMD-LSTM model with the first VMD of TSI data using *K*=m and the second VMD of the residual r using *K*=4 for prediction. "*I*" represents the prediction accuracy improvement yielded by the FVMD-LSTM and DFVMD-LSTM models over the LSTM model. "R_1" represents the correlation between the model's predicted results and the actual data.

From Table 1, it can be observed that the LSTM model exhibits a high correlation of 0.9644 with the original data, indicating a strong alignment between the predicted fluctuation trends and the actual data. However, the LSTM model still shows some deviation from the original data, with an *RMSE* value of 0.0581 $W.m^{-2}$ and a *MAE* value of 0.0496 $W.m^{-2}$. Compared to the LSTM model, both the FVMD-LSTM and DFVMD-LSTM models yield significant prediction accuracy improvements. The FVMD-LSTM model achieves average *RMSE* reductions ranging from 23.96 % to 31.23 %, with an average decrease of 29.05 %; average *MAE* reductions ranging from 31.93 % to 42.78 %, with an average decrease of 39.95 %; and average R1 increases ranging from 0.08 % to 0.67 %, with an average increase of 0.34 %. The DFVMD-LSTM model achieves average *RMSE* reductions ranging from 27.94 % to 35.93 %, with an average decrease of 32.07 %; average *MAE* reductions ranging from 37.86 % to 46.25 %, with an average decrease of 42.25 %; and average R1 increases ranging from 0.26 % to 0.81 %, with an average increase of 0.52 %. From the results, it can be concluded that the DFVMD-LSTM model demonstrates a certain level of improvement in overall predictive accuracy compared to the FVMD-LSTM model. When the *K* value for the first VMD decomposition in the DFVMD-LSTM model is the same as the *K* value in the FVMD-LSTM model, the predictive accuracy of the DFVMD-LSTM model is mostly slightly higher than that of the FVMD-LSTM model. This observation indicates that the DFVMD-LSTM model can further enhance the predictive accuracy of the FVMD-LSTM model by extracting oscillatory characteristics from the residual terms. However, since the DFVMD-LSTM model involves two rounds of VMD decomposition and is influenced by parameters such as *K*, achieving appropriate *K* values and other relevant parameters to effectively extract oscillatory features from the residual terms and minimize noise effects requires further in-depth exploration in future research.

In conclusion, the DFVMD-LSTM model shows a certain level of improvement in predictive accuracy compared to the FVMD-LSTM model. However, due to the significant influence of different *K* values on the predictive accuracy of the DFVMD-LSTM model, this paper selects the optimal *K* values for subsequent experiments. Specifically, the DFVMD-LSTM model

Table 2 Division of each dataset.

Dataset	TSI data	Training Set		Validation Set		Test Set	
		Time span	Length (years)	Time span	Length (years)	Time span	Length (years)
Dataset 1	*C1*	1981-2015	33.00	2015-2016	1.00	2016-2021	5.75
Dataset 2	*C1*	1992-2015	23.00	2015-2016	1.00	2016-2021	5.75
Dataset 3	*C1*	2003-2015	11.50	2015-2016	1.00	2016-2021	5.75
Dataset 4	*C1*	2009-2015	5.75	2015-2016	1.00	2016-2021	5.75
Dataset 5	*C2*	1978-1991	13.00	1991-1992	1.00	1992-1998	5.75
Dataset 6	*C2*	1978-1991	13.00	1991-1992	1.00	1992-2004	11.50
Dataset 7	*C2*	1978-1991	13.00	1991-1992	1.00	1992-2015	23.00

uses K=5 for the first VMD decomposition and K=4 for the second VMD decomposition.

4. EXPERIMENTAL RESULTS AND ANALYSIS

4.1. ANALYSIS OF MULTISOURCE DATA PREDICTION RESULTS

To analyze the prediction accuracy of the DFVMD-LSTM model under different feature data, considering that the utilized TSI time series have periods of 11.5 years (Montillet et al., 2022), this experiment divides the *C1* and *C2* datasets in 7 different ways as follows, and the division results are shown in Table 2.

In Table 2, the values under the term "Training Set" denote the duration and length of the training dataset. The values under "Validation Set" represent the duration and length of the validation dataset, while those under "Test Set" indicate the duration and length of the testing dataset. Each of these partitioned datasets is used as features in the ANN, LSTM, and DFVMD-LSTM models for experimentation purposes. The prediction results of each model are depicted in Appendix A (Figure A1 to Figure A7). In order to highlight the difference in prediction results, this paper adds prediction error to analyze. Specifically, datasets 1 to 4 are employed to investigate the prediction accuracy of each model for the same time span under different training durations and to identify the shortest training period. On the other hand, datasets 5 to 7 are utilized to examine the prediction accuracy of each model for different prediction time spans while maintaining the same training duration.

From Figure A1 to Figure A7, it can be observed that the LSTM model outperforms the traditional ANN model in predicting the long-term TSI time series. However, there are still deviations in the predicted oscillation amplitudes compared to the original data. On the other hand, the DFVMD-LSTM model not only maintains consistency in predicting trends with the initial data but also effectively reduces the prediction bias of the LSTM model, leading to higher predictive accuracy.

From Figure A1 to Figure A4, it can be observed that with decreasing training time, the deviations between the predictions of the ANN, LSTM, and DFVMD-LSTM models and the actual data gradually increase. For Dataset 1, both the LSTM and DFVMD-LSTM models exhibit a high degree of fit with the original TSI time series when using the full training data. However, when reducing the training set to 23 years (two TSI cycles), the goodness of fit decreases to some extent. Further reducing the training data to 5.75 years (half a TSI cycle) results in a significant deviation of the predictions from the original TSI time series. This indicates that reducing the training time decreases the prediction accuracy, and the training time should not be less than one TSI cycle. This finding confirms the importance of training time in the reconstruction of TSI time series. From the 95 % prediction intervals of each model, it can be inferred that the predictive results of the DFVMD-LSTM model are generally lower than those of the LSTM model, indicating a better fit with the true TSI data.

Considering the minimum training time and length of dataset *C2*, this experiment utilized the first 13 years of data from dataset *C2* as the training set and 1 year as the validation set. It predicted the time series for half a TSI cycle, one TSI cycle, and two TSI cycles, respectively. The predictive results are shown in Figures A5 to A7.

From Figure A5 to Figure A7, it can be concluded that the DFVMD-LSTM model provides superior predictive results when an adequate training dataset is available, showing a reasonably good alignment with the trend of the original TSI data. In contrast, the LSTM model, as the prediction horizon increases, exhibits larger deviations from the original TSI data, with even more pronounced discrepancies at extreme points. Analyzing the prediction errors, the LSTM model's errors are significantly influenced by the fluctuation trends, closely resembling the fluctuation trends of the original data. In contrast, the DFVMD-LSTM model shows smoother prediction errors, with most time points having lower errors compared to the LSTM model. The 95 % prediction intervals obtained from the model predictions indicate that the DFVMD-LSTM model's intervals are wider than those of the LSTM model. This suggests that the

Table 3 Accuracy evaluation indices of each model under different characteristics.

Model	Dataset	R_2	*RMSE* ($W.m^{-2}$)	*I* (%)	*MAE* ($W.m^{-2}$)	*I* (%)
ANN	Dataset 1	0.98	0.1314	/	0.1213	/
LSTM			0.0835	/	0.0712	/
VMD-LSTM			0.0406	51.31	0.0264	62.94
DFVMD-LSTM			0.0688	17.62	0.0533	25.15
ANN	Dataset 2	0.98	0.1339	/	0.1209	/
LSTM			0.1003	/	0.0890	/
DFVMD-LSTM			0.0889	11.30	0.0766	13.87
ANN	Dataset 3	0.97	0.2534	/	0.2472	/
LSTM			0.1114	/	0.1009	/
DFVMD-LSTM			0.0825	25.90	0.0696	31.01
ANN	Dataset 4	0.98	0.2324	/	0.2231	/
LSTM			0.1256	/	0.1080	/
DFVMD-LSTM			0.1297	-3.26	0.1052	2.60
ANN	Dataset 5	0.96	0.1167	/	0.0980	/
LSTM			0.1099	/	0.0908	/
DFVMD-LSTM			0.0967	12.05	0.0689	24.13
ANN	Dataset 6	0.97	0.1559	/	0.1178	/
LSTM			0.1347	/	0.1054	/
DFVMD-LSTM			0.1124	16.53	0.0817	22.45
ANN	Dataset 7	0.98	0.1371	/	0.1057	/
LSTM			0.1256	/	0.1009	/
DFVMD-LSTM			0.0962	23.41	0.0693	31.30

DFVMD-LSTM model captures larger fluctuations characteristics and fits the original data better. As the prediction horizon increases, the DFVMD-LSTM model's prediction errors do not show significant increases, indicating its stable predictive performance under sufficient training years. To achieve even higher predictive accuracy, increasing the model's training years and conducting thorough discussions on model parameters are necessary. Considering the TSI data's distinct and regular fluctuation frequencies and amplitudes, the data still demonstrates high predictive accuracy in long-term forecasting.

In conclusion, the DFVMD-LSTM model not only resolves the issue of information leakage in the VMD-LSTM model but also enhances the predictive accuracy of the LSTM model, reducing the prediction errors from the original data. The DFVMD-LSTM model exhibits favorable predictive performance for various TSI time series under different training periods and prediction ranges, demonstrating its strong adaptability and robustness. Thus, this method can be applied to other domains requiring high-precision time series forecasting.

4.2. MULTISOURCE DATA PREDICTION ACCURACY ANALYSIS

To further analyze the degree of model accuracy improvement, this paper uses the *RMSE*, the *MAE* and R_2 to evaluate the prediction results produced by three different models under different datasets, and the results are shown in Table 3.

In Table 3, R_2 represents the correlation between the DFVMD-processed data and the original data, while "*I*" represents the prediction accuracy improvement achieved over the LSTM model by each hybrid model. The table shows that the average correlation between the DFVMD-processed data and the original data is as high as 0.97, indicating that the proposed method resolves the issue of poor correlation between the data obtained after VMD and the original data. This suggests its applicability in the field of high-correlation multi-model prediction. Due to article length limitations, a comprehensive explanation of this topic will be covered in future research. From the prediction accuracies produced by each model under different datasets, it can be observed that the LSTM model, as a superior predictive model in long time series analysis cases, outperforms the traditional deep learning algorithm (the ANN model) in terms of prediction accuracy. Under Dataset 1, the VMD-LSTM model yields an *RMSE* reduction of 51.31 % and a *MAE* reduction of 62.94 % compared to the LSTM model. This indicates that the information leakage caused by VMD leads to an overestimation of the prediction accuracy in the VMD-LSTM model, making it unsuitable for practical applications. Hence, this study does not further investigate its performance on other datasets.

From the prediction results produced by each model on Dataset 1 to Dataset 4, it can be observed that as the training horizon decreases, the prediction accuracies of all models decrease to varying degrees,

reaching their lowest values when the training horizon is half a TSI cycle. Compared to the LSTM model, the DFVMD-LSTM model demonstrates improved prediction accuracy at different training horizons, with an average *RMSE* reduction of 12.89 % and an average *MAE* reduction of 18.16 %. This indicates that the DFVMD-LSTM model not only solves the issue of information leakage in the VMD-LSTM model but also achieves significant prediction accuracy improvements over the LSTM model. However, when the training set is only half a TSI cycle, the *RMSE* of the DFVMD-LSTM model increases by 3.26 %. This suggests that both the DFVMD-LSTM model and the LSTM model exhibit greater uncertainty in their predictions when the training set is insufficient. Therefore, when predicting TSI time series, the training horizon should be greater than one TSI cycle to ensure an adequately trained model.

From the prediction results produced by each model on Dataset 5 to Dataset 7, it can be observed that as the prediction horizon increases, the prediction accuracy of the LSTM model continuously decreases, while the DFVMD-LSTM model maintains more stability, indicating its stronger robustness. Compared to the LSTM model, the DFVMD-LSTM model exhibits an average *RMSE* reduction of 17.33 % and an average *MAE* reduction of 25.96 % across different prediction horizons. When predicting two TSI cycles, the DFVMD-LSTM model demonstrates the highest prediction accuracy improvement, with an *RMSE* reduction of 23.41% and a *MAE* reduction of 31.30 % compared to the LSTM model. This highlights that the DFVMD-LSTM model achieves higher prediction accuracy across different prediction horizons.

In conclusion, the DFVMD-LSTM model demonstrates superior predictive accuracy to that of the LSTM model in experiments involving different training and prediction durations for two types of TSI time series. The average *RMSE* and *MAE* reductions of 14.79 % and 21.50 %, respectively, in its predictive results validate the strong applicability and robustness of the DFVMD-LSTM model for time series forecasting. Furthermore, the data processed through DFVMD and the predicted results exhibit remarkably high correlations with the original data. Therefore, the DFVMD-LSTM model possesses great potential for applications in high-precision time series prediction domains.

5. CONCLUSION

In this paper, a dual-fusion variational mode decomposition-long short-term memory neural network (DFVMD-LSTM) model is proposed for TSI time series prediction, addressing the issue of information leakage and the low correlation between the modal components obtained through VMD and the original data. This model is based on changing the VMD method and processing the decomposed data to avoid the information leakage caused by simultaneous decomposition of the testing and training sets. The dual-fusion method used in DFVMD retains more effective information from the original signals, thereby improving the prediction accuracy of the model. The proposed DFVMD-LSTM model is tested and validated on multiple TSI datasets, and the experimental results demonstrate the following.

1. Compared to the single LSTM prediction model, the DFVMD-LSTM model exhibits *RMSE* reductions ranging from 27.94 % to 35.93 % and *MAE* reductions ranging from 37.86 % to 46.25 % for different values of *K*. In comparison with the FVMD-LSTM model, the DFVMD-LSTM model consistently demonstrates higher predictive accuracies and correlations for different *K* values. This indicates that the predictive accuracy of the DFVMD-LSTM model varies significantly with different *K* values, requiring the use of the grid search methodology to determine the optimal model parameter, *K*.
2. The DFVMD-LSTM model resolves the issue of information leakage in the VMD-LSTM model, and the *RMSE* and *MAE* of the model are reduced by averages of 12.89 % and 18.16 %, respectively, when the same prediction period is used for different training periods. When different prediction periods are used for the same training period, the *RMSE* and *MAE* are reduced by averages of 17.33 % and 25.9 6%, respectively. These results indicate that the DFVMD-LSTM model has strong adaptability and robustness to different TSI datasets. On all datasets, the *RMSE* and *MAE* of the DFVMD-LSTM model are reduced by averages of 14.79 % and 21.50 %, respectively, confirming that the DFVMD-LSTM model has high prediction accuracy and can be applied to the field of high-precision TSI time series prediction.
3. After applying DFVMD to the data, the average correlation coefficient R_2 between the processed data and the original data reaches a maximum value of 0.97. This indicates that in the context of high-correlation multifeatured prediction, the data processed through DFVMD can be used as substitutes for the original data to improve the accuracy of multi-feature prediction.
4. In addition, to further enhance the algorithm's predictive capabilities (especially for Long-term time series prediction), the training dataset could be expanded through methods such as simulating TSI observations or taking advantage of forthcoming longer datasets. Future work is essential to extend the forecast across multiple solar cycles, a crucial requirement for accurately modeling the solar influence on Earth's climate, as outlined in Herrera et al. (2015), Egorova et al. (2018) and Schmutz (2021).

FUNDING

This work was sponsored by National Natural Science Foundation of China (42104023, 42061077, 42064001), Major Discipline Academic and Technical Leaders Training Program of Jiangxi Province (20225BCJ23014), Jiangxi University of Science and Technology Postgraduate Education Teaching Reform Research Project (YJG2022006), Hebei Water Conservancy Research Plan (2022-28).

DATA AVAILABILITY STATEMENT

The TSI data are available at https://www.pmodwrc.ch/en/research-development/solar-physics/tsi-composite/ and http://www.issibern.ch/teams/solarirradiance

ACKNOWLEDGMENTS

We would like to express our gratitude to J.P. Montillet for providing the relevant TSI time series data and for their guidance and assistance throughout the writing of this paper.

REFERENCES

Adler, J. and Parmryd, I.: 2010, Quantifying colocalization by correlation: the Pearson correlation coefficient is superior to the Mander's overlap coefficient. Cytom. Part A, 77, 8, 733–742. DOI: 10.1002/cyto.a.20896

Almorox, J., Voyant, C., Bailek, N., Kuriqi, A. and Arnaldo, J.A.: 2021, Total solar irradiance's effect on the performance of empirical models for estimating global solar radiation: An empirical-based review. Energy, 236, 121486. DOI: 10.1016/j.energy.2021.121486

Ball, W.T., Unruh, Y.C., Krivova, N.A., Solanki, S., Wenzler, T., Mortlock, D.J. and Jaffe, A.H.: 2012, Reconstruction of total solar irradiance 1974–2009. Astron. Astrophys., 541, A27. DOI: 10.1051/0004-6361/201118702

Beer, J., Vonmoos, M. and Muscheler, R.: 2007, Solar variability over the past several millennia. In: Calisesi, Y., Bonnet, R.M., Gray, L., Langen, J., Lockwood, M. (eds.), Solar variability and planetary climates. Space Sciences Series of ISSI, vol 23., Springer, New York, NY., 67–79. DOI: 10.1007/978-0-387-48341-2_6

Chai, T. and Draxler, R.R.: 2014, Root mean square error (RMSE) or mean absolute error (MAE)–Arguments against avoiding RMSE in the literature. Geosci. Model Dev., 7, 3, 1247–1250. DOI: 10.5194/gmd-7-1247-2014

Chen, H., Lu, T., Huang, J., He, X., Yu, K., Sun, X., ... and Huang, Z.: 2023, An improved VMD-LSTM model for time-varying GNSS time series prediction with temporally correlated noise. Remote Sens., 15, 14, 3694. DOl: 10.3390/rs15143694

Dragomiretskiy, K. and Zosso, D.: 2013, Variational mode decomposition. IEEE Trans. Signal Process., 62, 3, 531–544. DOI: 10.1109/TSP.2013.2288675

Dudok de Wit, T., Kopp, G., Fröhlich, C. and Schöll, M.: 2017, Methodology to create a new total solar irradiance record: Making a composite out of multiple data records. Geophys. Res. Lett., 44, 3, 1196–1203. DOI: 10.1002/2016gl071866

Egorova, T., Schmutz, W., Rozanov, E., Shapiro, A.I., Usoskin, I., Beer, J., ... and Peter, T.: 2018, Revised historical solar irradiance forcing. Astron. Astrophys., 615, A85. DOI: 10.1051/0004-6361/201731199

Gao, B., Huang, X., Shi, J., Tai, Y. and Zhang, J.: 2020, Hourly forecasting of solar irradiance based on CEEMDAN and multi-strategy CNN-LSTM neural networks. Renew. Energy, 162, 1665–1683. DOI: 10.1016/j.renene.2020.09.141

Gueymard, C.A.: 2018, A reevaluation of the solar constant based on a 42-year total solar irradiance time series and a reconciliation of spaceborne observations. Sol. Energy, 168, 2–9. DOI: 10.1016/j.solener.2018.04.001

Han, L., Zhang, R., Wang, X., Bao, A. and Jing, H.: 2019, Multi-step wind power forecast based on VMD-LSTM. IET Renew. Power Gener., 13, 10, 1690–1700. DOI: 10.1049/iet-rpg.2018.5781

Herrera, V.V., Mendoza, B. and Herrera, G.V.: 2015, Reconstruction and prediction of the total solar irradiance: From the Medieval Warm Period to the 21st century. New Astron., 34, 221–233. DOI: 10.1016/j.newast.2014.07.009

Hochreiter, S. and Schmidhuber, J.: 1997, Long short-term memory. Neural Comput., 9, 8, 1735–1780. DOI: 10.1007/978-3-642-24797-2_4

Huang, S.: 2008, Surface temperatures at the nearside of the Moon as a record of the radiation budget of Earth's climate system. Adv. Space Res., 41, 11, 1853–1860. DOI: 10.1016/j.asr.2007.04.093

Jin, Y., Guo, H., Wang, J. and Song, A.: 2020. A hybrid system based on LSTM for short-term power load forecasting. Energies, 13, 23, 6241. DOI: 10.3390/en13236241

Liao, X., Liu, Z. and Deng, W.: 2021, Short-term wind speed multistep combined forecasting model based on two-stage decomposition and LSTM. Wind Energy, 24, 9, 991–1012. DOI: 10.1002/we.2613

Li, C., Wu, Y., Lin, H., Li, J., Zhang, F. and Yang, Y.: 2022. ECG denoising method based on an improved VMD algorithm. IEEE Sens. J., 22, 23, 22725–22733. DOI : 10.1109/jsen.2022.3214239

Li, Yu., Li, Ya., Chen, X. and Yu, J.: 2017, Denoising and feature extraction algorithms using NPE combined with VMD and their applications in ship-radiated noise. Symmetry, 9, 11, 256. DOI: 10.3390/sym9110256

Li, Z., Lu, T., He, X., Montillet, J.P. and Tao, R.: 2023, An improved cyclic multi model-eXtreme gradient boosting (CMM-XGBoost) forecasting algorithm on the GNSS vertical time series. Adv. Space Res., 71, 1, 912–935. DOI: 10.1016/j.asr.2022.08.038

Lv, L., Wu, Z., Zhang, J., Zhang, L., Tan, Z. and Tian, Z.: 2021, A VMD and LSTM based hybrid model of load forecasting for power grid security. IEEE Trans. Industrial Inform., 18, 9, 6474–6482. DOI: 10.1109/tii.2021.3130237

Majumder, I., Dash, P.K. and Bisoi, R.: 2018, Variational mode decomposition based low rank robust kernel extreme learning machine for solar irradiation forecasting. Energy Convers. Manag., 171, 787–806. DOI: 10.1016/j.enconman.2018.06.021

Mishchenko, M.I., Cairns, B., Kopp, G., Schueler, C.F., Fafaul, B.A., Hansen, J.E. ... and Travis, L.D.: 2007, Accurate monitoring of terrestrial aerosols and total solar irradiance: introducing the Glory mission. Bull. Am. Meteorol. Soc., 88, 5, 677–692. DOI: 10.1109/igarss.2010.5652996

Montillet, J.-P., Finsterle, W., Kermarrec, G., Sikonja, R., Haberreiter, M., Schmutz, W. and T. Dudok de Wit, T.: 2022, Data fusion of total solar irradiance composite time series using 41 years of satellite measurements. J. Geophys. Res., Atmospheres, 127, 13, e2021JD036146. DOI: 10.1029/2021JD036146

Pazikadin, A.R., Rifai, D., Ali, K., Malik, M.Z., Abdalla, A.N. and Faraj, M.A.: 2020, Solar irradiance measurement instrumentation and power solar generation forecasting based on Artificial Neural Networks (ANN): A review of five years research trend. Sci. Total Environ., 715, 136848. DOI: 10.1016/j.scitotenv.2020.136848

Qian, Z., Pei, Y., Zareipour, H. and Chen, N.: 2019, A review and discussion of decomposition-based hybrid models for wind energy forecasting applications. Appl. Energy, 235, 939–953. DOI: 10.1016/j.apenergy.2018.10.080

Sagheer, A. and Kotb, M.: 2019, Time series forecasting of petroleum production using deep LSTM recurrent networks. Neurocomputing, 323, 203–213. DOI: 10.1016/j.neucom.2018.09.082

Schmutz, W.K.: 2021, Changes in the total solar irradiance and climatic effects. J. Space Weather Space Clim., 11, 40. DOI: 10.1051/swsc/2021016

Shatat, M., Worall, M. and Riffat, S.: 2013, Opportunities for solar water desalination worldwide. Sustain. Cities Soc., 9, 67–80. DOI: 10.1016/j.scs.2013.03.004

Sherstinsky, A: 2020, Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. Phys. D: Nonlinear Phenom., 404, 132306. DOI: 10.1016/j.physd.2019.132306

Shi, X., Lei, X., Huang, Q., Huang, S., Ren, K. and Hu, Y.: 2018, Hourly day-ahead wind power prediction using the hybrid model of variational model decomposition and long short-term memory. Energies, 11, 11, 3227. DOI: 10.3390/en11113227

Sun, Z., Zhao, S. and Zhang, J.: 2019, Short-term wind power forecasting on multiple scales using VMD decomposition, K-means clustering and LSTM principal computing. IEEE Access, 7, 166917–166929. DOI: 10.1109/access.2019.2942040

ur Rehman, N. and Aftab, H.: 2019, Multivariate variational mode decomposition. IEEE Trans. Signal Process., 67, 23, 6039–6052. DOI: 10.1109/tsp.2019.2951223

Wan, C., Zhao, J., Song, Y., Xu, Z., Lin, J. and Hu, Z.: 2015, Photovoltaic and solar power forecasting for smart grid energy management. CSEE J. Power Energy Syst., 1, 4, 38–46. DOI: 10.17775/cseejpes.2015.00046

Wang, Y., Markert, R., Xiang, J. and Zheng, W.: 2015, Research on variational mode decomposition and its application in detecting rub-impact fault of the rotor system. Mech. Syst. Signal Process., 60, 243–251. DOI: 10.1016/j.ymssp.2015.02.020

Wang, Z., He, X., Shen, H., Fan, S. and Zeng, Y.: 2022. Multi-source information fusion to identify water supply pipe leakage based on SVM and VMD. Inform. Process. Manag., 59, 2, 102819. DOI: 10.1016/j.ipm.2021.102819

Wang, Y. and Wu, L.: 2016, On practical challenges of decomposition-based hybrid forecasting algorithms for wind speed and solar irradiation. Energy, 112, 208–220. DOI: 10.1016/j.energy.2016.06.075

Willmott, C.J. and Matsuura, K.: 2005, Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Clim. Res., 30, 1, 79–82. DOI: 10.3354/cr030079

Yadav, A., Jha, C.K. and Sharan, A.: 2020, Optimizing LSTM for time series prediction in Indian stock market. Procedia Comput. Sci., 167, 2091–2100. DOI : 10.1016/j.procs.2020.03.257

Yu, Y., Si, X., Hu, C. and Zhang, J.: 2019, A review of recurrent neural networks: LSTM cells and network architectures. Neural Comput., 31, 7, 1235–1270. DOI: 10.1162/neco_a_01199

Zang, H., Liu, L., Sun, L., Cheng, L., Wei, Z. and Sun, G.: 2020, Short-term global horizontal irradiance forecasting based on a hybrid CNN-LSTM model with spatiotemporal correlations. Renew. Energy, 160, 26–41. DOI: 10.1016/j.renene.2020.05.150

Zhang, M., Jiang, Z. and Feng, K.: 2017, Research on variational mode decomposition in rolling bearings fault diagnosis of the multistage centrifugal pump. Mech. Syst. Signal Process., 93, 460–493. DOI: 10.1016/j.ymssp.2017.02.013

Zhou, T., Wu, W., Peng, L., Zhang, M., Li, Z., Xiong, Y. and Bai, Y.: 2022, Evaluation of urban bus service reliability on variable time horizons using a hybrid deep learning method. Reliab. Eng. Syst. Saf., 217, 108090. DOI: 10.1016/j.ress.2021.108090

Zupan, J.: 1994, Introduction to artificial neural network (ANN) methods: what they are and how to use them. Acta Chim. Slov., 41, 327–327. DOI: 10.1385/1-59259-192-2:401

Zupan, J.:1994. Introduction to artificial neural network (ANN) methods: what they are and how to use them. Acta Chimica Slovenica, 41, 327–327. DOI: 10.1385/1-59259-192-2:401

APPENDIX A

Considering the consistent legends in Figures A1 to 6, a uniform description is given as follows:
The black curve represents the original TSI time series, denoted as "TRUE."
The red curve represents the predictive results and prediction errors (the difference between the ground truth and the predicted values) of the DFVMD-LSTM model for each dataset.
The green curve represents the predictive results and prediction errors of the LSTM model for each dataset.
The blue curve represents the predictive results and prediction errors of the ANN model for each dataset.
The light gray area indicates the 95 % prediction interval of the LSTM model, representing the range within which 95 % of the data points are expected to fall after fitting.
The light blue area indicates the 95 % prediction interval of the DFVMD-LSTM model.

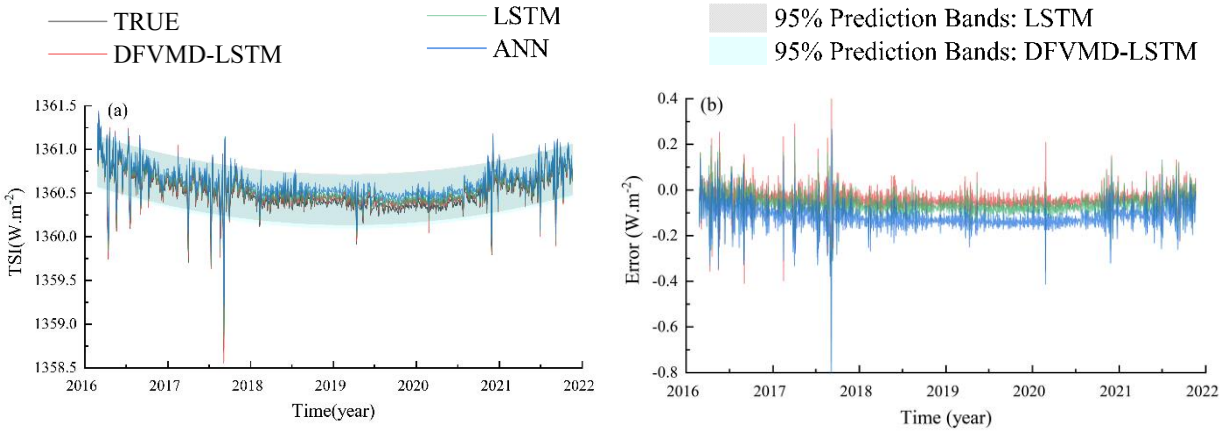

Fig. A1 Prediction results and errors of each model in dataset 1 (Figure (a) is prediction result, (b) is error check).

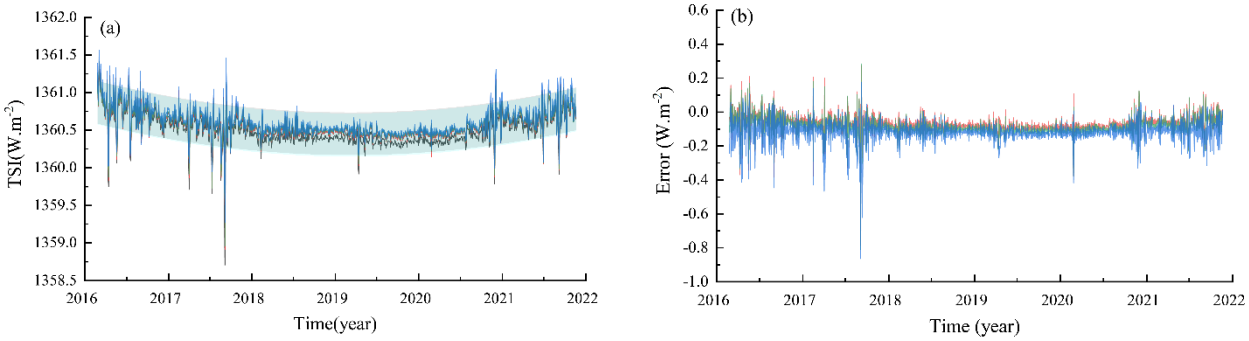

Fig. A2 Prediction results and errors of each model in dataset 2 (Figure (a) is prediction result, (b) is error check).

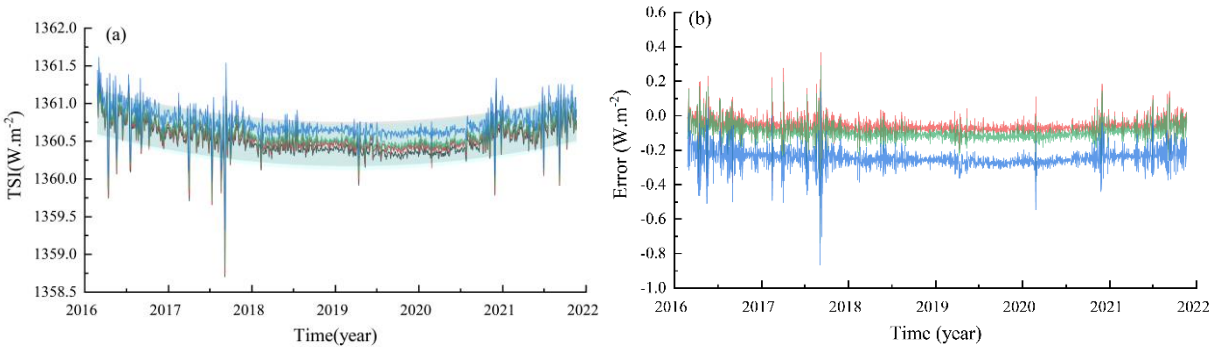

Fig. A3 Prediction results and errors of each model in dataset 3 (Figure (a) is prediction result, (b) is error check).

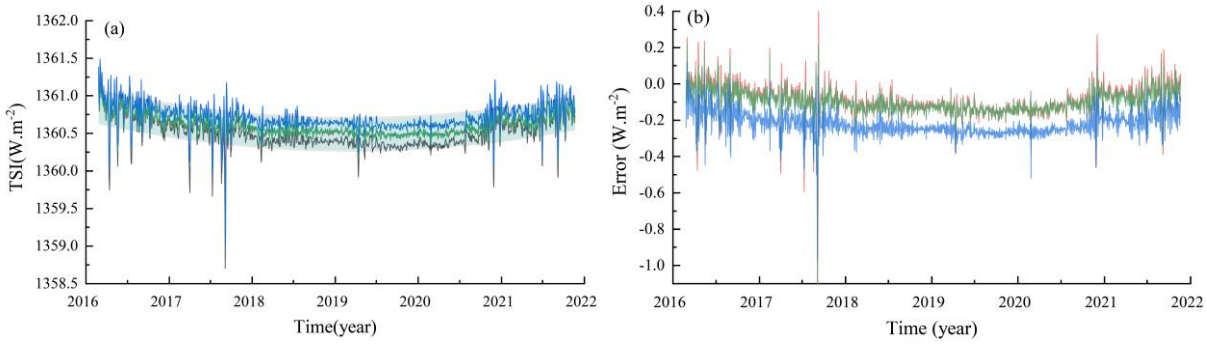

Fig. A4 Prediction results and errors of each model in dataset 4 (Figure (a) is prediction result, (b) is error check).

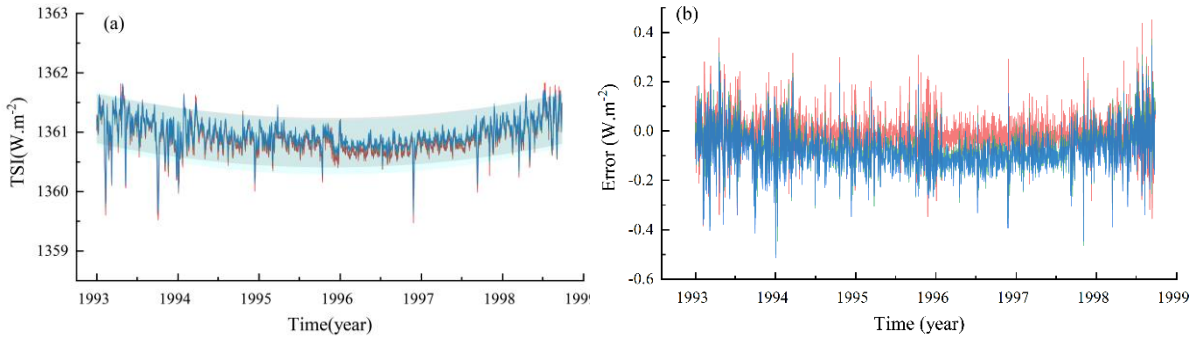

Fig. A5 Prediction results and errors of each model in dataset 5 (Figure (a) is prediction result, (b) is error check).

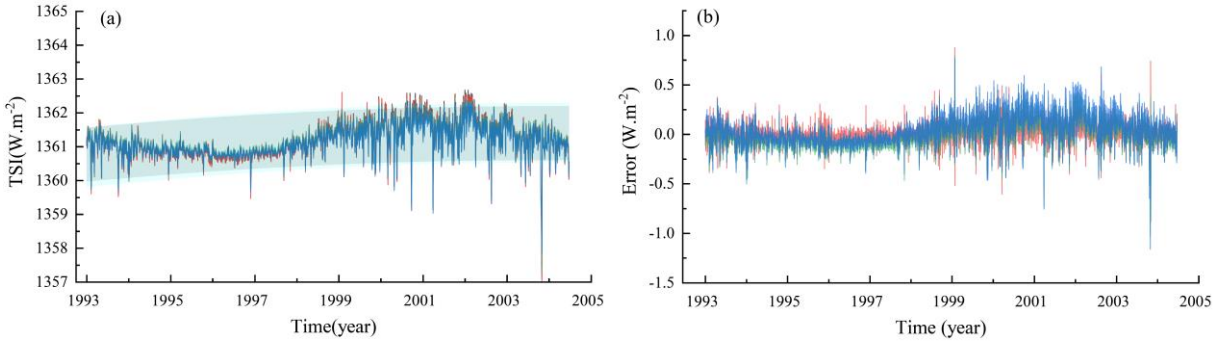

Fig. A6 Prediction results and errors of each model in dataset 6 (Figure (a) is prediction result, (b) is error check).

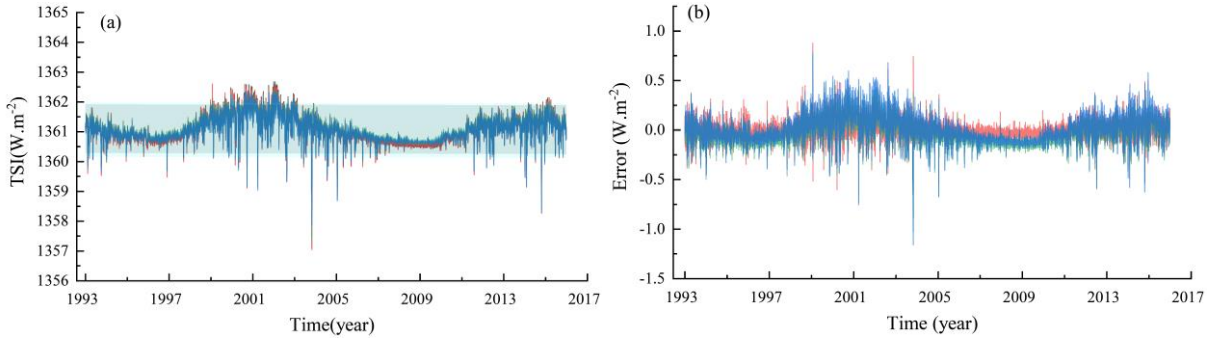

Fig. A7 Prediction results and errors of each model in dataset 7 (Figure (a) is prediction result, (b) is error check).