Journal of Network Intelligence ©2024 ISSN 2414-8105
Taiwan Ubiquitous Information Volume **9**, Number **4**, November **2024**

Cross-border E-commerce Economic Forecasting Based on Empirical Modal Decomposition and Graph Neural Networks

Abuduaini Abudureheman*

Guangdong University of Finance and Economics, Guangzhou 510320, P. R. China
18372075571@163.com

Guang-Ming Li

College of Arts and Humanities
Guangxi Vocational University of Agriculture, Nanning 530000, P. R. China
tgmlee@jnu.edu.cn

Wei Cao

College of Arts and Humanities
Guangxi Vocational University of Agriculture, Nanning 530000, P. R. China
caowei_1122@163.com

Feng Wang

College of Information Science
Mapúa University, Intramuros, 1002 Metro Manila, Philippines
uq3086@163.com

*Corresponding author: Abuduaini Abudureheman
Received January 2, 2024, revised May 9, 2024, accepted August 29, 2024.

Abstract. *As a novel mode of global commerce, cross-border e-commerce (CBEC) breaks the barriers between countries and regions, and has the advantages of easy transaction and cost saving. In this paper, a CBEC economic forecasting model (CEEF) on the ground of Empirical Modal Decomposition (EMD) and Graph Neural Network (GNN) is designed with the economic cost of CBEC as forecasting objective. Firstly, to address the issue of modal aliasing in traditional EMD algorithms, noise reduction is applied to the original signal before decomposition, and the screening criteria and iteration number are changed to reduce the decomposition error. Secondly, on the ground of the existing research, the indicators of the economic cost influencing factors of CBEC are selected, and the improved EMD algorithm is used to decompose the selected indicators to reduce the data dimensions, and then the features of the decomposed indicators are correlated by the Graph-Attention Neural Network (GANN), and the intrinsic features of the indicators are captured by the Graph-Convolutional Neural Network (GCNN), and the residuals of the EMD decomposition are combined with the reconstructed results, so that the final prediction results are outputted, thus improving the competence of the prediction. Thus, the competence of prediction is improved. Simulation results show that compared with the existing models, the CEEF method has lower Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE), the CEEF method is with higher forecasting accuracy.*
Keywords: Cross-border e-commerce; Economic forecasting; Empirical modal decomposition; Graph attention networks; Graph convolutional networks

---

1. **Introduction.** With the persistent growth of network technology and rapid progress of information services, e-commerce enterprises are emerging all over the world, and more and more offline physical enterprises are beginning to join, making the competition become more and more intense. Enhancing the competitiveness and core strength of enterprises and utilizing advanced machine learning technology make the economic cost manage of Cross-Border E-Commerce (CBEC) enterprises more scientific and efficient [1, 2]. Machine learning technology brings new opportunities and development to e-commerce enterprises, and is fully utilized in enterprise cost control, which can become an intangible wealth of the enterprise [3]. The wide application of diversified machine learning technology avoids many problems such as small information storage and insufficient analyzing ability that existed in the traditional data analysis in the past. On the basis of the rapid growth of the world economy, the e-commerce industry has gained certain development results. The development of e-commerce industry and logistics industry are nearly associated, and the operation and growth of cross-border e-commerce is inevitably dependent on logistics industry [4, 5]. The development speed of CBEC in the world has been increasing, however, under the joint effect of a lot of factors for instance surroundings and enterprise growth, CBEC enterprises yet cannot reasonably control economic costs.

The research value of CBEC economic forecast lies in providing decision-making basis and strategic guidance for stakeholders such as government, enterprises and investors. Through in-depth analysis of cross-border e-commerce market trends, consumer behavior, technological innovation and other factors, we can reveal the market development potential and challenges, and help to predict market size and growth rate and other key indicators. Machine learning has powerful capabilities in data processing, information analysis and forecasting modeling, which can effectively integrate multi-dimensional information such as big data, consumer behavior and competition in cross-border e-commerce market, conduct intelligent learning and accurate forecasting, improve the accuracy and accuracy of forecasting, and reduce the cost and time of forecasting.

1.1. **Related Work.** Fan et al. [6] based on the empirical analysis of online reviews by Kano model, identified the service indicators related to CBEC's expectation factors. He et al. [7] constructed a CBEC model by combining intelligent logistics communication technology, and proposed a uniform power-saving method on the ground of 3D metacellular universe, which balanced the power expenditure of logistics information unification and transmittance. Qi [8] constructed a CBEC enterprises' international competitiveness evaluation index system, and determined the indexes and the indexes' weights through the hierarchical analysis method. Yuan [9] used particle swarm algorithm to deal with the issue of CBEC using particle swarm algorithm for solving, verified the method's feasibleness, for the purpose of conferring a particular plan for the position of overseas warehouses. Luo et al. [10] combined Bayesian individualized classifying algorithm and pairwise reciprocity tensor putrefaction algorithm, constructed the sword orientation forecasting method, and predicted the sword orientation of the CBEC consumers. Giuffrida et al. [11] based on the system construction requirements of the B2C CBEC company. system construction requirements to construct and adjust the B2C cross-border e-commerce logistics system. Li and Li [12] used a deep forest classifier instead of the traditional neural network classifier to accomplish the task of commodity name recognition. It is able to recognize e-commerce trade names effectively and accurately.

Subsequently, Guo [13] raised problems by predicting the e-commerce user behavior process through a comprehensive analysis of cases, and gave corresponding suggestions. Abdulkarem and Hou [14] used the technology, environment, and organization (TOE) framework to study the effect of the environment on the four layers of CBEC adoption by small and medium-sized firms. Wang and Gao [15] offered a prediction method on the ground of support vector machine and BP neural network model that helps in decision making for the company. Lei and Qiu [16] optimized the traditional BP algorithm based on the model of BP algorithm and applied it to the prediction of goods purchasing intensity of a CBEC platform, which has high practical application value. Thanks to the growth of Graph Neural Network (GNN) technology, scholars have already applied it to the field of CBEC forecasting. Ren et al. [17] takes user behaviors as nodes and the frequency of product purchases as edge weights, constructs a relationship graph, and learns the representation using a graph attention network. Feng and Chen [18] propose an optimization-based GANN model. The model is able to optimize the window width and parameters, but the accuracy is not high. Chen and Wang [19] used a neural network model associated with the Spark cloud calculating platform to achieve brilliant forecasting of CBEC logistics cost, but the prediction accuracy is low. Chang et al. [20] conducted a comprehensive estimation of CBEC business models on the ground of an evaluation index system. Rui [21] established 17 performance evaluation indicators to play a positive part in CBEC by establishing an efficient and low-risk supply chain. Li et al. [22] summarized a deep neural network method advantageous to the growth of CBEC. Zhong et al. [23] presented a deep neural network way for the development of CBEC on the ground of BP neural network and SVM for CBEC, but the forecasting effect is not good. Since the time series generated in CBEC economic forecasting are usually nonlinear and complex data, traditional forecasting methods could not effectively capture the local instantaneous information of the series during forecasting, thus affecting the forecasting accuracy and stability. To effectively solve this problem, Zhang et al. [24] used Empirical Modal Decomposition (EMD) to decompose the CBEC economic cost series, and then the sequences obtained after decomposition were predicted using Artificial Neural Network (ANN) respectively, which improved the prediction accuracy and provided a novel and effective method for CBEC prediction. However, the method does not solve the modal aliasing problem of the EMD algorithm.

1.2. **Contribution.** Currently, the research on the simultaneous application of EMD and GNN to CBEC economic forecasting is on the low side, and the results of the forecasting are not satisfactory. In this paper, the EMD algorithm is used to process the nonsmooth data in CBEC economic forecasting to reduce the complexity of the INDEX series, and then the data are passed into the GNN to make predictions. Firstly, the original signal is processed with noise reduction before EMD decomposition utilizing the mirror extension method and wavelet threshold denoising method, and secondly, the impact indicators affecting the economic cost of CBEC are analyzed and selected, and after decomposition of the improved EMD algorithm, the information of the impact indicators is transmitted through the Graph Attention Neural Network (GANN), and the information of the neighboring nodes of different importance is adaptively converged to enrich the feature representation of economic cost., enriching the feature representation of economic cost. Then the Graph Convolution Neural Network (GCNN) is used to extract the intrinsic features of each indicator, and the final vector of predicted economic costs is obtained by activation function. Finally, the simulation outcome indicates that the CEEF model has high precision, recall, precision, and F1 value in CBEC economic cost, and can efficiently realize the economic cost prediction.

2. Theoretical Analysis.

2.1. **Graphical Convolutional Neural Network.** Graph Convolutional Neural Network (GCNN) [24] is to apply convolutional operations to graph structures. As shown in

Figure 1, each convolutional computation fuses the information of the node's neighbors once, which is referred to as "one hop" in graph convolutional neural networks. For each red graph node, the graph convolutional neural network expects it to fuse the information of the surrounding blue nodes. As a result, each node fuses the information of its neighbors over N hops, and finally, the features of these nodes are average pooled or maximally pooled to facilitate subsequent prediction. GCNN extracts features of graph structures by transforming to the spectral domain for convolution operations via Fourier transform and finally mapping it back to the null domain [25]. The normalized Laplace matrix feature decomposition is as follows.

$$y = V h_\vartheta(\Lambda) V^T x \tag{1}$$

$$L = V(\Lambda) V^T \tag{2}$$

where x is the stimulus, y is the output, L is the Laplacian matrix of the picture, V is the eigenmatrix of the Laplacian matrix of the picture, V^T is its transpose, Λ is the solidus matrix making up of the eigenvectors of the matrix, and $h_\vartheta(\cdot)$ is the frequency response function.

Let the input variable of GCNN be denoted as $x \in \mathbb{R}^M$, and perform the Fourier transform of the graph: $F(x) = V^T x$. Its Fourier inverse transform is $F^{-1}(x) = Vx$. The transformed input signal is convolved with the convolution kernel $h \in \mathbb{R}^M$.

$$(x * h) = F^{-1}(F(x) \odot F(h)) = V(V^T x \odot V^T h) = V h_\vartheta V^T x \tag{3}$$

where h_g is a diagonal matrix that is a function acting on the Laplace matrix identity.

$$h_\vartheta = \begin{pmatrix} \hat{h}(\lambda_1) & .... & 0 \\ ... & ... & ... \\ 0 & ... & \hat{h}(\lambda_n) \end{pmatrix} \tag{4}$$

Finally, the output of its model is represented as follows:

$$y = \delta(V h_\vartheta V^T x) \tag{5}$$

where h_ϑ is the parameter on which the model needs to be trained and δ is the nonlinear activation function.

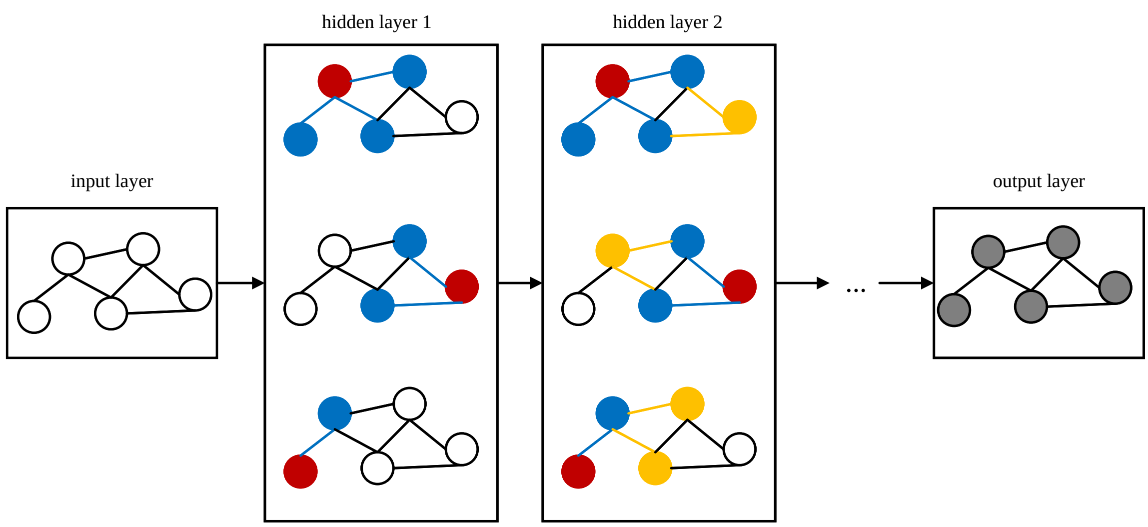

Figure 1. Graphical convolutional neural network

Economic data usually have complex correlation and interaction. Compared with BP neural network, GCN can capture the graph structure of economic data, and learn the characteristics of nodes and edges to better understand the relationship between variables in the economic system. This is helpful to improve the accuracy and explanatory ability of the prediction model.

2.2. **Empirical modal decomposition.** The core of empirical modal decomposition is to screen the Intrinsic Modal Function (IMF). EMD is a signal processing technique for decomposing nonstationary signals into multiple IMF. By adaptively determining the frequency and vibration mode of the eigenmode function, EMD can better capture the time-varying characteristics of the signal and has broad application potential. For non-smooth data, EMD decomposition also has defects, mainly due to the susceptibility to modal aliasing [26, 27]. The precise steps are as follows.

Step 1: Add white noise $h(s)$ to the original data $X(s)$, to get the new sequence $X'(s)$.

$$X'(s) = X(s) + h(s) \tag{6}$$

Step 2: Calculate all the upper and lower excessive points of $X'(s)$, draw the upper and lower envelopes $V(s)$ and $K(s)$, and then discover the average quality of the upper and lower envelopes to get the mean value envelope $H(s)$.

$$H(s) = \frac{V(s) + K(s)}{2} \tag{7}$$

Step 3: Compute the intermediate sequence $G(s)$.

$$G(s) = X'(s) - H(s) \tag{8}$$

Step 4: Decide if the intermediate sequence meets the situation of IMF, if so, the sequence is an IMF ingredient; otherwise, adopt the sequence as the foundation and re-do the analysis in step 2 to step 4 until the conditions are met.

Step 5: After obtaining the first IMF in the above steps, subtract the IMF component from $X'(s)$ and take it as a new sequence, and step 2 to step 4 to obtain m IMFs and one $R_m(s)$.

$$X'(s) = \sum_{j=1}^{m} IMF_j + R_m(s) \tag{9}$$

where the residual $R_m(s)$ is obtained by subtracting IMFn from the original signal $X(s)$.

Step 6: Repeat the noise addition process for M times, i.e., repeat step 1 to step 5 for M times.

Step 7: The final IMF component $J_i(s)$ is obtained by averaging the IMF components of the same ordinal number obtained from each of the above decompositions.

$$J_i(s) = \frac{1}{N} \sum_{j=1}^{N} J_{i,j}(s), \quad 1 \le i \le n+1 \tag{10}$$

3. **Optimization of empirical modal decomposition algorithms.** Modal aliasing in traditional EMD algorithm will lead to information overlapping and confusion between different IMFs. Such confusion may make it difficult for the prediction model to accurately capture the characteristics and changing trends of each mode, and make the prediction of long-term trends and short-term fluctuations of e-commerce economy unreliable. EMD algorithm produces modal aliasing can be summarized as the following two main reasons: one is due to the presence of noise in the signal, the original distribution of local extreme value points of the signal caused by interference; the second is the signal is accompanied

by a dense frequency, intermittent high-frequency weak signals, and other factors interfere. For the former, we can perform noise reduction on the original signal before EMD decomposition, while the latter can be achieved by changing the screening criteria and the number of iterations, signal frequency modulation and other methods. The overall optimization process is shown in Figure 2.

(1) Noise reduction processing of the initial signal. When the initial signal is wavelet decomposed, the wavelet decomposition coefficient of the effective signal will be significantly larger than the wavelet decomposition coefficient of the noise. Assuming that the wavelet transform is performed on the discrete signal $X(m)(m = 1, 2, ..., M)$.

$$HX(j,l) = 2^{-j/2} \sum_{m=1}^{M} X(m)\phi\left(2^{-j}m - l\right) \tag{11}$$

where $HX(j,l)$ is the wavelet decomposition coefficient (can be abbreviated as H_{jl}), j is the decomposition scale, and l is the position.

(2) Change the screening criterion and the amount of iterations. The iterative screening of the IMF component affects the decomposition effect of the EMD algorithm. Usually, the EMD decomposition determines whether the screening of an IMF is completed or not by the standard deviation SC.

$$SC = \sum_{m=0}^{M} \left[\frac{p_{jl}(m) - p_{j(l-1)}(m)}{p_{j(l-1)}(m)}\right]^2 < SC' \tag{12}$$

where SC' denotes the pre-set threshold for the end of sieving; M denotes the overall number of data; $p_{jl}(m)$ denotes the amplitude of sieving for the j-th IMF to the l-th time at data point m. Sieving ends when $SC < SC'$. The value of SC is related to the length of the signal being decomposed, and more iterations will be required to get the SC that satisfies the condition if the pre-set value of SC' is large. So, consider $R_m(s)$ as the $(m+1)$-th IMF, then:

$$X(s) = \sum_{j=1}^{m+1} IMF_j(s) \tag{13}$$

Squaring both sides yields Equation (14).

$$X(s)^2 = \sum_{j=1}^{m+1} IMF_j^2(s) + 2\sum_{j=1}^{m+1}\sum_{i=1}^{j-1}\left[IMF_j(s) \times IMF_i(s)\right] \tag{14}$$

The number of iterative screening of the eigenmodes has an impact on the EMD modal frequency resolution. For the purpose of quantitatively representing the modal resolution of two single-frequency combined signal models, it is defined as follows:

$$\tau(f) = \frac{\|IMF_1(s; f_1) - \cos(2\pi f_1 s)\|}{\|\cos 2\pi f_2 s\|} \tag{15}$$

where $\tau(f)$ is the modal decomposition error; $IMF_1(s; f_1)$ is the first IMF component of the signal model; $\|x\|$ is the number of paradigms in which the variable x is taken to be 2. It is usually assumed that the EMD separates the two dense modes when the decomposition error $\tau < 0.1$ is taken.

4. Cross-border e-commerce economic forecasting based on empirical modal decomposition and graph neural networks.

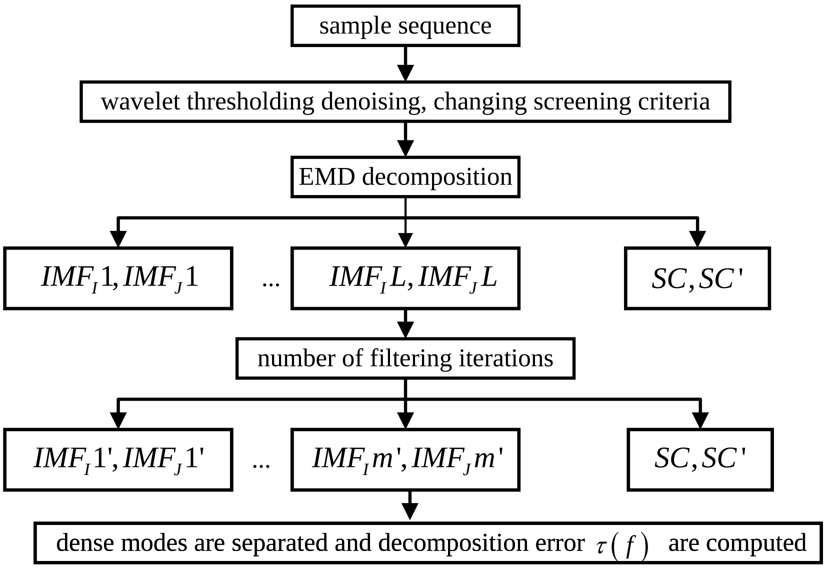

Figure 2. The enhanced process of EMD

4.1. **Cross-border e-commerce impact indicator selection and empirical modal decomposition.** This paper proposes a model of CBEC economic forecasting on the ground of EMD and GNN, and the overall framework is indicated in Figure 3. The model is chiefly composed of cross-border e-commerce impact indicator selection module, EMD decomposition module, GANN-based indicator feature data association module, GCN-based indicator intrinsic feature representation module and multi-classification output module.

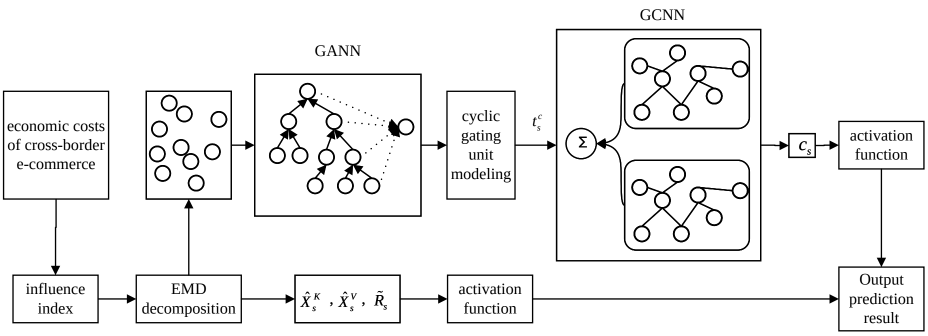

Figure 3. The overall framework of CEEF

Based on the existing research, it can be concluded that the indicators are selected from the three aspects of economic level, logistics level and platform construction level, and associated with the existing situation of CBEC growth and the four elements of development, "business flow, logistics, information flow and capital flow" [26], the indicators of the influencing factors are selected: Gross National Income, GDP per capita, per capita disposable income, mobile Internet users, total postal business, express delivery, CBEC enterprise registration, CBEC financing, CBEC user scale, and the number of CBEC investment are ten influencing indicators. Gross national income, per capita gross domestic product, per capita disposable income, mobile Internet users, total postal business, express delivery, CBEC-related enterprise registration, CBEC financing, CBEC user scale, CBEC investment number of the ten impact indicators, respectively, are recorded as follows: $X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_9, X_{10}$, and the economic cost of CBEC is still recorded as X_0.

An empirical modal decomposition is then performed for the above selected indicators.

(1) Decomposition of cross-border e-commerce impact indicators. Firstly, the modal aliasing of the original indicator interval data X_i (where $i = 1, 2, \ldots, 10$) are suppressed by using the optimized EMD algorithm to obtain the sequence $\tilde{X}_i$. Secondly, the upper and lower bounds of the interval time series $\tilde{X}_i$ are decomposed as a whole by the optimized EMD. The parts of the sequence that have been extended at the two ends of the sequence are truncated to obtain the decomposition of the l IMF inter-area components and one residual inter-area component, as shown as follows:

$$\tilde{H}_i(l) = \left[H_i^K(l), H_i^V(l)\right], l = 1, 2, \ldots, m \tag{16}$$

$$\tilde{R}_i = \left[R_i^K, R_i^V\right] \tag{17}$$

The purpose of the above is to reduce the boundary effects and improve the accuracy and reliability of the decomposition results, as well as to control the number and frequency range of the decomposition results, which will help in better analysis and prediction of the signal.

(2) Reconstruction of IMF interval components and residual components. Using the sample entropy to measure the complexity of each IMF interval component, the original IMF interval components are reorganized into high-frequency interval components, medium-frequency interval components and low-frequency interval components according to the value of the sample entropy of each component sequence, and m new IMF interval component sequences and one residual component are obtained.

$$\tilde{H}_i(m) = \left[H_i^K(m), H_i^V(m)\right] = \left[\sum_{j=1}^{n} h_j(m) H_{ij}^K(m), \sum_{j=1}^{n} h_j(m) H_{ij}^V(m)\right] \tag{18}$$

$$\tilde{R}_i = \left[R_i^K, R_i^V\right] = \left[\sum_{j=1}^{n} h_j R_{ij}^K, \sum_{j=1}^{n} h_j R_{ij}^V\right] \tag{19}$$

where h is the noise component.

4.2. **Data association of cross-border e-commerce indicator features based on GANN.** The architecture of the CBEC economic cost and the impact indicators of improved EMD decomposition can be described using a tree-like classification structure. As shown in Figure 4, the information of each node is passed through a specially designed graph attention network, which adaptively aggregates the information of its neighboring nodes of different importance to enrich the characteristic representation of the economic cost. The final output contains the coded representation of the impact indicators, which serves as the fundament for the prediction of CBEC economic costs.

First for non-leaf nodes, adaptively learn its correlation with each of its child nodes and weighted sum to update its association.

$$r_j = \prod_{l=1}^{L} \delta \left(\sum_{i \in d(i) \cup j} \beta_{ij}^l r_i \right) \tag{20}$$

where δ is the Relu function, $d(i)$ represents all the child nodes of node i; and C represents the attention weight coefficient of node j to its parent node i under the lth attention.

$$\beta_{ij}^{l} = \frac{\exp\left(\text{LeakyReLU}\left(b^T\left[r_j \parallel r_j\right]\right)\right)}{\sum_{i\in d(i)\cup j}\exp\left(\text{LeakyReLU}\left(b^T\left[r_j \parallel r_j\right]\right)\right)} \tag{21}$$

where $b \in \mathbb{R}^{2\times 64}$ is a single-layer feed-forward neural network; $\parallel$ denotes the connection of matrices; and the final attention coefficients are obtained after normalization by the Relu activation function.

A sequence of feature vectors of economic costs and their impact metrics is then input to the gated unit loop unit to capture the expected economic benefits. Attention is then adopted to decide the impact of the obscured level output information of the historical sequence with respect to the current metric representation. β^c is used to represent the weight coefficient of the economic cost under attention β.

$$\beta^c = \text{softmax}\left(F_\beta\left(h_1\right), F_\beta\left(h_2\right), ..., F_\beta\left(h_n\right)\right) \tag{22}$$

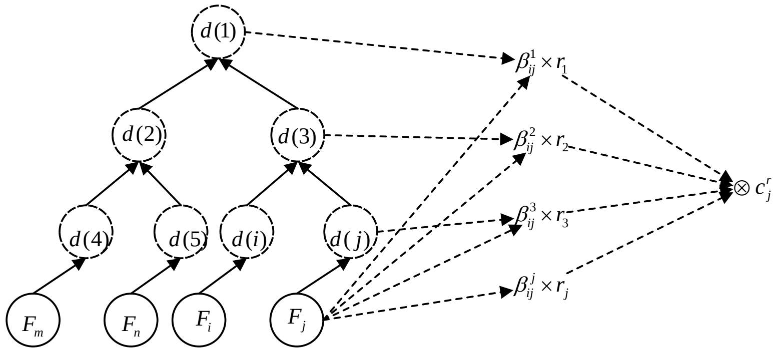

Figure 4. Indicator Characterization Data Correlation Chart

where F_β denotes the linear transformation function (weight vector dimension $\mathbb{R}^{64\times 1}$); h_s is the data output from the obscured level. The time series encoding the impact indicators are input into the gated loop unit G_α. The *tanh* operation is chosen as the activation function to obtain the attention weights α^c.

$$g_1, g_2, ..., g_s = G_\alpha(c_1^r, c_2^r, ..., c_s^r) \tag{23}$$

$$\alpha^c[j] = \tanh(F_\alpha(g_j)), j = 1, 2, ..., s \tag{24}$$

where $\alpha^c[j]$ denotes the weight coefficient of the j-th input and g_j is the hidden layer output. The representation of the economic cost time series relationship with cross-border e-commerce is obtained as follows.

$$t_s^c = \sum_{i=1}^{s}\beta^c[i]\alpha^c[i]\otimes c_i^r \tag{25}$$

where $\otimes$ denotes the element-by-element multiplication operation, $t_s^c \in \mathbb{R}^{64\times 1}$.

4.3. **Cross-border e-commerce economic forecasting based on EMD and GAN.** To accurately forecast the economic cost of CBEC, graph convolutional networks are used to capture the intrinsic features of each indicator. For convenient representation, in this paper, $B_x \in \mathbb{R}^{N_x\times N_m}$ is used to represent the interaction adjacency matrix B_D and co-occurrence adjacency matrix B_C of each indicator, and A_* is able to be transformed as $\tilde{A}_* = C^{-1}(A_* + I)C^{-1}$ where D is a diagonal matrix, I is a unit matrix. A_* is a symmetric normalized Laplace matrix. The intrinsic characteristics among the indicators are as follows:

$$Y_D = \tilde{B}_D \tanh(\tilde{B}_D m' \tilde{H}_D^1)\tilde{H}_D^2 \tag{26}$$

$$Y_C = \tilde{B}_C \tanh(\tilde{B}_C m' \tilde{H}_C^1)\tilde{H}_C^2 \tag{27}$$

where H denotes the IMF interval component after EMD decomposition; Y_D and Y_C denote the co-occurrence and intrinsic characteristics of the indicators, respectively. The attention coefficient μ_t is then obtained by combining the residual component $\tilde{R}_t$ of the EMD decomposition.

$$\mu_s = \text{soft max}\left((Y_D + hY_C)^T \cdot \tilde{R}_s\right) \tag{28}$$

After summation of attention weights to obtain the intrinsic eigenvector between the indicators c_s.

$$c_s = (Y_D + hY_C) \cdot \mu_s \tag{29}$$

After activation by the ReLU function, the final vector of predicted economic costs $\hat{X}_s$ is obtained.

$$\hat{X}_s = \vartheta\left[\hat{X}_s^D, \hat{X}_s^Y, \tilde{R}_s, t_s^c, c_s\right] \tag{30}$$

This paper uses binary cross-entropy loss function to measure the quality of economic cost prediction of cross-border e-commerce.

$$L_s = -\sum_s \sum_i V^j \left(\log\vartheta\left(\hat{X}_s^j\right) + (1 - \hat{X}_s^j)\log\vartheta\left(1 - \hat{X}_s^j\right)\right) \tag{31}$$

At the same time, in order to consider the internal characteristics of the indicators affecting the economic cost of CBEC, its loss function is as follows.

$$L_I = \sum_s \sum_{i,j} \left(B_C \odot \left(\hat{X}_s^T \hat{X}_s\right)\right)[i,j] \tag{32}$$

At last, the two loss functions are combined to obtain the ultimate joint loss function.

$$L_{\text{com}} = \lambda_1 L_s + \lambda_2 L_I \quad (\lambda_1 + \lambda_2 = 1) \tag{33}$$

5. Performance test and analysis.

5.1. **Comparison of economic cost prediction results of cross-border e-commerce with different models.** For the purpose of estimating the performance of the CBEC economic forecasting model on the ground of empirical mode decomposition and graph neural network, the CEEF forecasting method designed in this paper, the PCAP method proposed in literature [14] and the AMFC method proposed in literature [28] were established by selecting 1000 economic cost data of a large CBEC platform's 10 warehouses from 2012 to 2022 and relevant data of influencing factors. The prediction results of these three methods are compared. In the experiment, Python was chosen to write the script code and the implementation code was written in the Windows 10 operating system. Deep learning models use the TensorFlow library [29] as the back end, while machine learning models use the Scikit-learn library. The specific parameters of the experiment are indicated in Table 1.

Table 1. The specific parameters of the experiment

Parameters	Value	Parameters	Value
Network layer	4	Convolution kernel	2
Convolutional stride	1	Convolutional padding	"valid"
Activation function	"Tanh", "Relu"	Dropout	0.2
Training wheels	100		

The actual economic cost data of a CBEC business from 2014 to 2020 is implied in Table 2. To estimate the effectivity of the CEEF method designed in this article, PCAP method and AMFC method are selected as comparison methods. Considering the rigor and rationality of the experiment, the parameters of these comparison methods are all set according to the model in this paper. The forecasting experiment outcome of the comparison method is indicated in Figure 5.

Table 2. Data on the real economic costs of cross-border e-commerce

Vintage	2014	2015	2016	2017	2018	2019	2020
Actual economic cost (tens of millions yuan)	6.3	9.5	11.8	16.2	22.4	29.6	35.9

As can be seen in Figure 5, there is a certain error between the predicted value and the true value of each method, but the CEEF method has the smallest error among the forecasting value and the true value. The PCAP and AMFC models have poorer prediction results, which is due to the fact that the PCAP method only utilizes the BP neural network model to make simple forecasting of the cost data and ignores the intrinsic characteristics of the impact indicators, and the AMFC method does not carry out indicator data dimensionality reduction, resulting in biased prediction results.

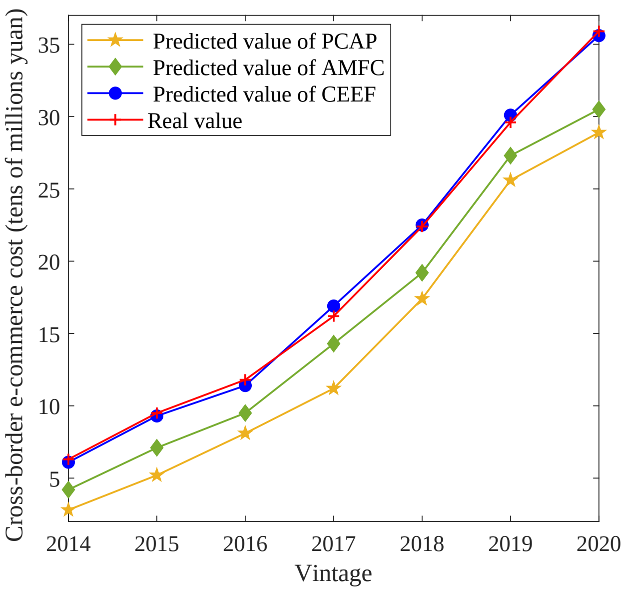

Figure 5. Economic cost forecast of cross-border e-commerce

To assess the prediction accuracy, the Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE) and Root Mean Square

Table 3. Comparison of the model's prediction results

Method	**MAE**	**MAPE**	**RMSE**
PCAP	0.01974	4.76924	0.03295
AMFC	0.01281	3.19672	0.01853
CEEF	0.00469	2.19637	0.01172

Percentage Error (RMSPE) were adopted to analyze the prediction results, as shown in Table 3.

From Table 3, CEEF method has the smallest error among the forecasting value and the real value, with the values of MAE, MAPE and RMSE of 0.00469, 2.19637 and 0.01172, respectively. Comparing this method with the best-performing comparative method AMFC, it can be found that the MAE decreases from 0.01281 to 0.00469, and the error is reduced by 63.39%; MAPE is reduced from 3.19672 to 2.19637, with an error reduction of 31.3%; RMSE is reduced from 0.01853 to 0.01172, with an error reduction of 36.8%, which fully demonstrates the effectiveness of the CEEF methodology in this paper. By comparing the two methods, it can be seen that the addition of empirical modal decomposition makes the model pay more attention to time series with different frequencies, which further improves the prediction results. The forecasting impact of the AMFC model is slightly better than that of the PCAP model, which is due to the fact that the AMFC method improves the ability of the data feature extraction through the convolution operation, which is conducive to the further learning of the potential relationships in the data, and thus enhances the prediction ability of the method. The prediction ability of AMFC method is slightly better than that of PCAP method.

5.2. **Comparison and analysis of model accuracy based on ROC curve and AUC value.** The statistical evaluation indexes used in this paper mainly include the calculated accuracy indexes, as well as the ROC curve and AUC value. As shown in Figure 6, the ROC curve is drawn with the false-positive rate as the horizontal coordinate and the T-true-positive rate as the vertical coordinate, in which the diagonal dashed line indicates the result of random prediction, and its AUC is 0.5. When the ROC curve is applied to the model evaluation, the smaller the horizontal coordinate and the bigger the vertical coordinate are, the higher the precision of the model is. In addition, the predictive performance of different models can be compared horizontally by calculating the AUC value of each model.

The AUC values of the CEEF and AMFC methods are all greater than 0.5, indicating that both methods have strong predictive ability and stability. The PCAP method has the smallest AUC value because it only uses the decision tree to establish the linkage between the impact indicators without empirical modal decomposition and feature correlation of the indicator data. The CEEF method is non-randomized in the processing of the data, and has good stability and the highest accuracy of the prediction results. The CEEF method is non-randomized, stable, robust, and has the highest accuracy in predicting the results.

In addition to visualizing the prediction ability of the comparison models through the ROC and AUC value curves, the prediction effect evaluation indexes of various methods can also be further analyzed in depth. The prediction effect evaluation indexes of each method are summarized as shown in Table 4.

As can be seen from Table 4, there is a significant difference in the effectiveness of different methods in predicting the economic costs of CBEC. In terms of combined recall and precision, the CEEF method has the strongest correct prediction ability, with an

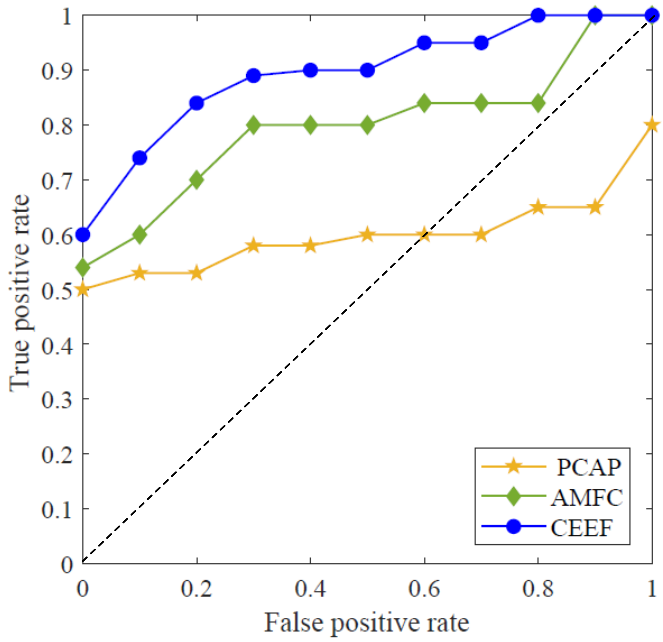

Figure 6. ROC curves and AUC values for different modelse

Table 4. Comparison of the model's prediction results

Method	AUC	Accuracy	Recall	Precision	F1-Score
PCAP	0.816	0.794	0.812	0.805	0.808
AMFC	0.859	0.843	0.872	0.853	0.862
CEEF	0.937	0.926	0.931	0.914	0.922

Accuracy of 0.926, a Recall of 0.931, a Precision of 0.914 and an F1 value of 0.922, followed by the AMFC method, with an Accuracy of 0.843, a Recall of 0.872, a Precision of 0.853, Recall is 0.872, Precision is 0.853, and F1 value is 0.862, which has some correct prediction ability, but cannot correctly find out the majority of positive samples, and is easy to predict the positive samples as negative samples. The PCAP method has the worst performance in both datasets, but all the indexes are greater than 0.5, which indicates that it has some classification ability, but has poor prediction effect.

6. **Conclusion.** Aiming at the issue of low forecasting accuracy in the existing CBEC economic prediction methods, this paper proposes a CBEC economic prediction method based on EMD and GNN. Firstly, the traditional EMD algorithm is improved, and the screening criterion and the number of iterations are changed before the EMD decomposition to reduce the decomposition error. Secondly, the influence indicators affecting the economic cost of CBEC are selected, and after these indicators are decomposed by the improved EMD algorithm, the graph attention network adaptively aggregates the information of neighboring nodes with different importance to enrich the feature representation of economic cost. Then the graph convolution network is used to further extract the intrinsic features of each indicator, and the final vector of predicted economic costs is obtained after activation by the Relu function. Finally, the performance of trade economic prediction is evaluated, and the simulation outcome indicates that the CEEF method has high forecasting accuracy in CBEC economic prediction and can be better applied to CBEC economic prediction.

Acknowledge. This work is supported by Guangzhou Huashang College 2021 Key Discipline Project: School Level Key Discipline-International Business (No. 2021HSXK05).

REFERENCES

[1] Y. Wang, Y. Wang, and S. H. Lee, "The effect of cross-border e-commerce on China's international trade: An empirical study based on transaction cost analysis," *Sustainability*, vol. 9, no. 11, 2028, 2017.

[2] Y. He, and J. WANG, "A panel analysis on the cross border e-commerce trade: Evidence from ASEAN countries," *The Journal of Asian Finance, Economics and Business*, vol. 6, no. 2, pp. 95-104, 2019.

[3] S. Ren, T.-M. Choi, K.-M. Lee, and L. Lin, "Intelligent service capacity allocation for cross-border-E-commerce related third-party-forwarding logistics operations: A deep learning approach," *Transportation Research Part E: Logistics and Transportation Review*, vol. 134, 101834, 2020.

[4] L. Guo, "Cross-border e-commerce platform for commodity automatic pricing model based on deep learning," *Electronic Commerce Research*, vol. 22, no. 1, pp. 1-20, 2022.

[5] Y. Shi, T. Wang, and L. C. Alwan, "Analytics for cross-border e-commerce: inventory risk management of an online fashion retailer," *Decision Sciences*, vol. 51, no. 6, pp. 1347-1376, 2020.

[6] M. Fan, Z. Tang, S. A. Qalati, K. Tajeddini, Q. Mao, and A. Bux, "Cross-Border E-Commerce Brand Internationalization: An Online Review Evaluation Based on Kano Model," *Sustainability*, vol. 14, no. 20, 13127, 2022.

[7] X. He, S. Meng, and J. Liang, "Analysis of cross-border E-Commerce logistics model based on embedded system and genetic algorithm," *Microprocessors and Microsystems*, vol. 82, 103827, 2021.

[8] W. Qi, "Path to Enhance the International Competitiveness of Cross-Border E-Commerce Logistics Enterprises-Research Based on AHP Analysis Method," *International Journal of Frontiers in Sociology*, vol. 4, no. 8, pp. 10-13, 2022.

[9] Q. Yuan, "The construction mechanism and algorithm of cross border E-commerce export logistics mode from the perspective of value chain," *Journal of Intelligent & Fuzzy Systems*, vol. 37, no. 3, pp. 3393-3400, 2019.

[10] Y. Luo, J. Ma, and C. Li, "Entity name recognition of cross-border e-commerce commodity titles based on TWs-LSTM," *Electronic Commerce Research*, vol. 20, pp. 405-426, 2020.

[11] M. Giuffrida, R. Mangiaracina, A. Perego, and A. Tumino, "Cross-border B2C e-commerce to Greater China and the role of logistics: a literature review," *International Journal of Physical Distribution & Logistics Management*, vol. 47, no. 9, pp. 772-795, 2017.

[12] G. Li, and N. Li, "Customs classification for cross-border e-commerce based on text-image adaptive convolutional neural network," *Electronic Commerce Research*, vol. 19, pp. 779-800, 2019.

[13] L. Guo, "Cross-border e-commerce platform for commodity automatic pricing model based on deep learning," *Electronic Commerce Research*, vol. 22, no. 1, pp. 1-20, 2022.

[14] A. Abdulkarem, and W. Hou, "The Influence of the Environment on Cross-Border E-Commerce Adoption Levels Among SMEs in China: The Mediating Role of Organizational Context," *SAGE Open*, vol. 12, no. 2, 2158, 2022.

[15] L. Wang, and X. Gao, "Construction of Cross-Border e-Commerce Financial Risk Analysis System Based on Support Vector Machine," *Journal of Industrial Integration and Management*, vol. 8, no. 01, pp. 25-38, 2023.

[16] Y. Lei, and X. Qiu, "Evaluating the investment climate for China's cross-border E-commerce: the application of back propagation neural network," *Information*, vol. 11, no. 11, 526, 2020.

[17] S. Ren, T. M. Choi, K. M. Lee, and L. Lin, "Intelligent service capacity allocation for cross-border-E-commerce related third-party-forwarding logistics operations: A deep learning approach," *Transportation Research Part E: Logistics and Transportation Review*, vol. 134, 101834, 2020.

[18] Z. Feng, and M. Chen, "Platformance-based cross-border import retail e-commerce service quality evaluation using an artificial neural network analysis," *Journal of Global Information Management*, vol. 30, no. 11, pp. 1-17, 2020.

[19] C. Chen, and F. Wang, "Exploring the Innovative Application of Azure Cloud Computing Platform in Cross-border E-commerce Operation:–Taking China and Southeast Asia Trade as an Example," *Frontiers in Computing and Intelligent Systems*, vol. 4, no. 2, pp. 21-26, 2023.

[20] Y. Chang, E. Iakovou, and W. Shi, "Blockchain in global supply chains and cross border trade: a critical synthesis of the state-of-the-art, challenges and opportunities," *International Journal of Production Research*, vol. 58, no. 7, pp. 2082-2099, 2020.

[21] C. Rui, "Research on classification of cross-border E-commerce products based on image recognition and deep learning," *IEEE Access*, vol. 9, pp. 108083-108090, 2020.

[22] J. Li, T. Wang, Z. Chen, and G. Luo, "Machine learning algorithm generated sales prediction for inventory optimization in cross-border E-commerce," *International Journal of Frontiers in Engineering Technology*, vol. 1, no. 1, pp. 62-74, 2019.

[23] H. Zhong, J. Zhang, and S. Zhang, "A combined prediction model of cross-border e-commerce export volume based on BP neural network and SVM," *International Journal of Technology, Policy and Management*, vol. 23, no. 3, pp. 310-328, 2023.

[24] H. Zhang, F. Jia, and J. X. You, "Striking a balance between supply chain resilience and supply chain vulnerability in the cross-border e-commerce supply chain," *International Journal of Logistics Research and Applications*, vol. 26, no. 3, pp. 320-344, 2023.

[25] F. Zhang, T.-Y. Wu, Y. Wang, R. Xiong, G. Ding, P. Mei, and L. Liu, "Application of Quantum Genetic Optimization of LVQ Neural Network in Smart City Traffic Network Prediction," *IEEE Access*, vol. 8, pp. 104555-104564, 2020.

[26] F. Zhang, T.-Y. Wu, J.-S. Pan, G. Ding, and Z. Li, "Human motion recognition based on SVM in VR art media interaction environment," *Human-centric Computing and Information Sciences*, vol. 9, 40, 2019.

[27] K. Wang, Z. Chen, X. Dang, X. Fan, X. Han, C.-M. Chen, W. Ding, S.-M. Yiu, and J. Weng, "Uncovering Hidden Vulnerabilities in Convolutional Neural Networks through Graph-based Adversarial Robustness Evaluation," *Pattern Recognition*, vol. 143, 109745, 2023.

[28] M. Barbosh, P. Singh, and A. Sadhu, "Empirical mode decomposition and its variants: A review with applications in structural health monitoring," *Smart Materials and Structures*, vol. 29, no. 9, 093001, 2020.

[29] H. Yang, and L. Yu, "A method of forecasting cross-border e-commerce stocking for SMEs based on demand characteristics and sequence trends under sustainable development strategy," *International Journal of Computational Systems Engineering*, vol. 7, no. 2-4, pp. 57-66, 2023.

[30] B. Pang, E. Nijkamp, and Y. N. Wu, "Deep learning with tensorflow: A review," *Journal of Educational and Behavioral Statistics*, vol. 45, no. 2, pp. 227-248, 2020.