

DVS: Deep Visibility Series and Application in Construction Cost Index Forecasting

Tianxiang Zhan^{a,b,*}, Yuanpeng He^{a,b,*}, Hanwen Li^{a,b,*}, Fuyuan Xiao^{b,c,**}

^a*College of Computer and Information Science College of Software, Southwest University, Chongqing, 400715, China*

^b*School of Big Data and Software Engineering, Chongqing University, Chongqing, 401331, China*

^c*National Engineering Laboratory for Integrated Aero-Space-Ground-Ocean Big Data Application Technology, China*

Abstract

Time series forecasting is a hot spot in recent years. Visibility Graph (VG) algorithm is used for time series forecasting in previous research, but the forecasting effect is not as good as deep learning prediction methods such as methods based on Artificial Neural Network (ANN), Convolutional Neural Network (CNN) and Long Short-Term Memory Network (LSTM). The visibility graph generated from specific time series contains abundant network information, but the previous forecasting method did not effectively use the network information to forecast, resulting in relatively large prediction errors. To optimize the forecasting method based on VG, this article proposes the Deep Visibility Series (DVS) module through the bionic design of VG and the expansion of the past research. By applying the bionic design of biological vision to VG, DVS has obtained superior forecasting accuracy. At the same time, this paper applies the DVS forecasting method to the construction cost index forecast, which has practical significance.

Keywords: Deep Learning, Complex Network, Visibility Graph, Time Series Forecasting, Construction Cost Index

1. Introduction

Time series forecasting is a research hot spot in the fields of mathematics, economy and so on. Time series forecasting technology has made many contributions to society such as disease prevention, risk forecasting, economic planning and so on. It is the goal of researchers to improve the efficiency and reduce the error of time series forecasting.

Some classic statistical models such as Autoregressive Integrated Moving Average Model (ARIMA) [1] can be combined with new technologies such as neural networks to reduce forecasting errors and improve efficiency. Time series forecasting based on Visibility Graph (VG) is a classic time series forecasting method

*The e-mail of Tianxiang Zhan: zhantianxiangswu@163.com; The e-mail of Yuanpeng He: heyuanpengswu@163.com; The e-mail of Hanwen Li: lihanwenswu@163.com

**Corresponding author: Fuyuan Xiao is with the School of Big Data and Software Engineering, Chongqing University, Chongqing 401331, China. (e-mail: xiaofuyuan@cqu.edu.cn; doctorxiaofy@hotmail.com)

which has small forecasting error [2, 3, 4]. However, due to the high time complexity of the VG algorithm combined with the random walk algorithm, effective prediction information cannot be fed back in time. To optimize the method based on VG, this paper combines VG and Convolutional Neural Network (CNN). The article proposes Enhanced Visibility Graph (EVG) and Deep Visibility Series (DVS) algorithm, which makes the visibility graph contains more information of network and rough time series, which improves the efficiency and further reduces the forecasting error.

In the experiment, DVS method is used to predict a actual problem Construction Cost Index (CCI). The CCI is an indicator that reflects the construction cost published by Engineering News Record (ENR) once a month [5, 6]. CCI forecasts are greatly meaningful in the construction and financial fields, and the classic data set of CCI time series forecasts effectively reflects the performance of forecasting methods.

The structure of this paper is as follows: The second part introduces the VG algorithm and previous work. The third part will introduce how to optimize the VG algorithm. The fourth part will forecast the CCI data set, and compare it with statistical forecasting methods, hybrid forecasting methods, and machine learning methods, and conduct ablation experiments to prove the role of DVS. The last part summarizes the article and discusses future work.

2. Previous research

This section will introduce previous research on VG. Firstly, the definition of time series T is as follows:

$$T = \{T_1, T_2, T_3, \dots, T_n\} \quad (1)$$

$$T_i = (t_i, v_i) \quad (2)$$

where t_i is the time of time point T_i and v_i is the value of the time point T_i .

2.1. Visibility Graph and forecasting method

2.1.1. Visibility Graph

VG is a algorithm of converting a time series into a network. By inputting the time series, it can output an adjacency matrix that represents the connection relationship between the time series inspection points [7].

Definition 1. *The visual relationship is a visual relationship constructed between two nodes in the time series based on the numerical value. The visual relationship has the same bionic characteristics as human vision. If an object is blocked behind an obstacle, humans cannot see the object, and there is no visual relationship between humans and*

the object. Humans can see obstacles, so obstacles are related to possession. So the definition of the relationship is as follows:

$$visual(T_i, T_j) = \begin{cases} 0, & \arg \max_{\theta \in [i,j], \theta \in Z} \tan(T_\theta, T_j) \neq i \\ 1, & \arg \max_{\theta \in [i,j], \theta \in Z} \tan(T_\theta, T_j) = i \end{cases} \quad (3)$$

The Formula.4 here means that if the slope between nodes i and j is the maximum value of the slope between node i and node j , then nodes i and j are visible, and the boolean state variable is 1, otherwise 0.

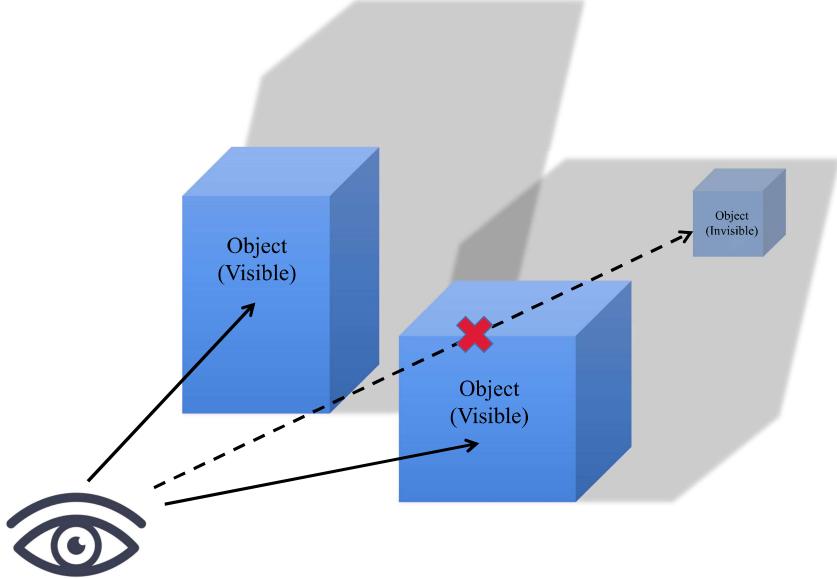


Figure 1: Bionic example of the VG algorithm

Fig.1 is a bionic example of the VG algorithm. For the observer, observation can bargain two objects, but the blocked object cannot be seen. The observer only knows whether the object can be seen or not, and the properties of the object such as color cannot be seen. For time series, the property that the VG algorithm cannot represent is the value corresponding to the time node. The visual relationship can construct a rough adjacency matrix A to represent the two-dimensional relationship in the entire time series.

Definition 2. *The definition of visual matrix A is as follows:*

$$A_{ij} = visual(T_i, T_j) \quad (4)$$

where visual matrix A contains the association relationship of the visual relationship.

Fig.2 below is an example of a VG. If the two nodes are visible, use the edge to connect the corresponding visual nodes. The connections between nodes in the time series are intuitively established.

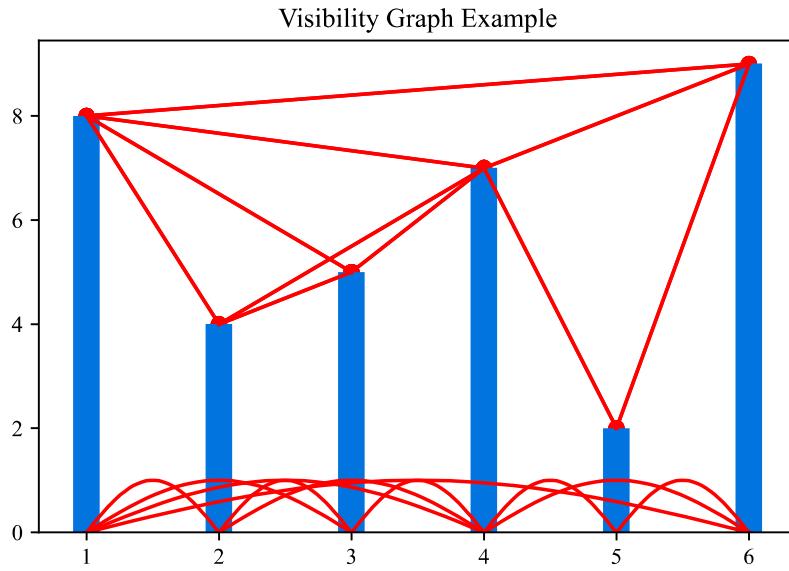


Figure 2: Example of VG ($T = \{(1,8), (2,4), (3,5), (4,7), (5,2), (6,9)\}$)

2.1.2. Forecasting based on Visibility Graph

In the previous research, after the time series was processed by the VG algorithm, the network generated by the VG was used to cooperate with the random walk algorithm. The purpose of the random walk algorithm is to find the similarity or the strength of the relationship between two nodes in the network [8]. Finally, by selecting a node or a group of nodes with the highest similarity or relationship strength to fit the last time node linearly, the final result is obtained by weighting according to time distance, node similarity or other distances such as Fig.3.

2.2. How the previous research work

For the previous research, the forecasting algorithm is divided into the following three steps:

1. Generate a visibility graph of the time series.
2. After random walk, find the point or a group of points with the highest similarity to the last node.
3. After the selected set of points is linearly fitted to the last node, the weighted sum of the fitting results is used as the final forecasting result.

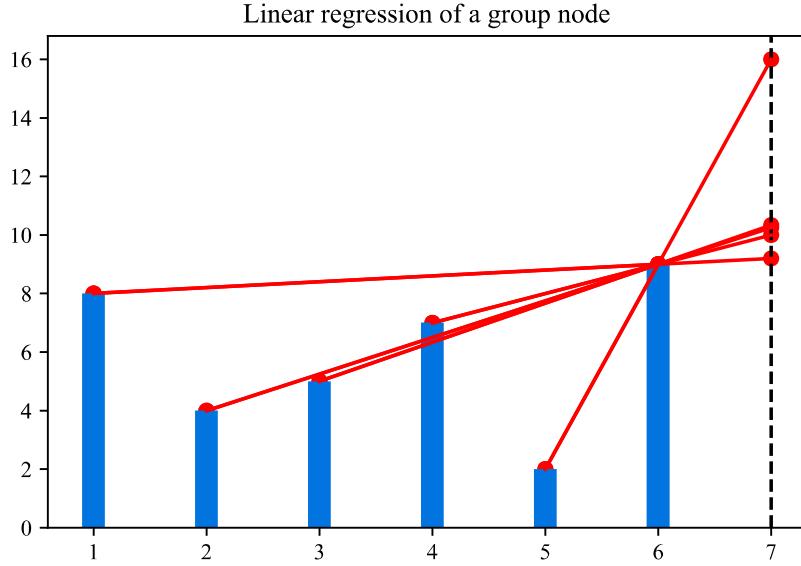


Figure 3: Linear Regression of each node and last node

Different from statistical linear fitting, the VG algorithm can be used to determine that some points are involved in the fitting instead of all points. At the same time, different time points are assigned a weight based on distance or similarity, which further reduces the forecasting error of the algorithm. The above analysis is the reason why the previous research is effective. For forecasting algorithms based on VG, how to efficiently use the nature of the network while retaining time series information is the core of the research.

2.3. Disadvantages of previous research

For the previous research, the feature of the network was preserved through the random walk algorithm, but in the end, only the time of the node was used, and other information was not used effectively. At the same time, the random walk algorithm requires a large number of cycles, and it stops when the state of the entire network tends to converge. In the end, only a few nodes in a row component of the adjacency matrix are used for forecasting. A large number of system resources are consumed in the process of random walk, and the information of the network is not used effectively.

3. Deep Visibility Series

This section will introduce the proposed DVS.

3.1. Enhanced visibility graph

For a time series T and its adjacency matrix A obtained by the VG algorithm, an enhanced visibility graph (EVG) algorithm is proposed.

Definition 3. EVG saves time series information in the network, and the output of EVG is an enhanced adjacency matrix B .

$$B_{ij} = \frac{A_{ij} * v_j}{\text{Degree}(T_i)} \quad (5)$$

$$\text{Degree}(T_i) = \sum_{j=1}^n A_{ij} \quad (6)$$

where the meaning of the Degree function is the node degree of node T_i in the VG network. Each element B_{ij} in matrix B represents the value at which node T_i can see node T_j . Because node T_i has chances to see more than one node, all the nodes it sees need to be divided by the degree of node T_i . Matrix B also reflects the network characteristics of the time series while retaining the element information of the time series.

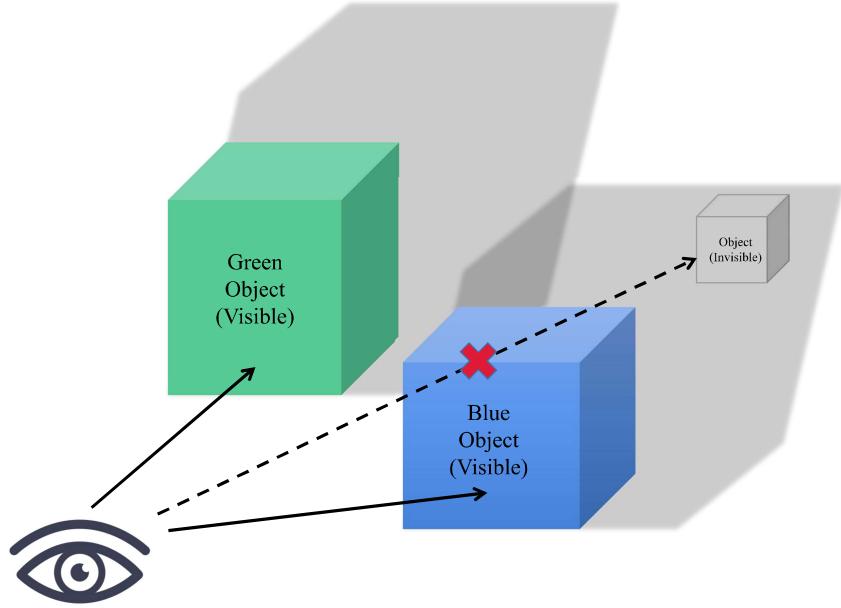


Figure 4: Bionic example of the EVG algorithm

The EVG algorithm is also a bionic algorithm. Fig.4 is a bionic example of the EVG algorithm. For the observer, the observer can not only see the object but also the color of the object. The VG algorithm is equivalent to human beings knowing that I can see the object, but not knowing what the object is. EVG has

increased the ability for humans to see exactly what an object is. Fig.5 is an example of the previous VG and time series conversion to EVG, and observer can recognize the color of objects.

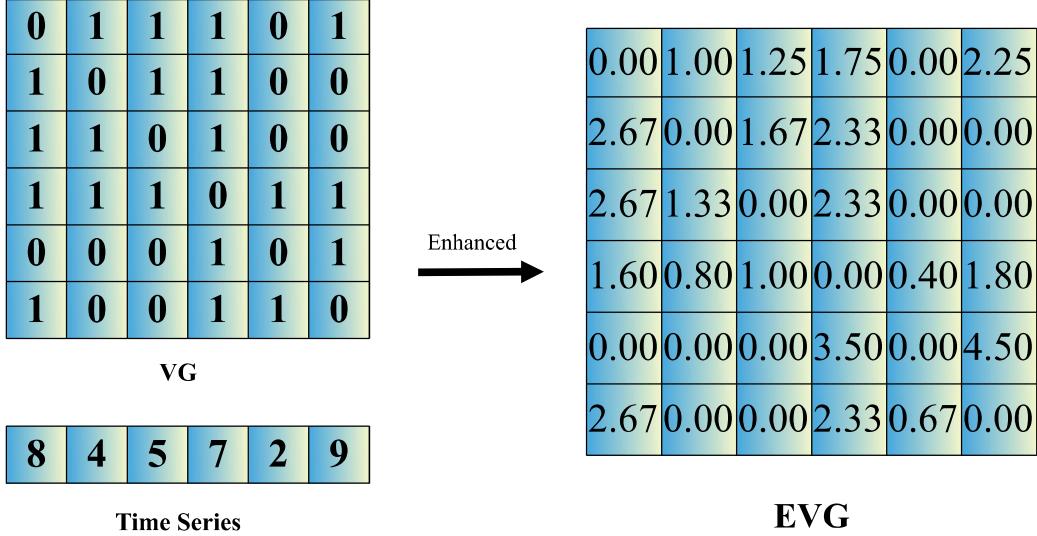


Figure 5: Example of transform from time series and VG to EVG ($T = \{(1,8), (2,4), (3,5), (4,7), (5,2), (6,9)\}$)

3.2. Deep Visibility Series Forecasting

Through the EVG algorithm, the problem of retaining network information is solved. Another problem is how to use the matrix generated by the existing EVG.

3.2.1. How to determine the form of information for forecasting

The matrix generated by EVG can be regarded as a picture. For pictures, a common operation is to use Convolutional Neural Network (CNN) or Artificial Neural Network (ANN) for picture recognition, which is a classification problem [9]. The time series forecasting problem is a non-discrete variable forecasting problem, and the matrix generated by EVG cannot be processed by means of pictures.

Previous research has chosen to abandon most of the information in dealing with network information, allowing limited network information to be expressed in time series. The final forecasting operation is still the serialized information. On the other hand, the purpose of time series forecasting is not to over-strengthen the extraction of network information, but to enhance the expression of the sequence, so here is a way to compress the sequence generated by the EVG.

Definition 4. When compressing the matrix generated by the EVG, the operation cannot be converted into a sequence such as Hash mapping by simply using a mapping method, because the original time series has three characteristics:

1. *Orderliness*: The order of the time series cannot be changed at will.
2. *Limited length*: The number of time points is limited, one time corresponds to one value, and the number of values can only be equal to the number of times.
3. *Interpret ability*: The value of each event has its own physical meaning and cannot be missing.

Therefore, the compression algorithm uses the bionic idea here, which is defined as follows:

$$zip_i = \sum_{i=1}^n B_{ij} \quad (7)$$

where matrix B is the matrix generated by EVG. The $i - th$ element in the compressed sequence represents the sum of the visible values of the time node T_i . For humans, EVG distinguishes what each object is.

Therefore, under the action of the EVG module and the ZIP module, the time series is converted into a time series that retains network information. Fig.6 shows the compression process of the time series matrix generated from EVG.

Definition 5. The combination of EVG module and ZIP module constitutes a Deep Visibility Series (DVS) module.

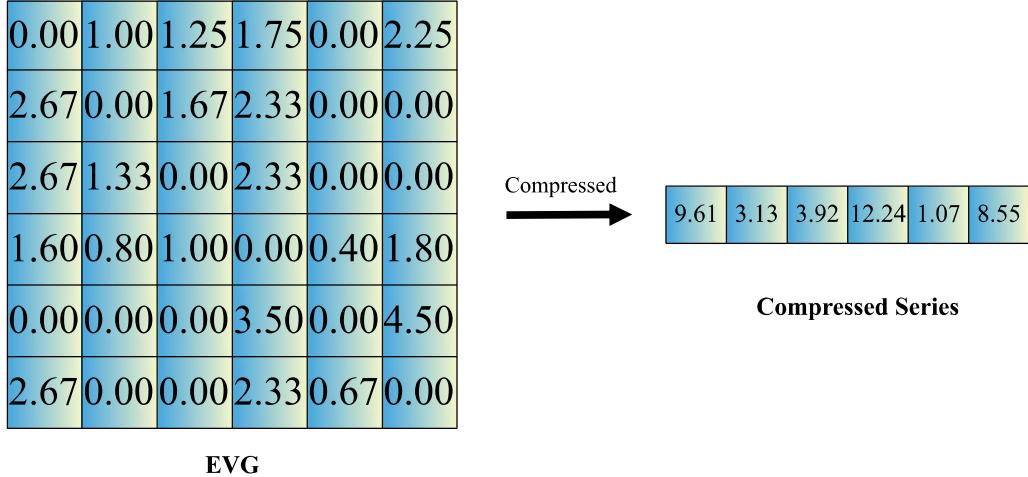


Figure 6: Example of compressing process ($T = \{(1,8), (2,4), (3,5), (4,7), (5,2), (6,9)\}$)

3.2.2. How to forecast

In the compression process, although the boundary information between time nodes is lost, the visible features are retained. Here, CNN is used to receive the sequence generated by DVS for deep learning. Compared with ANN, CNN can effectively extract the required features in complex information, which

is represented here as the required time series information and visual network information. ANN, CNN, Long Short-Term Memory (LSTM) Network can all be used as training networks for DVS. The algorithm of DVS is as follows.

Algorithm 1 Deep Visibility Series Forecasting

```

1: Calculate matrix A as:
2: for Each time point  $T_i$  do
3:   for Each time point  $T_j$  do
4:

$$A_{ij} = \text{visual}(T_i, T_j)$$


5: Calculate matrix B as:
6: for Each time point  $T_i$  do
7:   for Each time point  $T_j$  do
8:

$$B_{ij} = \frac{A_{ij} * v_j}{\text{Degree}(T_i)}$$


9: Calculate vector zip as:
10: for Each time point  $T_i$  do
11:

$$\text{zip}_i = \sum_{j=1}^n B_{ij}$$


12:

$$\text{zip} \xrightarrow{\text{input}} \text{CNN} \xrightarrow[\text{test}]{\text{train}} \hat{y}_{n+1}$$

return  $\hat{y}_{n+1}$ 

```

4. Application in Construction Cost Index Forecasting

4.1. Data set information

This experiment uses the CCI data set. The CCI data set has 295 construction cost data values, and the time includes the period from January 1990 to July 2014. The experimental CCI data set is processed, and the time series are split with a window size of 30 to generate 264 sets of sequences [5, 6]. The ratio of training set to test set in the experiment is 8:2.

4.2. Model parameters

In the experimental part, the network structure used this time is shown in Fig.7.

The network used in the experiment is composed of DVS module and CNN, where CNN has two convolutional layers. The network used in the experiment uses the Cyclic Learning Rate (CLR) algorithm to dynamically adjust the learning rate while using the Adam optimizer to accelerate learning. Tab.1 shows the model parameters of the DVS+CNN network.

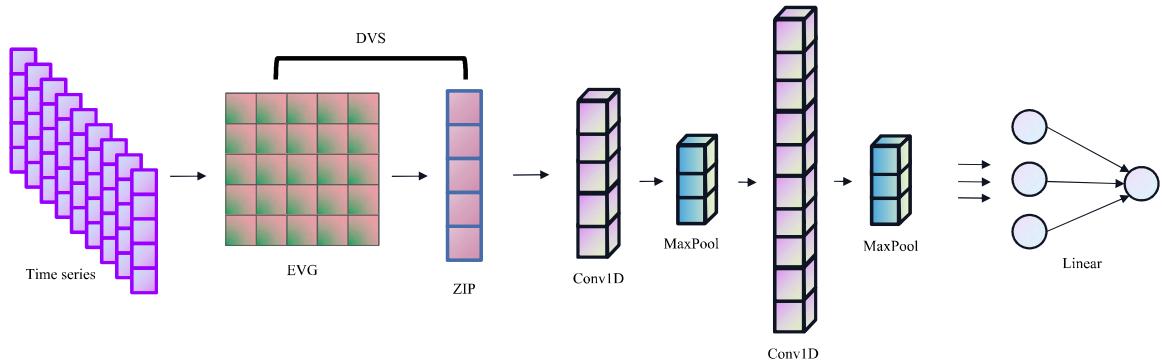


Figure 7: Example of VG ($T = \{(1, 8), (2, 4), (3, 5), (4, 7), (5, 2), (6, 9)\}$)

Table 1: DVS+CNN training network parameters

Structure	Training Iteration	Learning Rate	Loss Function	Optimization Function	
DVS					
Convolution Layer 1:					
In-Channel = 1					
Kernel Size = 3					
Out-Channel = 8					
Pooling Layer1:					
Pool Size=2					
Convolution Layer 2:					
DVS+CNN	In-Channel = 8	100	CLR $(10^{-4}, 10^{-12})$	MSE	Adam
	Kernel Size = 3				
	Out-Channel = 16				
	Pooling Layer 2:				
	Pool Size=2				
	Flatten Layer				
	Linear Layer				

4.3. Experimental results

4.3.1. Experimental error standard

To evaluate the forecasting of DVS+CNN and comparison method, five measures of error is error standard in the experiment: mean absolute difference (MAD) [10] , mean absolute percentage error (MAPE) [11] , symmetric mean absolute percentage error (SMAPE) [12], root mean square error (RMSE) [13] , and normalized root mean squared error (NRMSE) [14] .

$$MAD = \frac{1}{N} \sum_{t=1}^N |\hat{y}(t) - y(t)| \quad (8)$$

$$MAPE = \frac{1}{N} \sum_{t=1}^N \frac{|\hat{y}(t) - y(t)|}{y(t)} \quad (9)$$

$$SMAPE = \frac{2}{N} \sum_{t=1}^N \frac{|\hat{y}(t) - y(t)|}{\hat{y}(t) + y(t)} \quad (10)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N |\hat{y}(t) - y(t)|^2} \quad (11)$$

$$NRMSE = \frac{\sqrt{\frac{1}{N} \sum_{t=1}^N |\hat{y}(t) - y(t)|^2}}{y_{max} - y_{min}} \quad (12)$$

where $\hat{y}(t)$ is the predicted value, $y(t)$ is the true value and N is the total number of $\hat{y}(t)$.

4.3.2. Deep Visibility Series forecasting result

After dividing the CCI data set into a training set and a test set, the DVS+CNN network trains the CCI training set. After 100 iterations, the CCI test set is used as input for forecasting. Fig.8 is the forecasting effect of DVS+CNN on the test set. The forecasting result of DVS+CNN is close to the actual value of the CCI training set. At the same time, the forecasting value of DVS+CNN is stable and there is no jitter.

4.3.3. Comparison of previous research and classical forecasting method

In terms of comparison methods, the experiment selected multiple methods from five different perspectives for comparison. The five perspectives are: previous perspectives, statistical methods, machine learning methods, hybrid methods and ablation experiments.

Mao and Xiao's Method is a prediction method using VG in previous research [4]. According to the introduction of past research, Mao and Xiao's method adopts a method of weighting the results of a pair of points linear fitting according to the distance.

Fig.9 shows the predicted value of DVS+CNN, the predicted value of Mao and Xiao's method and the actual value of CCI. Different from the prediction method in the previous study, the prediction of

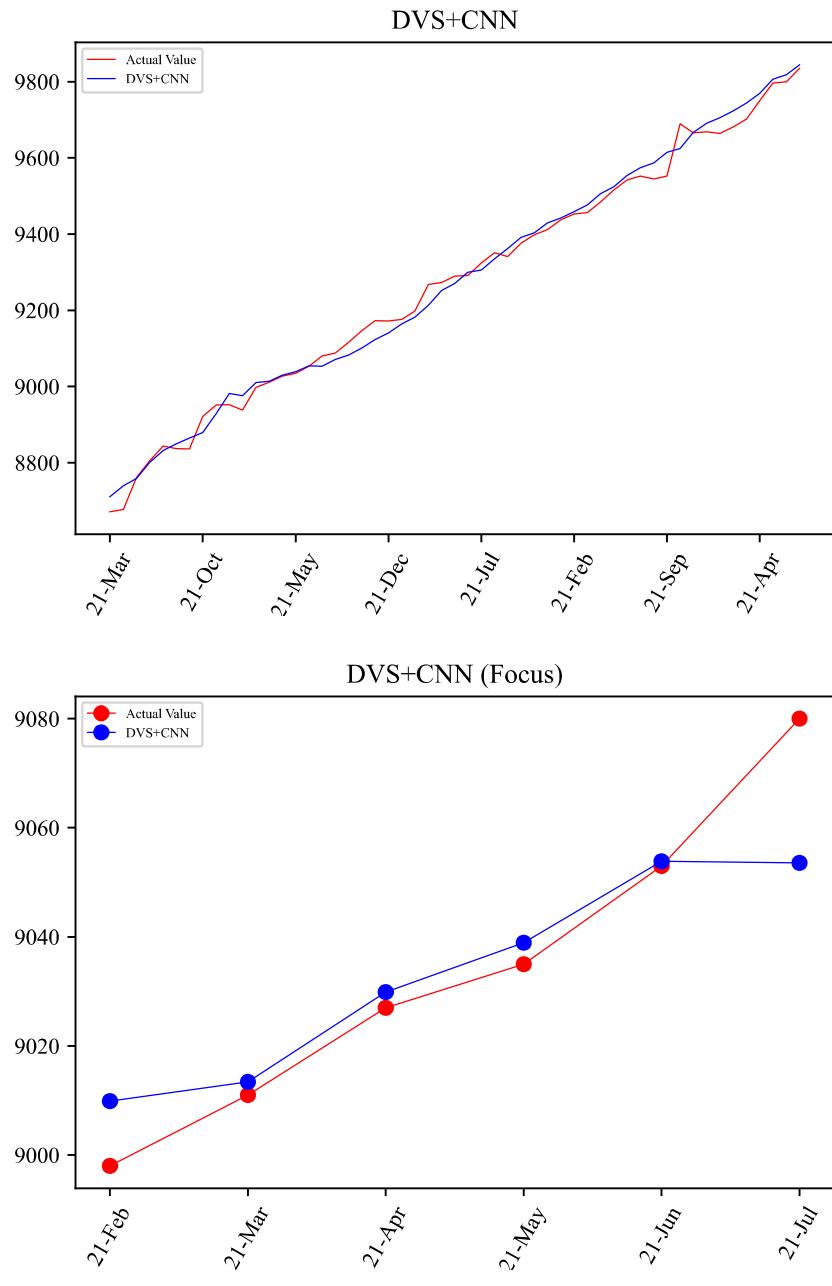
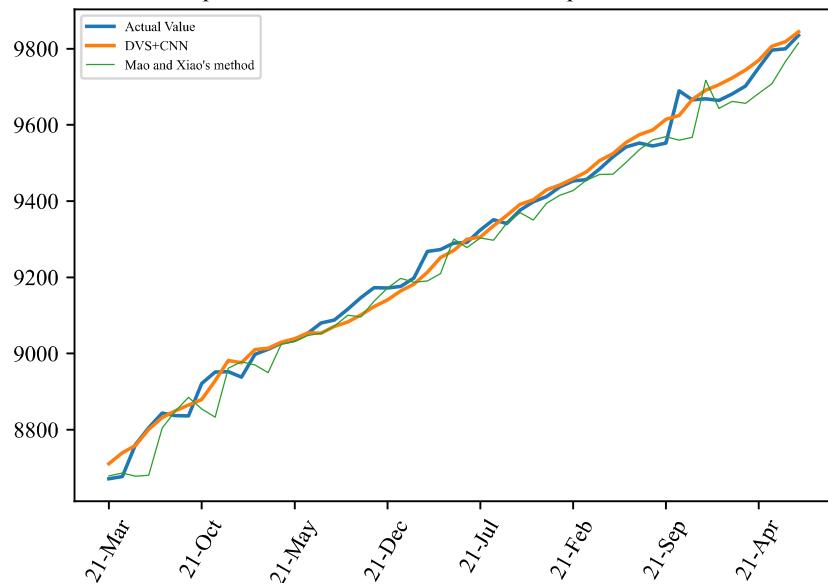


Figure 8: DVS+CNN forecasting curve and CCI actual value

Comparison between DVS+CNN and previous research



Comparison between DVS+CNN and previous research (Focus)

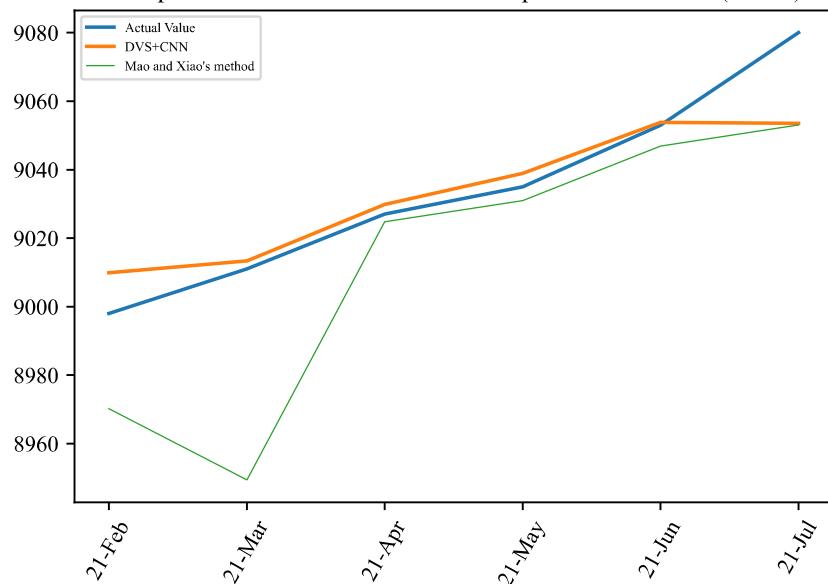


Figure 9: DVS+CNN forecasting curve and previous research forecasting curve

DVS+CNN is closer to the actual value, and the prediction result of DVS+CNN is less jitter than the previous study. The specific forecasting error between DVS+CNN and previous research methods is shown in Tab.2.

Table 2: Forecasting error between DVS+CNN and previous research methods

	MAD	MAPE	SMAPE	RMSE	NRMSE
DVS+CNN	22.8781	0.2467	0.2468	28.4596	84.5042
Mao and Xiao's method	36.2932	0.3936	0.3924	48.5065	143.8612

For the comparison between DVS+CNN and statistical methods, the statistical methods are selected: Simple Moving Average (SMA) (K=1) [15], Autoregressive Integrated Moving Average model (ARIMA) [1], Seasonal Autoregressive Integrated Moving Average model (Seasonal ARIMA) [16] and ExponenTial Smoothing (ETS) [17].

Fig.10 shows the predicted value of DVS+CNN, the predicted value of statistical methods and the actual value of CCI. The statistical method and DVS+CNN have less jitter, and the trend of the predicted curve is close to the actual curve. But compared with statistical methods, DVS+CNN is closer. The specific forecasting error between DVS+CNN and statistical methods is shown in Tab.3.

Table 3: Forecasting error between DVS+CNN and statistical methods

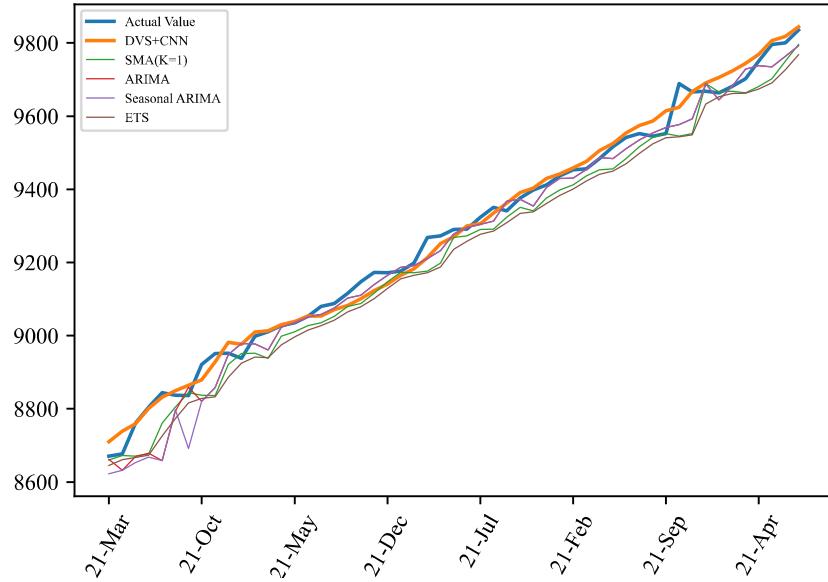
	MAD	MAPE	SMAPE	RMSE	NRMSE
DVS+CNN	22.8781	0.2467	0.2468	28.4596	84.5042
ARIMA	33.5646	0.3683	0.3669	49.401	145.05
Seasonal ARIMA	37.0188	0.4084	0.4066	54.5338	159.4385
SMA(K=1)	45.1132	0.4905	0.4886	56.2097	166.7715
ETS	57.8689	0.6299	0.6273	65.8998	196.6236

For the comparison between DVS+CNN and machine learning methods, the machine learning methods are selected: Decision Tree Regression (DTR) [18], Ordinary least squares Linear Regression (Linear) [19], Lasso model fit with Least Angle Regression (Lasso) [20], Support Vector Machines Regression (SVM) [21], Bayesian Ridge Regression (Bayesian) [22] and Logistic Regression (Logistic) [23, 24].

Fig.11 shows the predicted value of DVS+CNN, the predicted value of machine learning methods and the actual value of CCI. Machine learning methods are not as stable as statistical methods. For example, the prediction effect of DTR has a large error. Some methods of machine learning have low prediction errors, such as Lasso. Compared with machine learning methods, DVS+CNN has the highest prediction accuracy and strong stability. The specific forecasting error between DVS+CNN and machine learning method is shown in Tab.4.

For the comparison between DVS+CNN and hybrid methods, the statistical method is selected: ARIMA+ANN [25] and ETS+ANN [26]. Hybrid forecasting methods usually combine a statistical forecasting method with a deep learning method. In the hybrid method, the deep learning network generally chooses ANN

Comparison between DVS+CNN and statistical forecasting methods



Comparison between DVS+CNN and statistical forecasting methods (Focus)

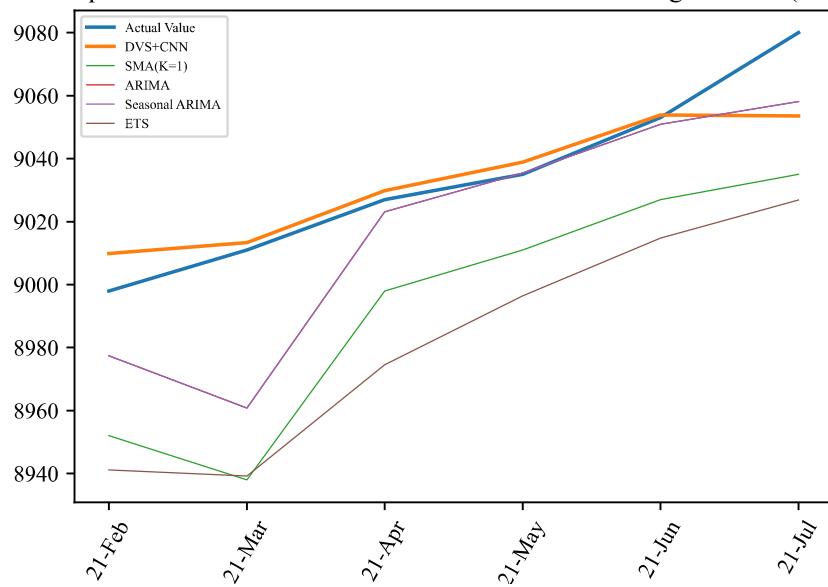
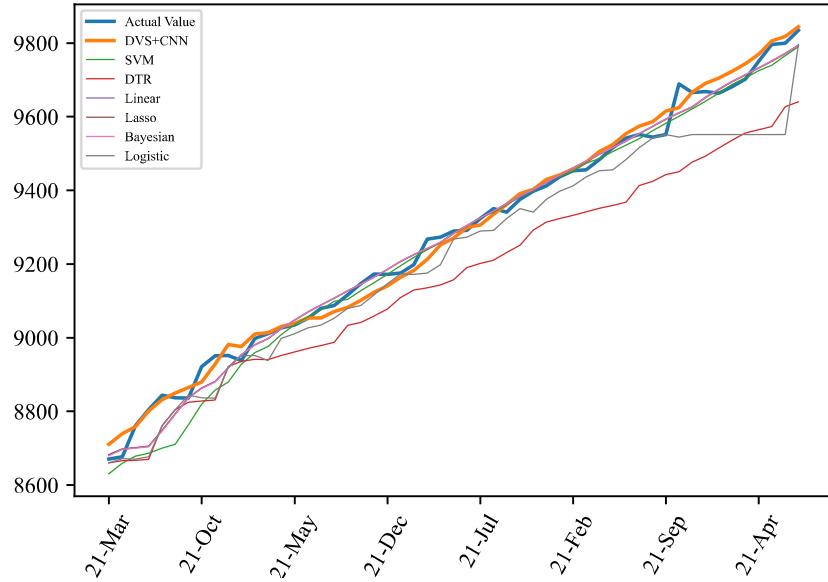


Figure 10: DVS+CNN forecasting curve and statistical method forecasting curve

Comparison between DVS+CNN and machine learning regression methods



Comparison between DVS+CNN and machine learning regression methods (Focus)

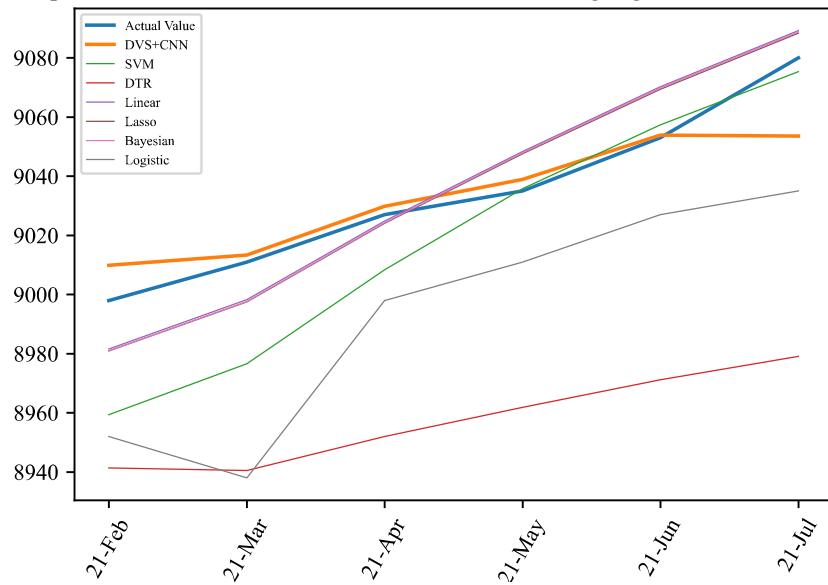


Figure 11: DVS+CNN forecasting curve and machine learning method forecasting curve

Table 4: Forecasting error between DVS+CNN and machine learning methods

	MAD	MAPE	SMAPE	RMSE	NRMSE
DVS+CNN	22.8781	0.2467	0.2468	28.4596	84.5042
Lasso	23.2985	0.2542	0.2537	32.7596	98.2587
Linear	23.3821	0.255	0.2546	32.5878	97.7052
Bayesian	23.4487	0.2558	0.2553	32.9991	98.7953
SVM	30.8839	0.3412	0.3399	46.6137	136.8565
Logistic	62.283	0.6711	0.6671	84.0805	249.4629
DTR	111.3944	1.2038	1.1952	122.5928	391.5623

to process the residuals in the statistical prediction, and then completes the prediction through synthesis, which has a lower prediction error. DVS+CNN is also an optimized time series prediction method. DVS optimizes the VG algorithm. By comparing DVS+CNN and hybrid methods, it can effectively show the advantages of DVS+CNN.

Fig.12 shows the predicted value of DVS+CNN, the predicted value of Mao and Xiao’s method and the actual value of CCI. Compared with pure ARIMA and ETS, the hybrid method further reduces the prediction error. Although the hybrid method has a lower prediction error, compared with DVS+CNN, DVS+CNN has higher prediction accuracy. The specific forecasting error between DVS+CNN and hybrid methods is shown in Tab.5.

Table 5: Forecasting error between DVS+CNN and hybrid methods

	MAD	MAPE	SMAPE	RMSE	NRMSE
DVS+CNN	22.8781	0.2467	0.2468	28.4596	84.5042
ETS+ANN	25.0547	0.2719	0.2715	33.8315	99.7235
ARIMA+ANN	29.7483	0.3226	0.3221	40.4425	119.2806

By comparing previous studies, statistical methods, machine learning methods and statistical methods, the superiority of DVS+CNN has been demonstrated. In order to rigorously further illustrate the superiority of DVS, the experimental part includes ablation experiment. In the ablation experiment, DVS and LSTM, CNN and ANN are separated, and the effect of the DVS module is proved by comparing the experimental errors between the DVS module and the non-DVS module. In the experiment, the model parameters of ANN [27] , CNN [28] and LSTM [29] are shown in Tab.6.

Fig.13 shows the predicted value of ablation experiment value and the actual value of CCI. Among them, because the LSTM network has a large prediction error when it is trained for 100 epochs, the images of LSTM and DVS+LSTM are not drawn in Fig.13. The specific forecasting error in the ablation experiment is shown in Tab.5. The error of LSTM is much higher than other neural network prediction errors because 100 epoch is not enough for the training volume of LSTM network, and the training efficiency of LSTM is lower than that of ANN and CNN. Comparing the ANN and CNN using the DVS module with the ANN and CNN without the DVS module, the prediction error of the neural network using the DVS module

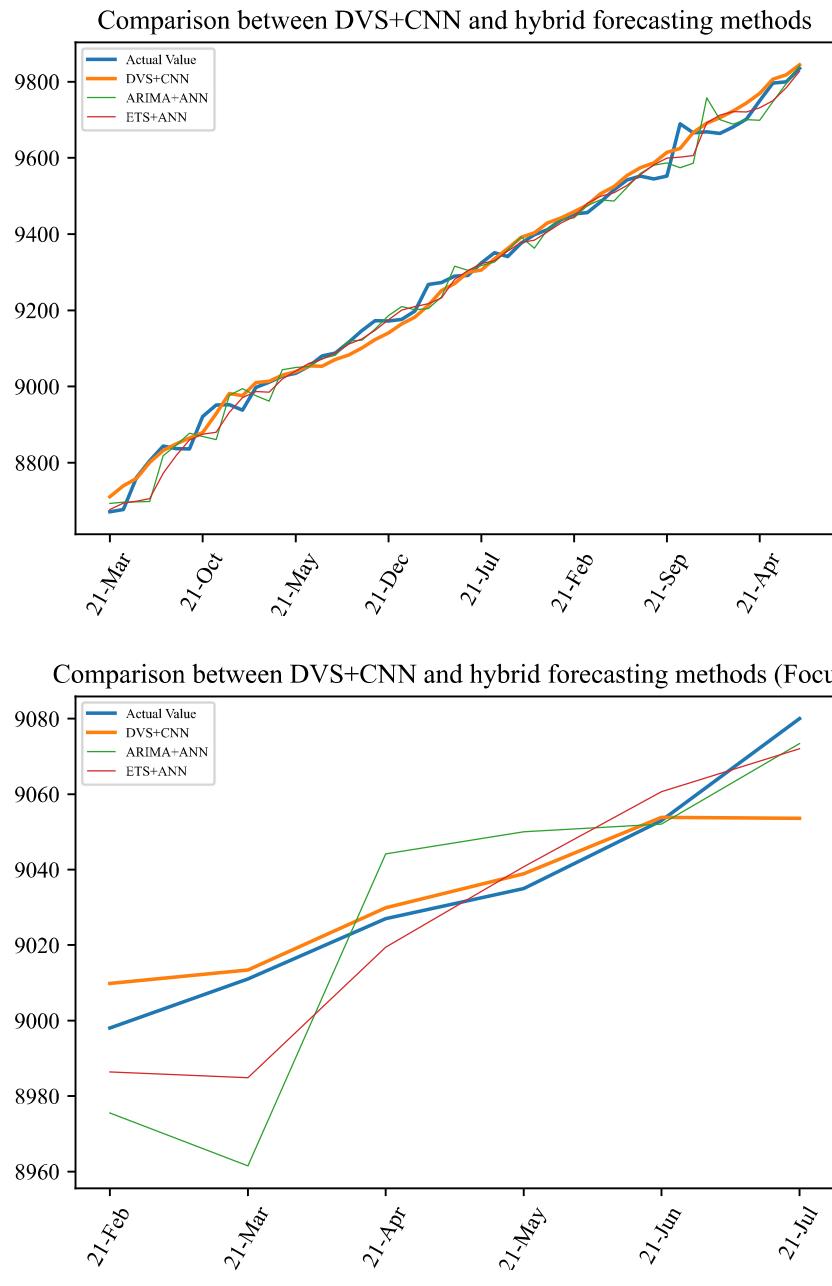


Figure 12: DVS+CNN forecasting curve and comparison method forecasting curve

Table 6: Comparison of method parameters for ablation experiments

	Structure	Training Iteration	Learning Rate	Loss Function	Optimizer Function
ANN	Hidden Layer=100 Activation=Relu Output Layer=1	100	0.01	MSE	Adam
LSTM	Hidden Layer=50 Activation=Relu Input Timestep=3 Output Timestep=1 Dense Layer=1	100	0.01	MSE	Adam
CNN	Convolution Layer: Filters=64 Kernel Size=2 Activation=Relu Pooling Layer: Pool Size=2 Dense Layer=100 Dense Layer=1	100	0.01	MSE	Adam

is further reduced. Among them, DVS has the greatest improvement in the prediction accuracy of CNN networks. DVS+CNN has the best effect among all the comparison methods.

Table 7: Forecasting error in the ablation experiments (DVS+LSTM has a large prediction error in small training epochs, which is indicated by “—”)

	MAD	MAPE	SMAPE	RMSE	NRMSE
DVS+CNN	22.8781	0.2467	0.2468	28.4596	84.5042
DVS+ANN	24.6384	0.2674	0.2672	29.821	88.2419
ANN	26.3411	0.286	0.2854	35.0024	104.294
CNN	30.7757	0.334	0.3332	40.1989	119.5102
LSTM	376.1784	41944.6893	7.8407	1796.6998	1814.6961
DVS+LSTM	-	-	-	-	-

4.4. Analysis

4.4.1. Comparative experiment analysis

This experiment contains a total of six sub-experiments. The comparison between the design of DVS+CNN and the previous research shows that DVS+CNN has greatly improved the previous research. The comparison of design statistical methods and machine learning methods with DVS+CNN shows that DVS+CNN has a good forecasting effect, which is better than traditional methods. The hybrid method is an improved method based on statistical methods. By comparing different improvement methods, the rationality of the improvement of DVS+CNN is explained. Finally, an ablation experiment is set up, which proves that the progress of DVS+CNN is due to the DVS module, which once again proves the superiority of the DVS module.

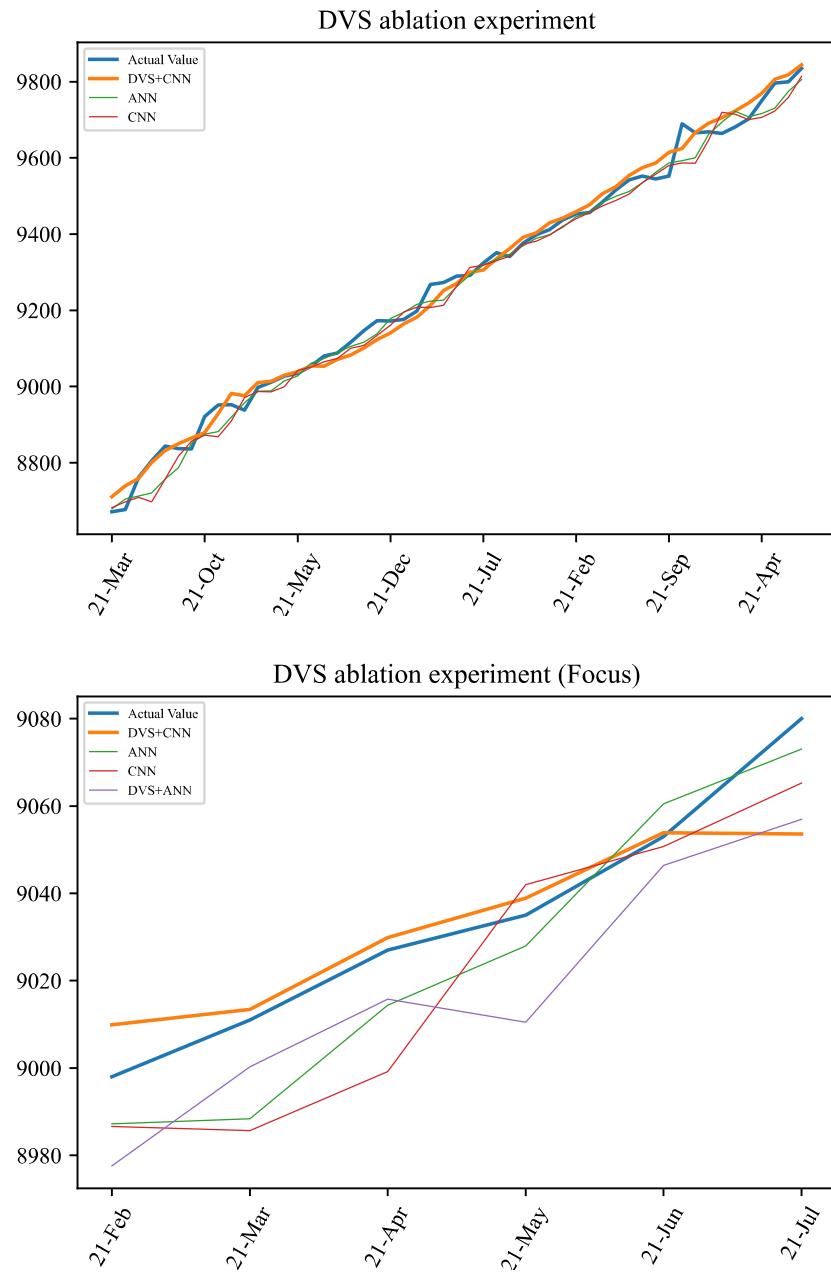


Figure 13: DVS+CNN forecasting curve and comparison method forecasting curve

4.4.2. Deep Visibility Series superiority analysis

The third section in the article mentioned how DVS is designed. Like most neural network models, the design of DVS is also a bionic design. Although the original VG and random walk algorithm utilize the visible feature of nodes, in the end only appropriate nodes are selected for weighting, without involving the specific nature of the network, and it uses less potential information. DVS strengthens the information carried by the visible view through EVG, and uses the principle of biological vision to save more network information in the time series. The time complexity of DVS is $O(n^2)$ and the time complexity is equal to that of a simple neural network. DVS cooperates with CNN to extract effective network information and time series information. The neural network forecasting method without DVS module does not cover the potential network information when forecasting the time series leading to larger forecasting error.

5. Conclusion

In order to improve the previous research and optimize the forecasting method based on VG, the DVS module is proposed, combining the VG algorithm with CNN. The EVG module in the DVS solves the problem of incomplete utilization of the original VG and time series forecasting information, greatly improving the forecasting efficiency and reducing the forecasting error. At the same time, DVS prediction has a good forecasting effect on the CCI data set, which has contributed to CCI prediction. In future work, DVS will continue to improve by combining bionic design and using biological vision principles.

Authorship contributions

Tianxiang Zhan: Conceptualization, Methodology, Software, Writing – original draft.

Yuanpeng He: Visualization, Writing – review & editing.

Hanwen Li: Visualization, Writing – review & editing.

Fuyuan Xiao: Writing – review & editing, Project administration, Funding acquisition, Supervision.

Declarations

Conflict of interest Authors Tianxiang Zhan, Yuanpeng He, Fuyuan Xiao declare that they have no conflict of interest.

Funding This study was funded by National Natural Science Foundation of China (No. 62003280) and Chongqing Talents: Exceptional Young Talents Project (CQYC202105031).

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

- [1] F.-M. Tseng, H.-C. Yu, G.-H. Tzeng, Combining neural network model with seasonal time series arima model, *Technological forecasting and social change* 69 (1) (2002) 71–87.
- [2] R. Zhang, B. Ashuri, Y. Deng, A novel method for forecasting time series based on fuzzy logic and visibility graph, *Advances in Data Analysis and Classification* 11 (4) (2017) 759–783.
- [3] P. Xu, R. Zhang, Y. Deng, A novel visibility graph transformation of time series into weighted networks, *Chaos, Solitons & Fractals* 117 (2018) 201–208.
- [4] S. Mao, F. Xiao, Time series forecasting based on complex network analysis, *IEEE Access* 7 (2019) 40220–40229.
- [5] S. M. Shahandashti, B. Ashuri, Forecasting engineering news-record construction cost index using multivariate time series models, *Journal of Construction Engineering and Management* 139 (9) (2013) 1237–1243.
- [6] S. Hwang, Time series models for forecasting construction costs using time series indexes, *Journal of Construction Engineering and Management* 137 (9) (2011) 656–662.
- [7] L. Lacasa, B. Luque, F. Ballesteros, J. Luque, J. C. Nuno, From time series to complex networks: The visibility graph, *Proceedings of the National Academy of Sciences* 105 (13) (2008) 4972–4975.
- [8] W. Liu, L. Lü, Link prediction based on local random walk, *EPL (Europhysics Letters)* 89 (5) (2010) 58007.
- [9] Y. Huang, X. Mao, Y. Deng, Natural visibility encoding for time series and its application in stock trend prediction, *Knowledge-Based Systems* 232 (2021) 107478.
- [10] S. Yitzhaki, et al., Gini's mean difference: A superior measure of variability for non-normal distributions, *Metron* 61 (2) (2003) 285–316.
- [11] A. De Myttenaere, B. Golden, B. Le Grand, F. Rossi, Mean absolute percentage error for regression models, *Neurocomputing* 192 (2016) 38–48.
- [12] C. Tofallis, A better measure of relative prediction accuracy for model selection and model estimation, *Journal of the Operational Research Society* 66 (8) (2015) 1352–1362.
- [13] R. J. Hyndman, A. B. Koehler, Another look at measures of forecast accuracy, *International journal of forecasting* 22 (4) (2006) 679–688.
- [14] M. V. Shcherbakov, A. Brebels, N. L. Shcherbakova, A. P. Tyukov, T. A. Janovsky, V. A. Kamaev, et al., A survey of forecast error measures, *World Applied Sciences Journal* 24 (24) (2013) 171–176.
- [15] S. Guan, A. Zhao, A two-factor autoregressive moving average model based on fuzzy fluctuation logical relationships, *Symmetry* 9 (10) (2017) 207.
- [16] F.-M. Tseng, G.-H. Tzeng, et al., A fuzzy seasonal arima model for forecasting, *Fuzzy Sets and Systems* 126 (3) (2002) 367–376.
- [17] B. Billah, M. L. King, R. D. Snyder, A. B. Koehler, Exponential smoothing model selection for forecasting, *International journal of forecasting* 22 (2) (2006) 239–247.
- [18] T. Hastie, R. Tibshirani, J. Friedman, The elements of statistical learnin, Cited on (2009) 33.
- [19] G. D. Hutcheson, Ordinary least-squares regression, L. Moutinho and GD Hutcheson, *The SAGE dictionary of quantitative management research* (2011) 224–228.
- [20] J. Taylor, R. Lockhart, R. J. Tibshirani, R. Tibshirani, Post-selection adaptive inference for least angle regression and the lasso, *arXiv preprint arXiv:1401.3889* 354 (2014).
- [21] H. Tong, D.-R. Chen, L. Peng, Analysis of support vector machines regression, *Foundations of Computational Mathematics* 9 (2) (2009) 243–257.
- [22] W. Xu, X. Liu, F. Leng, W. Li, Blood-based multi-tissue gene expression inference with bayesian ridge regression, *Bioinformatics* 36 (12) (2020) 3788–3794.
- [23] D. W. Hosmer Jr, S. Lemeshow, R. X. Sturdivant, *Applied logistic regression*, Vol. 398, John Wiley & Sons, 2013.

- [24] A. Defazio, F. Bach, S. Lacoste-Julien, Saga: A fast incremental gradient method with support for non-strongly convex composite objectives, in: Advances in neural information processing systems, 2014, pp. 1646–1654.
- [25] C. N. Babu, B. E. Reddy, A moving-average filter based hybrid arima-ann model for forecasting time series data, Applied Soft Computing 23 (2014) 27–38.
- [26] S. Panigrahi, H. S. Behera, A hybrid ets-ann model for time series forecasting, Engineering applications of artificial intelligence 66 (2017) 49–59.
- [27] P. H. Borghi, O. Zakordonets, J. P. Teixeira, A covid-19 time series forecasting model based on mlp ann, Procedia Computer Science 181 (2021) 940–947.
- [28] Z. Zeng, H. Xiao, X. Zhang, Self cnn-based time series stream forecasting, Electronics Letters 52 (22) (2016) 1857–1858.
- [29] J. Cao, Z. Li, J. Li, Financial time series forecasting model based on ceemdan and lstm, Physica A: Statistical Mechanics and its Applications 519 (2019) 127–139.