Journal of Computers Vol. 33 No. 2, April 2022, pp. 25- 36
doi: 10.53106/199115992022043302003

Short-term PV Power Prediction based on VMD-SE-TCAN Model

Yun Wu, Dan-Nan Zhang, Jie-Ming Yang, Zhen-Hong Liu*

School of Computer Science, Northeast Electric Power University, Jilin 132012, China
{838558160, 965435538, 670172713, 33249648}@qq.com

Received 25 May 2021; Revised 20 September 2021; Accepted 20 October 2021

Abstract. Focusing on the accuracy of short-term photovoltaic power prediction, this paper proposes a VMD-SE-TCAN short-term photovoltaic prediction model. First, the Sample Entropy (SE) is used to fuse the different frequency components generated by the Variational Modal Decomposition (VMD) to reduce the complexity of prediction model and effectively alleviate the under-decomposition or over-decomposition of traditional decomposition algorithm; then the TCAN model based on Temporal Convolution Neural network (TCN) and Attention Mechanism is used to predict each component, the TCN is used to capture temporal dependencies in prediction, and the Attention Mechanism improves the impact of key weather features on different components; finally, the prediction results of each component are superimposed to output photovoltaic prediction power. Through experiments and comparison verification, the combined prediction method effectively improves the accuracy of photovoltaic power prediction, and can better guarantee the reliable operation of the power system.

Keywords: VMD, TCN, SE, attention mechanism, photovoltaic power prediction

1 Introduction

Affected by the constraints of resource reserves and fossil energy environmental issues, the development and utilization of renewable energy has become an inevitable trend of energy transformation worldwide. As a promising form of clean energy application, photovoltaic power prediction has become an important direction of energy transformation [1-2]. With the implementation of large-scale photovoltaic grid-connected in various countries in the world, the proportion of photovoltaic grid-connected power generation in the grid is increasing day by day, and it has brought significant environmental and economic benefits [3-4]. However, due to the inherent intermittency, randomness and volatility of photovoltaic power prediction, and the greater impact of meteorological factors [5], it has brought many difficulties and risks to the safe operation of the power system when it is connected to the grid [6-7]. Therefore, improving the accuracy of photovoltaic power prediction is of great significance to the long-term development of photovoltaic power prediction.

The current research on photovoltaic power prediction forecast mainly includes physical methods and historical photovoltaic data analysis methods. The physical methods mainly include the method of predicting photovoltaic power prediction by modeling photovoltaic cells and inverters, using weather information and sun trajectories as input data [8-9]. The physical method does not require historical operating data, but requires detailed physical information of photovoltaic power plants, and the prediction accuracy is low, so it is rarely used in practice [10]. For historical data analysis methods, relevant theories and models are used to predict photovoltaics, such as SVM [11], Markov [12-13], Extreme Learning Machines (ELM) [14], neural networks [15-16] and other methods. Among them, the Long-Short Term Memory (LSTM) neural network is widely used in the complex photovoltaic time series prediction field because it can mine the spatial and temporal correlation between the output and the relevant input variables [17]. Literature [18] propose a hybrid model of Convolutional Neural Network (CNN) and LSTM for prediction, CNN was used to extract the feature vector, and the feature vector was input into LSTM in time series to obtain the final photovoltaic power output; In literature [19], a prediction model based on LSTM and Attention Mechanism is proposed, firstly, the Attention Mechanism is used to assign weight coefficients to the input sequence, and then the output of each unit of LSTM is calculated according to the forward propagation method, and the error between the real value and the predicted value is calculated by the backward propagation method, finally, the weight gradient is calculated according to the error term and the model weight is updated, and good results are achieved, but the use of LSTM for prediction has the problem of long model training time [20].

Photovoltaic power is a nonlinear time series with strong volatility and some scholars use signal decomposition method to decompose it into a series of sub-modes for prediction. Literature [21] uses empirical decom-

* Corresponding Author

position method (EMD) to decompose photovoltaic power prediction into relatively stable IMF components of different frequencies, and inputs the minimum support vector machine (LSSVM) for prediction. Literature [22] uses stationary wavelet transform (SWT), multiple LSTMs and deep neural network (DNN) are combined with an adaptive model for prediction. Among them, multiple LSTM use the photovoltaic power and other features decomposed by SWT to make multi-step predictions. The deep neural network combines the predicted value of LSTM with temperature data. The final prediction results are obtained through the combination of other factors. Literature [23] uses a combined algorithm of wavelet decomposition and BP neural network to predict the short-term power output of photovoltaic power in the same weather type as sunny, cloudy and rainy. However, EMD is easy to cause modal mixing phenomenon, and the wavelet decomposition may bring harmonics that the original physical quantity does not have, resulting in an increase in prediction errors. The literature [24-25] used VMD to decompose the photovoltaic component and achieved good prediction results. However, the number of modes after decomposition is too large, resulting in complex prediction model and low computational efficiency, which brings many difficulties for short-term photovoltaic prediction.

In order to improve the calculation efficiency and prediction accuracy of short-term photovoltaic power prediction model, this paper proposes a VMD-SE-TCAN short-term photovoltaic prediction model. The main innovations are as follows:

(1) Because of the randomness of photovoltaic power generation sequence, it is difficult to model it. The prediction after decomposition of photovoltaic power by signal decomposition method can effectively reduce the prediction error caused by the non-stationarity of photovoltaic data, but it also brings problems such as complex prediction model and low computational efficiency. Therefore, this paper proposes to use the Sample Entropy to fuse the different frequency components generated by the Variational Modal Decomposition, which reduces the data volatility and also reduces the model prediction complexity.

(2) The traditional LSTM has achieved good results in photovoltaic prediction, but the model's training time is long. TCN has good performance in predicting time series and has been applied in many important applications field. Therefore, this paper proposes to use TCN to predict each photovoltaic component after fusion.

(3) Due to the different influence of different weather characteristics on photovoltaic power fluctuation characteristics, this paper adds the feature Attention Mechanism to the TCN prediction model to assign different weights to different weather characteristics to improve the prediction accuracy of the model.

Through experimental and comparative verification, the combined prediction method effectively improves the accuracy of photovoltaic power prediction, and can better guarantee the reliable operation of the power system.

2 Related Algorithms

2.1 Variational Modal Decomposition

VMD is an intrinsic, adaptive, and non-recursive decomposition technology, which can effectively deal with non-stationary and nonlinear signals, reduce the complexity of prediction and improve the prediction accuracy [24]. The algorithm adaptively decomposes the initial complex signal into K modal sets with finite bandwidth by iterative search variational model, so as to solve the signal noise, highlight the local characteristics of the signal and avoid modal aliasing.

(1) The input signal is obtained as the corresponding one-sided spectrum by Hilbert transform, as well as the spectrum of each mode is modulated to the corresponding fundamental band for all the modal components, then the output signal of each component subject to constraints is equation (1) [25-26].

$$\begin{cases} \min\limits_{\{u_k\}\{\omega_k\}} \left\{ \sum\limits_{k=1}^{K} \| \partial_t \left[\left(\delta(t) + \dfrac{j}{\pi t} \right) \circ u_k(t) \right] e^{-j\omega_k t} \|_2^2 \right\} \\ s.t. \sum\limits_{k=1}^{K} u_k = f(t) \end{cases}. \tag{1}$$

where: $\{u_k\} = \{u_1, \cdots u_k\}$, $\{\omega_k\} = \{\omega_1, \cdots \omega_k\}$ is the set of modes and the set of center frequencies, respectively; K is the number of components; ∂_t is the partial derivative; $\delta(t)$ is the unit shock function, t is the time; $\circ$ is the convolution operation; and $f(t)$ is the input signal.

(2) To solve the above formula, the constrained problem needs to be transformed into an unconstrained problem to be solved, establishing the function as shown in equation (2).

$$L(\{u_k\},\{\omega_k\},\lambda)=\alpha\sum_{k=1}^{K}\left\|\partial_t\left[\left(\delta(t)+\frac{j}{\pi t}\right)\circ u_k(t)\right]e^{-jw_k t}\right\|_2^2+\left\|f(t)+\sum_{k=1}^{K}u_k(t)\right\|_2^2+\left\langle\lambda(t),f(t)-\sum_{k=1}^{K}u_k(t)\right\rangle. \tag{2}$$

where: α is the introduced quadratic penalty term, and λ is the Lagrange multiplier.

(3) Using the alternating direction multiplier method to solve equation (2) to obtain the components u_k and their central frequencies ω_k , the update method is shown in equation (3) and equation (4).

$$\hat{u}_k^{n+1}(\omega)=\frac{\hat{f}(\omega)-\sum_{i\neq k}\hat{u}_k^n(\omega)+\frac{\hat{\lambda}^n(\omega)}{2}}{1+2\alpha(\omega-\omega_k^{(n)})^2}. \tag{3}$$

$$\omega_k^{n+1}=\frac{\int_0^\infty \omega\,|\,\hat{u}_k^{n+1}\omega\,|^2\,d\omega}{\int_0^\infty |\,\hat{u}_k^{n+1}\omega\,|^2\,d\omega}. \tag{4}$$

where: $\hat{f}(\omega)$, $\hat{u}_k(\omega)$ and $\hat{\lambda}(\omega)$ are the fourier transforms of $f(t)$, $u_k(t)$ and $\lambda(t)$; n is the number of iterations; ω is the frequency.

2.2 Sample Entropy (SE)

In the early 20th century, scholars such as Richman proposed the concept of Sample Entropy (SE) as a measure of time series complexity. The advantage of SE is that the classification of complex systems can show the deterministic and stochastic characteristics of a finite data time series. The specific steps are as follows.

(1) N data points form $N-m+1$ a sequence of dimensional m vectors $S_m(1),\cdots,S_m(N-m+1)$, where: $S_m(i)=\{s(i),s(i+1),\cdots,s(i+m-1)\}$ $i=1,2,\cdots N-m+1$. These vectors represent m consecutive values of s starting from the i -th point.

(2) Define the absolute value of the maximum difference in the distance of two vectors $S_m(i)$, $S_m(l)$ as $d\left[S_m(i),S_m(l)\right]$, then:

$$d\left[S_m(i),S_m(l)\right]=\max_{0\le k\le m-1}\left(\left|s(i+k)-s(l+k)\right|\right). \tag{5}$$

(3) Count the number B_i $(1\le l\le N-\mathrm{m}+1,\ l\neq i)$ of l where the distance between $S_m(i)$ and $S_m(l)$ is less than r , and define:

$$B_i^m(r)=\frac{1}{N-m}B_i. \tag{6}$$

$$B^m(r)=\frac{1}{N-m+1}\sum_{i=1}^{N-m+1}B_i^m(r). \tag{7}$$

$B^m(r)$ is the probability that two sequences match m at a given threshold value r .

4) Increase the number of dimensions to $m+1$, and repeat the above steps with.

$$B^{m+1}(r)=\frac{1}{N-m}\sum_{i=1}^{N-m}B_i^{m+1}(r). \tag{8}$$

5) Then the Sample Entropy is defined as

$$SampEn(m,r,N)=-\ln[\frac{B^{m+1}(r)}{B^m(r)}]. \tag{9}$$

where: N is the total number of time series and takes a finite value; $m \in Z$; $r \in R$ denotes the given threshold value.

2.3 Temporal Convolutional Neural Network (TCN)

The traditional convolutional neural network adopts local connection and weight sharing to process the original data at a higher level and more abstract, which can effectively extract the internal characteristics of the data. Unlike the traditional convolutions, TCN is a sequence modeling structure that combines one-dimensional full convolution network, causal convolution and dilated convolution, effectively avoiding the gradient disappearance or gradient explosion problems faced by recurrent neural networks, with advantages of parallel computing, low memory consumption, and flexibility to control the length of sequence memory by changing the sensory field [17]. TCN takes advantage of the powerful characteristics of convolution, which can extract features across time steps, so it allows sampling samples with input intervals when it convolution, and its structure is shown in Fig. 1.

In the TCN structure diagram, d is the sampling interval of each layer input. When $d=2$, the output of each two time steps in the previous layer is the input of this layer. It can be seen that the higher the level is, the greater the d is, so the effective window size of TCN increases exponentially with the increase of the level. Therefore, TCN can use less network layer to get a large sensory field.

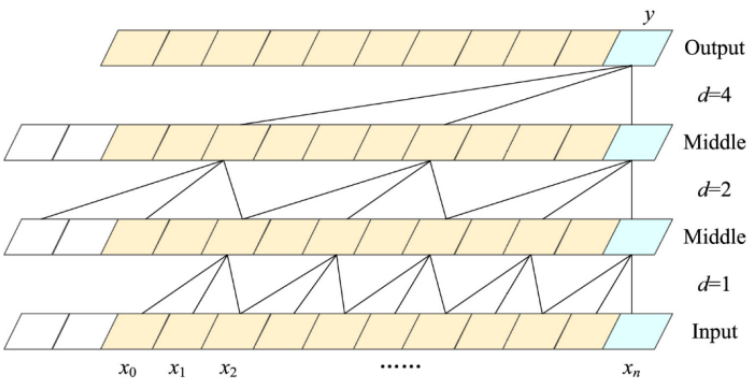

Fig. 1. Structure diagram of TCN

According to the convolution structure of TCN, the memory length of the network, that is, the fearure learning range of each neuron is related to the kernel size, expansion coefficient and model depth. Let the input sequence be x and a series of convolution kernels be $f:\{0,...,k-1\}$, then the expression of the output value of the s-th neuron after expansion convolution calculation is as follows:

$$F(s)=(x*_d f)(s)=\sum_{t=0}^{k-1}f(i)\cdot x_{s-d\cdot i}. \tag{10}$$

Where k is the size of convolution kernel; d is the expansion coefficient; '$*$' is convolution operation; $f(i)$ is the i th element in convolution kernel; $x_{s-d\cdot i}$ is the sequence element multiplied by the corresponding element in convolution kernel.

2.4 Attentional Mechanisms

In the traditional Encoder-Decoder model, while it processes the time series, the encoder encodes the input sequence into a hidden vector with a fixed length, and assigns the same weight to the hidden vector. The decoder decodes the output based on hidden vector. When the length of the input sequence increases, the same weight of the component results in no discrimination of the input sequence and the performance of the model decreases.

Attention Mechanism improves the performance of the Encoder-Decoder model. Its essence is a simulation of the form of human brain attention, which receives a large amount of information in human daily life through auditory, visual and other ways. However, humans choose a small part of useful information from these large amounts of information and focus on processing, weakening or ignoring other information, so as to work orderly in the information bombing environment. Attention Mechanism learns from this way, by changing the weight of the method, using high weight focus on important information, low weight ignores irrelevant information, and constantly adjust the weight to select important information in different cases, so it has high scalability and robustness [17, 27].

The Attention Mechanism assigns different weights to the hidden vectors at different moments of the input sequence, and merges the hidden vectors into new hidden vectors according to their importance to be input to the decoder, so the Encoder-Decode model has the ability to differentiate, and the performance is improved.

3 VMD-SE-TCAN Prediction Model Construction

To improve the prediction accuracy of short-term PV power generation, a VMD-SE-TCAN PV power prediction model is proposed in this paper. The model mainly consists of VMD decomposition layer, component fusion layer, TCAN prediction layer, and the structure is shown in Fig. 2.

Firstly, the PV power data in the initial training set is decomposed in the VMD layer to obtain the K PV decomposition components; then the Sample Entropy of each component is calculated in the component fusion layer, and the similar components are fused to obtain the K' PV components; K' TCAN prediction model combining TCN and Attention Mechanism are constructed, and the K' PV components and the weather features in the training set are used as the TCAN prediction model to obtain the final prediction results. Finally, the K' final prediction results are obtained by superimposing the predicted values.

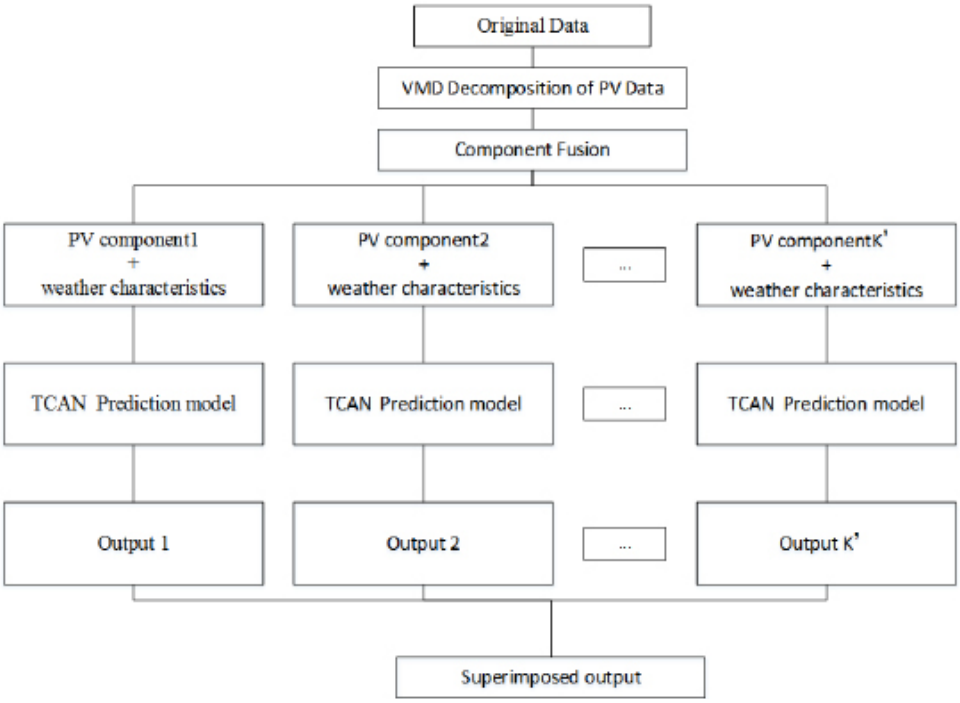

Fig. 2. Short-term PV forecasting model of VMD-SE-TCAN

3.1 VMD Decomposition and Component Fusion

Photovoltaic power is non-stationary due to the influence of meteorological factors. VMD can decompose non-stationary photovoltaic power signals and effectively reduce the complexity of photovoltaic power time series. Firstly, the photovoltaic power data in the initial training set is decomposed into VMD layer, and K photovoltaic decomposition components are obtained. Among them, the mode number K of decomposition component is the key to decomposition. When the value of K is too small, the mode is under-decomposed, and the data decomposition is not sufficient. Some important information in the original signal will be filtered, affecting the prediction accuracy. When the value of K is too large, the mode is over-decomposed and the center frequency distance of adjacent components is close, leading to modal repetition and excessive noise.

In the process of decomposition of VMD, if the decomposition is sufficient, the Sample Entropy difference between the components is obvious; if the decomposition is excessive, the Sample Entropy values of the components are close to each other. In view of the above correlation between VMD and its result component entropy, this paper analyzes whether the extracted component has over-decomposition phenomenon after using the K components obtained by the center frequency. If there is, the similar components are fused to obtain the optimal number of components K' to improve the decomposition effect of VMD.

3.2 TCN Prediction Model based on Attention Mechanism

The extended convolution structure used in the TCN model can well capture the long-term dependence and solve the problem of LSTM gradient explosion. However, when the input time series is long, the influence of different weather characteristics on the photovoltaic power generation is different. For example, the influence of irradiance on the photovoltaic power generation is greater than that of temperature. Therefore, this paper introduces the Attention Mechanism to set different weights for different neurons in the TCN hidden layer, and strengthen the influence of key characteristics on photovoltaic treatment, so as to obtain better prediction results. The TCAN structure is shown in Fig. 3.

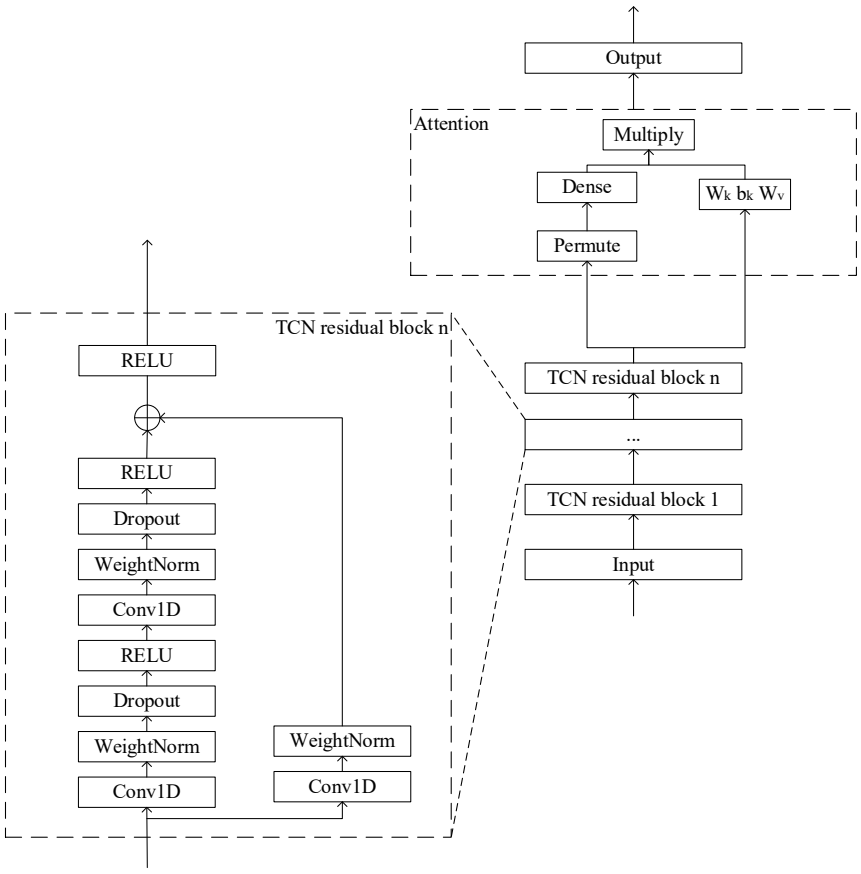

Fig. 3. Prediction model of TCAN

This paper uses the TCAN model combining TCN and Attention Mechanism to predict. The model is composed of multiple TCN residual modules and one Attention layer. The TCN residual module is composed of two parallel branches. The left branch adds the weight normalization layer (WeightNorm) and the loss layer (Dropout) after each one-dimensional convolution layer (Conv1D). The weight normalization layer can accelerate the calculation speed of the model, and the loss layer is randomly inactivated by neurons to prevent over-fitting of the model. The right branch is a one-dimensional convolution layer and weight normalization layer. The outputs of the two branches are input into the final activation function by summation. In order to make the model fully trained, this model sets seven TCN residual modules.

The Attention layer is added after the output of the last TCN residual module and attention is computed using dot product multiplication. The output vector of the TCN is first dimensionally rearranged in the Permute layer; Multiply multiplies the output of the Dense fully connected layer with the weight matrix of the TCN input in dot product and outputs it [20]. Where Attention Mechanism weights are calculated as:

$$Attention = softmax(\frac{QK^T}{\sqrt{L}})V \ . \tag{11}$$

$$K^{N\times 1\times L} = I^{N\times 1\times L} \cdot W_K{}^{N\times L\times L} \cdot b_K{}^{N\times 1\times L} \ . \tag{12}$$

$$V^{N\times 1\times L} = I^{N\times 1\times L} \cdot W_V{}^{N\times L\times L} \ . \tag{13}$$

where: Q is the attention term to be computed, K,V is the matrix of key-value pairs, L is the given scalar; I is the original time series input; W_K , b_K , W_V is the trainable matrix.

4 Example Analysis

The experimental data in this paper are obtained from the photovoltaic power prediction data of DKASC Solar Energy Center from January to December 2016 [28], containing photovoltaic power prediction and meteorological characteristics such as solar irradiance, temperature, humidity, wind speed and rainfall. Since the numerical dimensions of different variables are quite different, the effect of the model will be adversely affected during training, so the historical data are normalized.

The hardware configuration of the experimental platform is as follows: processor-Intel Xeon E5, GPU graphics card-Tesla P100; software configuration is as follows: development software-Pycharm, development language-Python.

4.1 VMD Decomposition and Component Fusion

In the process of VMD decomposition of the photovoltaic power signal, the choice of the number K of modes determines the resolution of the signal. If the value K is too low, the VMD decomposition is not sufficient, and if the value K is too high, the signal will be over-decomposed. This article uses the center frequency observation method to make judgments, and then determine the value of K . The VMD penalty item is set to 120.

It can be seen from Table 1 that when $K \geq 9$, the change of the center frequency of the IMF component tends to be stable. Therefore, the number of components is determined to be 9. The decomposed IMF components are shown in the Fig. 4.

The number of components determined by the center frequency observation method, although the center frequency tends to be stable. But in fact, when the value of the mode is too small, the mode will be under-decomposed, and some important information in the original signal will be filtered, which affects the prediction accuracy; on the other hand, when the value of the mode is too large, adjacent modes the center frequencies of the components will be close together, which will lead to defects such as modal repetition and excessive extra noise. Therefore, in order to avoid modal repetition and excessive additional noise, further confirm the number of modalities, calculate the Sample Entropy of the decomposed components IMF1-IMF9 according to formula (9), and

then fuse them according to their entropy characteristics. The entropy results are shown in Table 2.

Table 1. The center frequencies of different components

K	u_1	u_2	u_3	u_4	u_5	u_6	u_7	u_8	u_9	u_{10}
1	1.080									
2	1.050	0.226								
3	1.050	0.228	0.633							
4	1.021	0.232	0.359	0.646						
5	0.057	0.216	0.309	0.443	0.668					
6	0.048	0.214	0.254	0.360	0.616	0.724				
7	0.045	0.216	0.252	0.356	0.443	0.633	0.742			
8	0.042	0.116	0.252	0.356	0.443	0.633	0.742	0.810		
9	0.041	0.106	0.232	0.344	0.430	0.504	0.628	0.716	0.817	
10	0.040	0.105	0.228	0.322	0.401	0.489	0.536	0.607	0.711	0.816

Table 2. Sample entropy of VMD component

VMD component	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9
Sample Entropy	0.4045	0.1080	0.0418	0.0439	0.0435	0.0253	0.0995	0.1073	0.0241

The closer the entropy value of the IMF sample, the more common its sequences are, and the better the fusion. It can be seen from Table 2 that the entropy values of IMF2, IMF7 and IMF8 samples are close, the entropy values of IMF3, IMF4 and IMF5 samples are close, and the entropy values of IMF6 and IMF9 samples are close, so they are fused separately. Since IMF1 has no component close to its Sample Entropy, it does not need to be fused with other components. In this paper, the components with similar entropy are superimposed, and finally four photovoltaic components are obtained. The components after fusion are shown in Fig. 5.

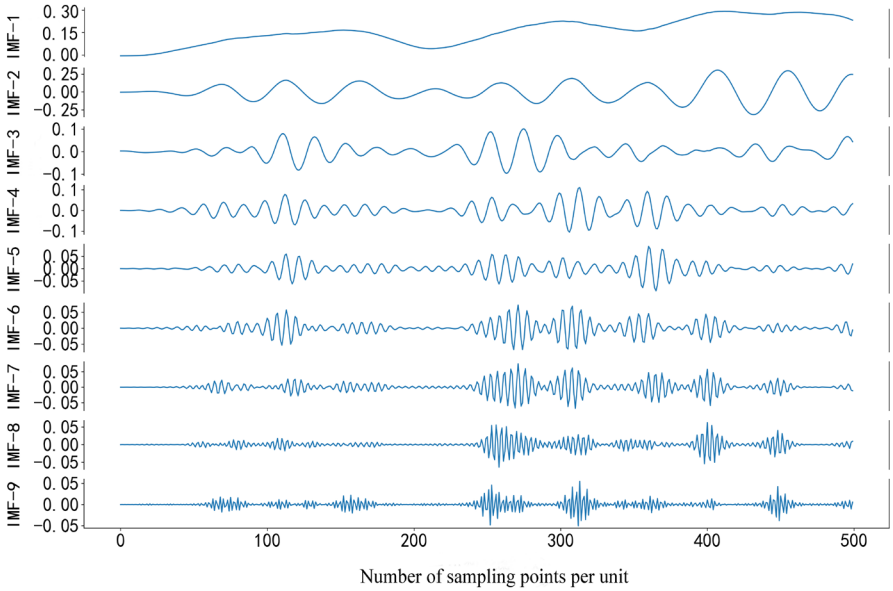

Fig. 4. VMD initial decomposition results

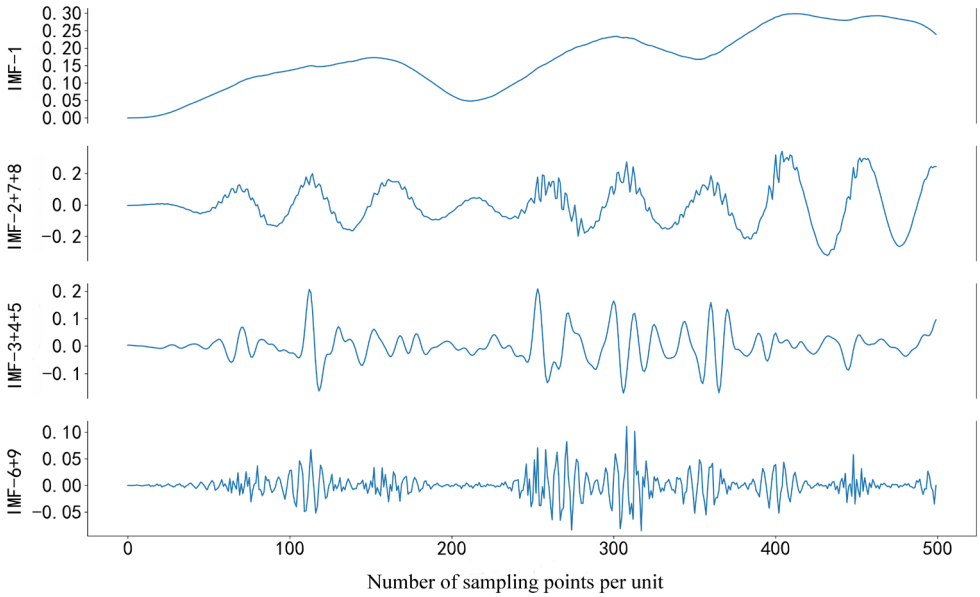

Fig. 5. Component fusion results

4.2 VMD-SE-TCAN Model Prediction

In order to verify the effect of Sample Entropy and Attention Mechanism in the model of this paper, the ablation experiment method is used in this paper for comparison and verification, and the TCN model in the experiment uses Adam as the optimizer and Epoches is set to 30.

This paper selects the coefficient of determination (R^2) and root mean square error (RMSE) as the error analysis indicators. R^2 responds to the accuracy of model prediction, and generally, a larger R^2 indicates a better model fit; RSME responds to the accuracy of model prediction, and the smaller the RSME value, the higher the model accuracy. R^2 is shown in formula (13), and RMSE is shown in formula (14).

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y}_i)^2}. \tag{13}$$

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

n is the number of predicted points; y_i is the actual value; $\overline{y}_i$ is the average value in the sequence; $\hat{y}_i$ is the predicted value [17].

The results of a two-day ablation experiment selected for this paper are shown in Fig. 6, and the performance comparison is shown in Table 3.

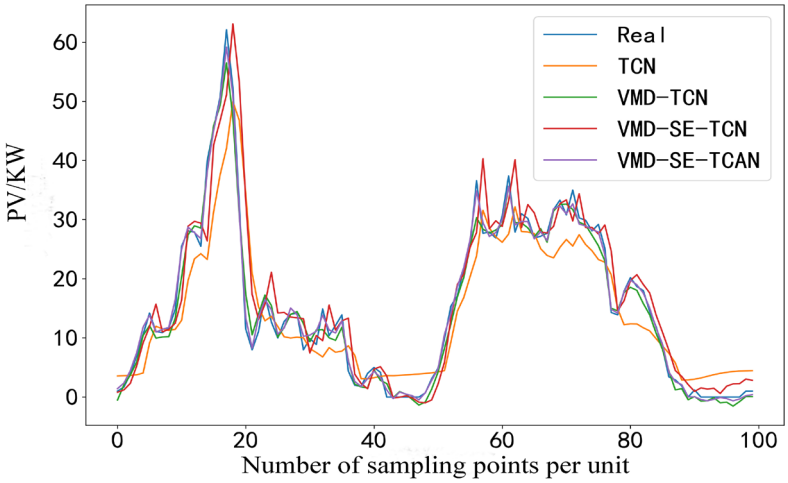

Fig. 6. Model ablation prediction results

Table 3. Comparison of ablation performance of models

Algorithm	R^2	RSME
TCN	0.897	0.083
VMD-TCN	0.910	0.074
VMD-SE-TCN	0.926	0.069
VMD-SE-TCAN	0.932	0.058

Through the analysis of Fig. 6 and Table 3, it can be seen that:

The R^2 value of TCN model is 0.897, RSME value is 0.083, and the R^2 value of VMD-TCN model is 0.910, RSME value is 0.074, which shows that the TCN model using VMD has better prediction effect;

The R^2 value of VMD-SE-TCN model is 0.926 and RSME value is 0.069 after using Sample Entropy to fuse components, which shows that the effect of component prediction is better after using Sample Entropy to fuse components, and can effectively alleviate the problem of low prediction accuracy caused by VMD over decomposition;

The R^2 value of VMD-SE-TCAN model with Attention Mechanism is 0.932 and RSME value is 0.058, which is better than VMD-SE-TCN model without Attention Mechanism. It shows that adding Attention Mechanism in TCN can better strengthen the influence of key features on prediction and improve the prediction effect of the model.

4.3 Comparison of Prediction Results of Different Models

To verify the performance of the proposed model, the commonly used prediction models LSTM, SVM and ELM are used as references, and compared with the proposed model VMD-SE-TCAN in this paper.

In the experiments, the proposed TCAN model adopts 7 TCN residual modules in the network structure, the number of convolutional kernels in the convolutional layer is set to 64, and the dropout rate of the model is set to 0.3. To ensure the consistency, 7 hidden layers are set in the LSTM model, the number of neurons in the hidden layer is also set to 64, and the dropout rate of the model is set to 0.3; the SVM adopts radial basis function; the number of neurons in the hidden layer of ELM is set to 64. The prediction results of each model for a two-day period selected in this paper are shown in Fig. 7, and the prediction performance is shown in Table 4.

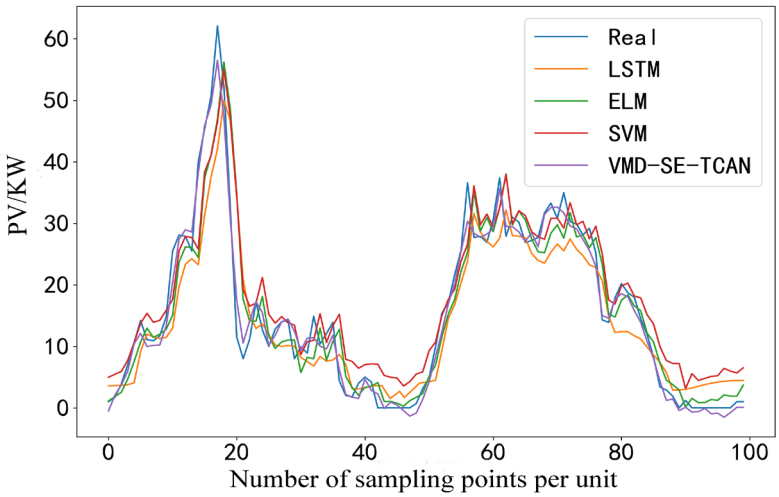

Fig. 7. Prediction results of different models

Table 4. Performance comparison of different models

Algorithm	R^2	RSME
LSTM	0.872	0.125
SVM	0.894	0.099
ELM	0.905	0.083
VMD-SE-TCAN	0.935	0.062

From Fig. 7 and Table 4, it can be seen that the RSME values of the model proposed in this paper are reduced by 50.40%, 37.37%, and 25.30%, respectively, compared with LSTM, SVM, and ELM methods, and the R^2 values are greater than those of other models, indicating that the prediction effect of the model proposed in this paper is better than that of the commonly used models, with better robustness and stability, and more suitable for short-term PV prediction.

5 Conclusion

Aiming at the problem of low prediction accuracy of photovoltaic power generation, a short-term photovoltaic prediction model based on VMD-SE-TCAN is proposed and verified by real data sets. Firstly, Sample Entropy is used to fuse different frequency components generated by VMD. Then, the TCAN model combined with Temporal Convolution Neural network and Attention Mechanism is used to predict each component. Finally, the prediction results of each component are superimposed to output photovoltaic prediction power. Through experiments and comparative verification, the combined prediction method effectively improves the accuracy of photovoltaic power prediction. The accuracy of photovoltaic power generation prediction can be further improved by weather clustering prediction and TCN network parameter optimization.

6 Acknowledgement

This work was supported by the Science and Technology Research Project of Jilin Province Education Department (JJKH20210100KJ).

References

[1] K.-Y. Bae, H.-S. Jang, D.- K. Sung, Hourly Solar Irradiance Prediction Based on Support Vector Machine and Its Error Analysis, IEEE Transactions on Power Systems 32(2)(2017) 935-945.

[2] D.-F. Yang, Q.-K. Ma, Y. Wang, M.-J. Peng, Optimal Configuration of Stand-Alone Microgrid Capacity Considering Volatility of Photovoltaic Output, Journal Of Northeast Electric Power University 39(4)(2019) 19-26.

[3] M. Ceci, R. Corizzo, F. Fumarola, D. Malerba, A. Rashkovska, Predictive Modeling of PV Energy Production: How to Set Up the Learning Task for a Better Prediction? IEEE Transactions on Industrial Informatics 13(3)(2017) 956-966.

[4] Y.-F. Huang, W. He, S.-Y. Liu, Research on Energy Optimization of Home Microgrid for Intelligent Power Consumption, Journal of Northeast Electric Power University 40(4)(2020) 29-34.

[5] L.-M. Yin, L. Wang, G. Lei, J.-X. Jiang, Optimal Dispatch of Microgrid Considering Demand Response and Comprehensive Battery Loss, Journal of Northeast Electric Power University 40(2)(2020) 37-48.

[6] L.-T. Li, Y. Luo, Prediction of distributed photovoltaic generation based on GA-BP, Journal of Sichuan University of Science & Engineering (Natural Science Edition) 30(5)(2017) 31-36.

[7] H.-M. Sheng, J. Xiao, Y.-H. Cheng, Q. Ni, S. Wang, Short-Term Solar Power Forecasting Based on Weighted Gaussian Process Regression, IEEE Transactions on Industrial Electronics 65(1)(2018) 300-308.

[8] T. Ma, H.-X. Yang, L. Lu, Solar photovoltaic system modeling and performance prediction, Renewable & Sustainable Energy Reviews 36(2014) 304-315.

[9] Y.-P. Cai, S.-Y. Lu, T.-F. Chu, L.-S. Chang, J.-S. Liu, J. Liu, Study on Capacity of Distribution Network with Photovolatic Power Generation, Journal of Northeast Electric Power University 39(1)(2019) 9-15.

[10] E. Ogliari, A. Dolara, G. Manzolini, S. Leva, Physical and hybrid methods comparison for the day ahead PV output power forecast, Renewable Energy 113(2017) 11-21.

[11] X.-Y. Wang, D. Luo, Y.-L. Sun, C. Li, J. Li, Combined Forecasting Method of Daily Photovoltaic Power Generation in Microgrid Based on ABC-SVM and PSO-RF Models, Acta Energiae Solaris Sincia 41(3)(2020) 177-183.

[12] B. Chen, J.-H. Li, Combined probabilistic forecasting method for photovoltaic power using an improved Markov chain, IET Generation, Transmission & Distribution 13(19)(2019) 4364-4373.

[13] M.-J. Sanjari, H.-B. Gooi, Probabilistic Forecast of PV Power Generation Based on Higher Order Markov Chain, IEEE Transactions on Power Systems 32(4)(2017) 2942-2952.

[14] M.-K. Behera, N. Nayak, A comparative study on short-term PV power forecasting using decomposition based optimized extreme learning machine algorithm, Engineering Science and Technology, an International Journal 23(1)(2020) 156-167.

[15] C. Zhang, J.-B. Bai, K. Lan, X.-X. Huan, C. Fan, X. Xia, Photovoltaic Power Generation Prediction Based on Data Mining and Genetic Wavelet Neural Network, Acta Energiae Solaris Sincia 42(3)(2021) 375-382.

[16] J. Tan, C.-H. Deng, W. Yang, N. Liang, F.-J. Li, Ultra-short-term Photovoltaic Power Forecasting in Microgrid Based on Adaboost Clustering, Automation of Electric Power Systems 41(21)(2017) 33-39.

[17] Y.-F. Chen, M. Wen, K. Zhang, S. Yu, Short-term photovoltaic output prediction based on similar day matching and TCN attention, Electrical Measurement & Instrumentation. <http://kns.cnki.net/kcms/detail/23.1202.TH.20200727.1612.024.html>, 2020 (accessed 21.05.01).

[18] J.-X. Lu, Q.-P. Zhang, Z.-H. Yang, M.-F. Tu, J.-J. Lu, H. Peng, Short-term Load Forecasting Method Based on CNN-LSTM Hybrid Neural Network Model, Automation of Electric Power Systems 43(8)(2019) 131-137.

[19] Z.-F. Lin, L.-L. Cheng, G.-H. Huang, Electricity consumption prediction based on LSTM with attention mechanism, IEEJ Transactions on Electrical and Electronic Engineering 15(4)(2020) 556-562.

[20] L. Pantiskas, K. Verstoep, H. Bal, Interpretable Multivariate Time Series Forecasting with Temporal Attention Convolutional Neural Networks, in: Proc. 2020 IEEE Symposium Series on Computational Intelligence (SSCI), 2020.

[21] S. Yang, D.-S. Luo, H.-Y. He, J.-W. Yang, S.-Y. Hu, Output Power Forecast of PV Power System Based on EMD-LSSVM Model, Acta Energiae Solaris Sincia 37(6)(2016)1387-1395.

[22] J. Ospina, A. Newaz, M.-O. Faruque, Forecasting of PV plant output using hybrid wavelet-based LSTM-DNN structure model, IET Renewable Power Generation 13(7)(2019)1087-1095.

[23] L.-W. Yang, X.-Q. Gao, J.-X, Jiang, Q.-Q. Lyu, Z.-C. Li, Short-term Photovoltaic Output Power Prediction Based on Wavelet Transform and Neural Network, Acta Energiae Solaris Sincia 41(7)(2020)152-157.

[24] W.-G. Li, W.-W. Xu, X.-G. Wang, Fault Line Selection Method for Active Distribution Network Based on VMD Energy Proportion, Journal of Northeast Electric Power University 39(5)(2019) 23-33.

[25] J.-X. Yang, S. Zhang, J.-C Liu, J.-Y. Liu, Y. Xiang, X.-Y. Han, Short-term Photovoltaic Power Prediction Based on Variational Mode Decomposition and Long-short Term Memory with Dual-stage Attention Mechanism, Automation of Electric Power Systems 45(3)(2021)174-182.

[26] X.-Y. Yu, Y.-M. Zhao, N.-N. Yang, T. Yue, C.-Y. Gao, Photovoltaic power generation forecasting based on VMD-SE-LSSVM and iterative error correction, Acta Energiae Solaris Sincia 41(2)(2020) 310-318.

[27] H. Ren, X.-G. Wang, Review of attention mechanism, Journal of Computer Applications 41(z1)(2021) 1-6.

[28] The Desert Knowledge Australia Solar Centre. <http://dkasolarcentre.com.au/locations/alice-springs> (accessed 20.10.23).