# A resilient model for trade volume forecasting under economic uncertainty: Addressing challenges in the global supply chain


**Guanglan Zhou[1], Ziyi Wu[2]**

[1] Zhejiang Gongshang University, School of Management Engineering, Department of E-Commerce, China, ORCID: 0000-0002-5993-0638, guanglanzhou@zjgsu.edu.cn;

[2] Zhejiang Gongshang University, School of Management Engineering, Department of E-Commerce, China, ORCID: 0009-0000-1348-7530, 18510590338@163.com (corresponding author).



**Abstract:** *In recent years, the escalating economic uncertainty arising from Sino-US trade frictions has made the accurate forecasting of trade volume a crucial yet challenging task, with wide-ranging implications for the stability of the global supply chain. Precise trade forecasts are essential for supporting strategic decision-making and ensuring resilience across industries that rely on international trade. To address this challenge, this study introduces an innovative predictive model, the principal component analysis-simulated annealing-backpropagation neural network (PCA-SA-BPNN), specifically developed to enhance forecasting accuracy within this volatile economic landscape. The model utilizes principal component analysis (PCA) to reduce the dimensionality of extensive datasets collected from search engines, simplifying the data while retaining critical information. Simultaneously, simulated annealing (SA) is applied to optimize the backpropagation neural network (BPNN), effectively addressing the local optimization challenges often impair traditional backpropagation neural network models, which can hinder prediction accuracy. The effectiveness of the PCA-SA-BPNN model is demonstrated through comprehensive comparative experiments, demonstrating its superior performance compared to other models, including principal component analysis-adaptive differential evolution-backpropagation neural network (PCA-ADE-BPNN) and principal component analysis-backpropagation neural network (PCA-BPNN) models, as well as standalone XGBoost and BPNN models. The PCA-SA-BPNN model achieves notably lower mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE) values, with an $R^2$ approaching 1, underscoring its superior predictive performance. This research thus offers valuable insights into how combining dimensionality reduction, optimization techniques, and neural network networks can significantly enhance trade volume forecasting amidst economic uncertainties. Furthermore, it provides valuable insights into the interplay between predictive accuracy, model efficiency, and resilient decision-making within global supply chain management, contributing to both theoretical advancements and practical applications in the field.*



**Keywords:** *Sino-US trade frictions, PCA-SA-BPNN, dimensionality reduction, model optimization, predictive accuracy.*

**JEL Classification:** *C38, C45, C53, F17.*

***APA Style Citation:*** Zhou, G., & Wu, Z. (2026). A resilient model for trade volume forecasting under economic uncertainty: Addressing challenges in the global supply chain. *E&M Economics and Management*, *29*(1), 207–224. https://doi.org/10.15240/tul/001/2026-1-013

## Introduction

Over the past 20 years, globalization has rapidly expanded (Munir & Ameer, 2018), driving significant shifts in global political and economic landscapes (Panigrahi et al., 2018). China has emerged as the world's second-largest economy, and the economic gap between China and the United States is narrowing at an accelerating pace. However, the complex and evolving relationship between these two major economies continues to be a key source of uncertainty for China.

The Sino-US trade friction is one of the most serious economic challenges China has faced in recent years (Cheng et al., 2023). The logistics industry has emerged as a new catalyst and a driving force for national economic development in China (Zhou et al., 2024). However, trade tensions have led to rising costs, disrupting multinational operations and creating bottlenecks in China's supply chains. This threatens China's role as a global manufacturing hub and poses risks to the global industrial network.

This study addresses the research question (*RQ*): *How can trade volume be more accurately predicted amid the economic uncertainty from Sino-US trade friction on the supply chain?* Here, "accurately" means the model should reliably forecast trade volume, with results offering valuable insights for regulatory bodies and government policies.

Forecasting is vital for supply chain management, helping businesses manage risks and align with demand. Accurate trade volume forecasts enable companies to adapt to market changes. As a major trade player, China's supply chain trends significantly influence the global economy and trade relations, thereby supporting sustainable cooperation. To address this need, this study proposes the principal component analysis-simulated annealing-backpropagation neural network (PCA-SA-BPNN) model to enhance trade volume forecasting accuracy.

BPNN offers a unique advantage over traditional models that rely on statistical and distribution adjustments (Lyu & Zhang, 2019). However, traditional BPNNs often face local optimization issues due to dependence on initial parameters and weights (Wang et al., 2015).

To address these limitations, the model combines PCA and SA with BPNN. PCA reduces input data dimensionality, especially from real-time sources like Baidu, retaining key information, reducing overfitting, and improving generalization. It also identifies major factors affecting trade volume, aiding strategic decisions. SA enhances the model by optimizing BPNN weights and parameters, allowing occasional acceptance of suboptimal solutions to avoid local minima. With each iteration, the probability of accepting a suboptimal solution decreases, ensuring algorithm stability. The combination of PCA and SA helps the model better capture complex non-linear patterns, resulting in more accurate forecasts.

Despite some studies that have explored either PCA or SA in isolation with BPNNs, few have attempted to combine these techniques into a unified PCA-SA-BPNN framework. This study aims to fill that gap by applying the hybrid model to forecast trade volume under the economic uncertainty caused by Sino-US trade friction. The findings are expected to provide practical insights for policymakers and enterprises seeking to mitigate risks and enhance operational efficiency. The remainder of this paper will introduce the theoretical background, results and related discussions of this study.

## 1 Theoretical background

### 1.1 Leveraging search engine data for accurate forecasting

Collecting search engine data captures user interests, offers real-time insights, and enhances prediction accuracy. Research shows a clear link between social behavior and search data.

Tourism forecasting is crucial due to the sector's unique traits, like inherent variability and the inseparability of production and consumption, attracting widespread attention. Search engine data reveals traveler behavior and intentions, prompting many researchers to use internet big data to forecast tourism levels (Li et al., 2021; Sun et al., 2019). Yang et al. (2014) predicted local hotel occupancy rates at a destination in the United States based on Internet traffic data from DMO websites, and found that when using Internet traffic data as a time-sensitive leading indicator of visitor behavior, performance improved significantly. Bigné et al. (2019) investigated the impact of DMO on Twitter activity and hotel stay during a short vacation period, and found that some Twitter activities (such as retrophile and user replies, tourism tweets and DMO feeds) could significantly predict local hotel stay. Wen et al. (2019) developed an ARIMA-ANN hybrid model based on the Baidu index to predict Hong Kong travel demand from China. The predictive effect of this model is superior to that of the component model. Bi et al. (2020) used LSTM and combined historical data, search engine data and weather data at the same time to forecast the daily number of tourists in Jiuzhaigou and the Huangshan regions. Their findings demonstrate the powerful predictive performance of this method compared to other models.

Internet search engine data can also provide a new perspective for product sales data prediction, and the data used to predict must not only include consumer emotional attitude but also be forward-looking (Zhang et al., 2022). Kulkarni et al. (2012) use search engine data to improve the accuracy of sales forecast. Their basic assumption is that there is a positive correlation between product sales and the number of searches, similar to the study on the relationship between the volume of Google searches and the scale of influenza outbreaks (Lazer et al., 2014). Yu et al. (2019) proposed an oil consumption prediction model using Google search data, which outperformed models without this data in global oil sales forecasts.

## 1.2 Optimizing trade volume forecasts

Efforts have been made to make these forecasts scientific, building theories upon a solid statistical footing. The existing forecasting studies can be divided into statistical methods and artificial intelligence methods.

Statistical methods include time series forecasting, Kalman filtering, and the grey model (GM). Time series forecasting is widely used in finance, network data flow prediction, and power load forecasting, but all face the common issue of using past data to predict the future. Jiancheng and Sheng (2011) proposed an adaptive filtering algorithm combining the extended Kalman filter (EKF) and innovative point adaptive estimation to improve GPS measurement accuracy, optimizing the algorithm for real-time performance in POS machines with large initial heading errors. GM (1,1) is a basic forecasting model for gray information and small datasets, but it performs poorly in cases with weak exponential rules (Pei & Liu, 2021). The fractional gray model (FGM (1,1)) was developed to address this (Wu et al., 2013), offering higher accuracy under certain conditions.

Different from statistical methods, artificial intelligence methods are data-driven and do not require accurate mathematical models. They use the adaptive and self-learning capabilities of AI algorithms to predict future deformation trends by extracting network structure parameters from historical data. Key AI methods include support vector machine (SVM), long short-term memory neural networks (LSTM), and convolutional neural networks (CNN). SVM minimizes structural risk, balancing VC-confidence intervals and empirical error, outperforming other models in generalization. SVM training solves a quadratic programming problem with unique, optimal solutions. LSTM, a type of recurrent neural network (RNN), excels in long-term predictions and learning (Muzaffar & Afshari, 2019). CNN is good at capturing spatial or sequential features in data (Zhou & Wang, 2025). It removes noise from input data, enhancing feature extraction for better predictive modeling (Livieris et al., 2020).

In their investigation of the detrimental effects of COVID-19 on the functioning of international supply chains, Kazançoğlu et al. (2022) employed the developed system dynamics (SD) model, taking into account specific parameters, and evaluated the effects of COVID-19 on the global supply chain using Turkey and China as case studies. The impact of COVID-19 on the global supply chain is reflected in the volume of import and export as well as the logistics of international trade between nations. Therefore, this paper takes the total import and export volume of foreign trade in China's supply chain as the data source of trade volume.

Forecasting import/export trade volume has been widely studied, with models based on statistical methods and artificial neural networks (ANN). Sokolov-Mladenović et al. (2016) used ANN for economic growth estimation, while Alam (2019) employed ANN and ARIMA to forecast Saudi Arabia's trade volume. The findings demonstrated the effectiveness of ANN, ARIMA (1,1,2), and ARIMA (0,1,1) in forecasting the Kingdom of Saudi Arabia's annual total imports and exports. While statistical models struggle with the non-linear nature of trade volume data, ANN models can handle non-linearity but may suffer from overfitting, leading to less accurate predictions. Some scholars suggest combining different models to improve prediction accuracy and stability.

## 1.3 Advanced predictive models

The BPNN model is designed to address non-linear problems, enabling random non-linear mapping and autonomous learning of input-output relationships. As a mature, adaptive, and fault-tolerant non-linear mapping method, it is widely used in fields such as finance (Chaâbane, 2014), electric power systems, economy (Claveria & Torra, 2014), and water resources (López-Lineros et al., 2014), often replacing traditional forecasting methods.

However, problems still exist in the BPNN model, one of which is over-fitting due to the lack of generalization ability (Kratzert et al., 2018). To address the limitations of linear algorithms and overcome the issue of local optima, non-linear optimization techniques like the simulated annealing (SA) algorithm have been integrated with BPNN. Researchers tackled a prediction challenge in the electric industry by utilizing a hybrid approach that integrated the traditional BPNN with the SA optimization algorithm. This method successfully forecasted several parameters in wire electrical discharge processes, including cutting speed, average roughness (Ra), and maximum roughness (Rt). In the field of metallurgy, Bahrami and Doulati Ardejani (2016) applied a similar hybrid technique combining BPNN and SA to predict pyrite content in coal washing waste piles. Comparable approaches were also employed in other areas, such as water quality assessment and earthquake forecasting.

An excessive amount of input variables can complicate the BP network structure and increase the training load, reducing the learning rate. One of the most popular data analysis techniques in machine learning and artificial intelligence is principal component analysis (PCA) (Marukatat, 2023). PCA can increase the BP neural network's operating speed and simplify its structure. Numerous studies have integrated PCA and BPNN across various fields. Rajath Kumar et al. (2015) proposed a face recognition system based on PCA and BPNN. The multivariate data set of face photos was made simpler using PCA technology, and BPNN was utilized for training and learning to accomplish effective face recognition. The outcomes demonstrate how much faster the PCA-based facial recognition system is, but the recognition accuracy is affected, while the BPNN is vice versa, and the system combining the two technologies is preferred.

Traditional forecasting methods use simple linear models, limiting accuracy. Real-time search engine data enhances forecasting but large volumes slow processing. PCA reduces dimensionality, improving model generalization, while SA optimizes BPNN parameters. Most studies combine BPNN with PCA or SA, but few integrate both for economic trade volume prediction. This study proposes a PCA-SA-BPNN hybrid model to improve forecasting under economic uncertainty, aiding supply chain planning and future trade operations.

## 2 Research methodology

Fig. 1 displays the PCA-SA-BPNN flow chart. This research introduces SA to improve the BP network and uses PCA to extract numerous unrelated principal components. Lastly, the BP neural network model that has been tuned is applied to predict the trade volume under economic uncertainty in the supply chain.

### 2.1 Principal component analysis (PCA)

The primary goal of principal component analysis (PCA) is to minimize sample data loss and information loss by substituting new independent principal components for the multi-dimensional variables in the original data through linear transformation. A primary component can be used to represent the original sample's nature. The first principal component reflects the change in the raw data, while the miscellaneous principal component might reflect the prototype's many features (Tsai & Yang, 2015). The main calculation process of PCA is as follows (Li & Zhao, 2015):

**(1) Standardize raw data to eliminate dimensional side effects**

The unprocessed data are:

$$X = (x_{ij})_{n\times p} = (X_1, X_2, \dots, X_p) \quad (1)$$

A standardized formula is:

$$Z_{ij} = \frac{(x_{ij} - \bar{x}_j)}{S_j},\ i = 1,2,\dots,n;\ j = 1,2,\dots,p \quad (2)$$

Amongst them:

$$\bar{x}_j = \sum_{i=1}^{n} \frac{x_{ij}}{n}, S_j^2 = \sum_{i=1}^{n} \frac{(x_{ij} - \bar{x}_j)^2}{(n-1)} \quad (3)$$

The standardized matrix is:

$$Z = (z_{ij})_{n\times p} = (Z_1, Z_2, \dots, Z_p) \quad (4)$$

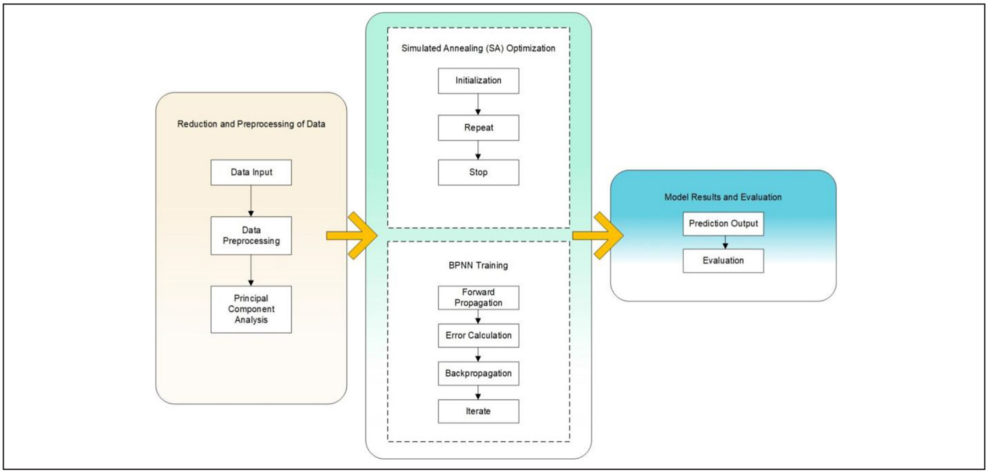


**Fig. 1: Calculation flowchart of PCA-SA-BPNN**

Source: own

**(2) Finding the matrix of correlation coefficients**

$$R = \left(\frac{1}{n-1}\right) Z'Z \tag{5}$$

$Z'$ is the transpose of $Z$, and $R$ is an $n \times n$ symmetric matrix with all 1 data on the diagonal.

**(3) Calculate the correlation coefficient matrix's eigenvalues and eigenvectors**

After obtaining the eigenvalues $\lambda_i$ via $|\lambda E - R| = 0$, $\lambda_i$ are sorted by size, where $\lambda_1 \geqq \lambda_2 \geqq \ldots \geqq \lambda_p \geqq 0$.

By using $(\lambda E - R)X = 0$, the eigenvectors $ZX_i$ are obtained:

$$ZX_i = (ZX_{i1}, ZX_{i2}, \ldots, ZX_{ip})' \tag{6}$$

**(4) Ascertain the number of main components**

The primary component's rate of contribution is:

$$\alpha_i = \frac{\lambda_i}{\sum_{i=1}^{p} \lambda_i} \tag{7}$$

The previous $m$ major components' combined contribution rate is:

$$CV_i = \frac{\sum_{i=1}^{m} \lambda_i}{\sum_{i=1}^{p} \lambda_i} \tag{8}$$

Generally, $m$ principal components with cumulative contribution rate $CV_i \geq 85\%$ are selected.

**(5) Find the primary component's expression**

$$F_i = ZX_{i1} \times Z_1 + ZX_{i2} \times Z_2 + \cdots + ZX_{ip} \times Z_p \tag{9}$$

**(6) Ascertain the role of the comprehensive evaluation**

$$F = \frac{\lambda_1 F_1 + \lambda_2 F_2 + \cdots + \lambda_m F_m}{\lambda_1 + \lambda_2 + \cdots + \lambda_m} \tag{10}$$

## 2.2 Simulated annealing (SA)

SA, introduced by Kirkpatrick in "optimization by simulated annealing," demonstrated its effectiveness in solving combinatorial problems like TSP by using a physical annealing analogy. The algorithm escapes local minima by accepting worse solutions at higher temperatures. Its application expanded through a practical simulation for TSP and later to complex optimization problems. The theoretical foundations were further enhanced with stochastic relaxation in image processing, connecting to Gibbs distributions and Bayesian restoration.

SA is based on thermodynamic principles, where a material at a high temperature has a greater likelihood of transitioning between different states, but as the material cools, the probability of transitioning into less stable states diminishes. In optimization, this is represented by exploring the solution space at higher “temperatures” (or randomness), allowing the algorithm to accept worse solutions to escape local minima. As the temperature decreases, the algorithm becomes more focused on converging to a stable and optimal solution.

To summarize, the key formulas from recent work on SA are:

**(1) Energy/cost difference**

$$\Delta E = E(S_{new}) - E(S_{old}) \tag{11}$$

This formula measures the change in the objective function when transitioning from the current solution $S_{old}$ to a new solution $S_{new}$.

**(2) Acceptance probability**

$$P = e^{-\frac{\Delta E}{T}} \tag{12}$$

The acceptance probability depends on the current temperature $T$ and the energy difference $\Delta E$. A higher temperature or smaller $\Delta E$ increases the likelihood of accepting worse solutions.

**(3) Cooling schedule**

$$T = \alpha T \tag{13}$$

The temperature is reduced by a factor of α after each iteration to gradually transition from exploration to exploitation.

## 2.3 Back propagation neural network (BPNN)

In 1986, Rumelhart, McClelland and other scientists proposed the concept of BPNN. One of the most popular neural network models is this multi-layer feed-forward neural network, which was trained using the error back propagation algorithm. A BP neural network has three layers: input, hidden, and output. The input layer

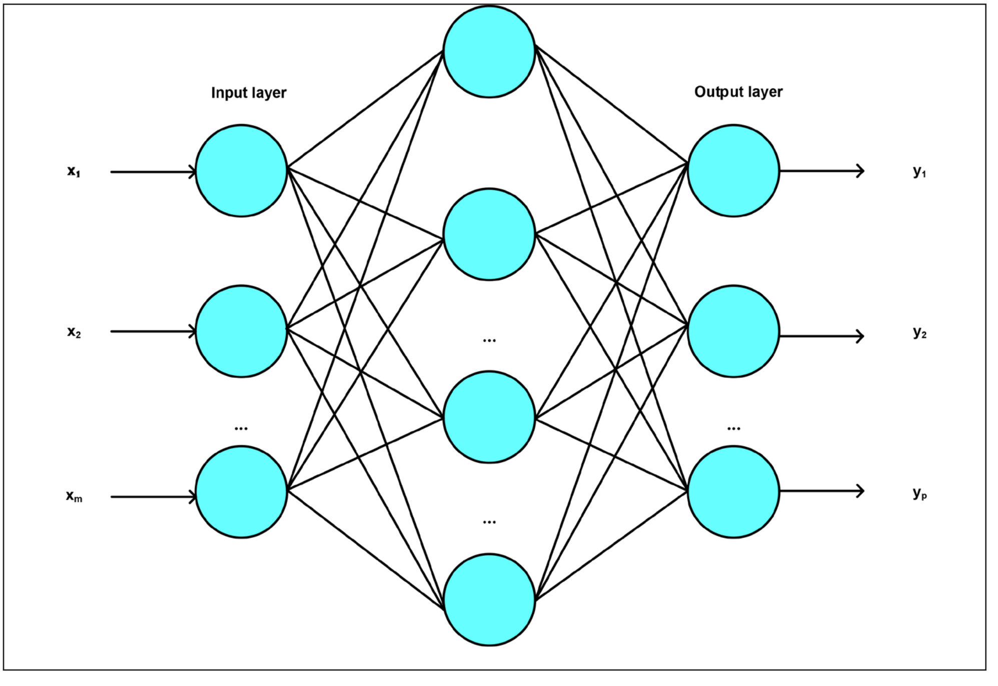


**Fig. 2: Structure of the single hidden layer neural network model**

Source: own

uses known data, and the output layer predicts results. The simplest BPNN has one hidden layer, but it can have more. The learning process involves forward propagation of information and backward propagation of errors. If the output is incorrect, the error is backpropagated to adjust the network. The erroneous signal modifies the weights and thresholds of neurons in each layer to satisfy the predetermined requirements as it propagates back through the network structure via the original connection channel (Chen et al., 2009; Mkadem & Boumaiza, 2009). The network structure then modifies the weights and thresholds of neurons in each layer to satisfy the prerequisites and backpropagates the erroneous signal in accordance with the original connection path.

Fig. 2 shows the BPNN model with single hidden layer, which belongs to the most basic BPNN model (Li et al., 2018).

It is assumed that there are *m* neurons, *h* neurons, and *p* neurons in the input layer, hidden layer, and output layer, respectively. Equations (14–15) represent the input and output values of the hidden layer, while Equations (16–17) represent the input and output values of the output layer.

$$I_j = \sum w_{ji} \times x_i + b_j \quad (i = 1, 2, 3, ..., m;\ j = 1, 2, 3, ..., h) \tag{14}$$

$$y_j = f_h\,(I_j)\ (j = 1, 2, 3, ..., h) \tag{15}$$

$$I_o = \sum w_{0j} \times y_j + b_o \quad (j = 1, 2, 3, ..., h;\ o = 1, 2, 3, ..., p) \tag{16}$$

$$y_o = f_p\,(I_o)\ (o = 1, 2, 3, ..., p) \tag{17}$$

The connection weights of the input layer and the hidden layer is represented by the symbol $w_{ji}$. The connection weights of the hidden layer and the output layer is represented by the symbol $w_{oj}$. The input values of the input layer is represented by the symbol $x_i$. The input values of the hidden layer and the output layer are represented by $I_j$ and $I_O$, respectively. The transfer functions of the hidden layer and the output layer are denoted by $f_h$ and $f_p$, respectively, while the threshold values of the hidden layer and the output layer are represented by $b_j$ and $b_o$. Tansig or logsin is the common activation function of the hidden layer, while purelin is the activation function of the output layer (Wang et al., 2015).

$$\mathrm{tansig}(x) = \frac{2}{1 + e^{-2x}} - 1 \tag{18}$$

$$\mathrm{logsin}(x) = \frac{1}{1 + e^{-x}} \tag{19}$$

$$\mathrm{purelin}(x) = x \tag{20}$$

# 3 Results and discussion

## 3.1 Results

This paper uses the National Bureau of Statistics of China to search and collate the trade volume on the Chinese supply chain from August 2018 to August 2023 (a total of 5 years, 61 months), and uses it as the neural network model's output variable.

Input variables, or keywords related to the trade volume under economic uncertainty due to the Sino-US trade friction on the supply chain, are obtained from data on monthly Baidu searches. Before trade activities, most businesses use search engines like Baidu, China's largest search engine. The Baidu index (http://index.baidu.com), based on extensive data since August 2018, tracks keyword search volumes related to trade, reflecting user demand, characteristics, and public opinion. The monthly data used in this research is the product of the monthly average data and the number of days in a month.

This paper uses the demand graph of Baidu index and related word recommendation technology to identify potential keywords that enterprises are interested in when conducting economic trade on the supply chain.

**(1) Determine the search keywords for the initial list.** Merchants need to make decisions in all aspects when conducting transactions in the supply chain. When considering global supply chain trade relations, Tian and Sarkis (2023) mainly considers the cutting-edge technologies such as block chain and sustainable production and the policy implications of consumption. Lin et al. (2023) found that increased trade policy uncertainty reduced export volume and technological innovation channels, and inhibited the improvement of supply chain efficiency. Therefore, the efficiency of supply chain transactions is also affected by trade policy, technological innovation and economy. After comprehensive analysis, this paper finally takes information, technology, cooperation, management, economy and policy as the main factors of economic trade in the supply chain, and the 17 key words are shown in Tab. 1.

**(2) Expand search keywords.** This paper uses the demand graph of Baidu index and related word recommendation technology to generate about 466 keywords on the basis of the initial list of search keywords, and then deletes some invalid or highly overlapping keywords, finally obtains 429 keywords. This step's objective is to broaden the keyword set to as nearly represent the enterprise's search requirements as feasible (Yang et al., 2015).

**Tab. 1: Six main factors related to transactions in the supply chain and their keywords**

| Factors | Keywords |
|---|---|
| Information | Information sharing |
| | Information symmetry |
| | Information system |
| Management | Concept management |
| | Talent management |
| | Risk management |
| Technology | Digital technology |
| | Technological innovation |
| | Intelligence |
| Economy | Price |
| | Logistics cost |
| | Industrial structure |
| Cooperation | Contract trust |
| | Development intention |
| | Cooperation relationship |
| Policy | Laws and regulations |
| | Social responsibility |

Source: own

**(3) Calculate the correlation between keywords and the trade volume under economic uncertainty on the supply chain.** This paper multiplied the monthly search data of each keyword and the days of each month to obtain the monthly data of each keyword, then conducted correlation analysis between the monthly data of each keyword and the trade volume data. The correlation coefficient of the two was obtained. We determine the greatest correlation coefficient between the trade volume and the keywords with a lag order of 0 ~ 6 by taking into account the lag between the search engine data and the trade volume.

**(4) Identify the last keyword.** The Pearson correlation coefficient was used in this study due to its simplicity, interpretability, and computational efficiency, which make it particularly suitable for analyzing high-dimensional search data. By capturing linear correlations between keywords and economic activities or policy changes, it provides a clear and efficient way to identify relevant features. In this context, the absolute value of 0.6 was chosen as the threshold for the Pearson correlation coefficient between keywords and trade volume under economic uncertainty in the supply chain, resulting in the identification of 18 key keywords (Yang et al., 2015). These keywords are shown in Tab. 2.

Tab. 2: **18 keywords and their correlation coefficients with the trade volume under economic uncertainty on the supply chain**

| Keywords | Correlation coefficients |
|---|---|
| Information cocoon | 0.758** |
| Digital transformation | 0.744** |
| Data governance | 0.747** |
| Achievement transformation | 0.777** |
| Science and technology media | −0.727** |
| Safety guarantee | 0.764** |
| Talent team construction | 0.745** |
| Improve the mechanism | 0.780** |
| B2b information | −0.695** |
| Platform economy | 0.693** |
| Laws and regulations database | 0.695** |
| ESG | 0.690** |
| Resources | −0.686** |
| Application optimization | −0.680** |
| Resource sharing | −0.649** |
| Transformation of scientific and technological achievements | 0.604** |
| Integration | 0.641** |
| Foresight | 0.601** |

Note: * indicates a significant correlation at the 0.05 level (double-tailed); ** indicates a significant correlation at level 0.01 (double-tailed).

Source: own

Tab. 3: **KMO value and Bartlett sphericity test**

| KMO sample appropriateness measure | | **0.907** |
|---|---|---|
| Bartlett sphericity test | Approximate chi-square | 1,540.257 |
| | Degree of freedom | 153.000 |
| | Significance | 0.000 |

Source: own

In order to test whether the data is suitable for PCA, KMO test and Bartlett sphericity test are performed on the data. The results are shown in Tab. 3. The KMO value is 0.907, greater than 0.7, and the significance is less than 0.05, demonstrating that PCA is supported by the data. The next step can be taken.

In order to make the model less complex, this paper standardize the monthly data of 18 keywords collected to reduce dimensions, and finally obtains three principal components.

Tab. 4 illustrates how three common factors are retrieved based on the eigenvalue larger than 1 principle, with a total variance contribution rate of 85.073%. Therefore, the three principal components can have a major impact on the trade volume under economic uncertainty.

Determining the training-to-test ratio is crucial before using neural networks. Bichri et al. (2024) employed an 8:2 training/test split to examine the impact of data distribution on model performance. The authors argued that this split

**Tab. 4: Contribution rates and cumulative contribution rates of the three principal components**

| Components | Contribution rate (%) | Cumulative contribution rate (%) | Characteristic value |
|---|---|---|---|
| Component 1 | 70.192 | 70.192 | 12.635 |
| Component 2 | 9.314 | 79.506 | 1.677 |
| Component 3 | 5.567 | 85.073 | 1.002 |

Source: own

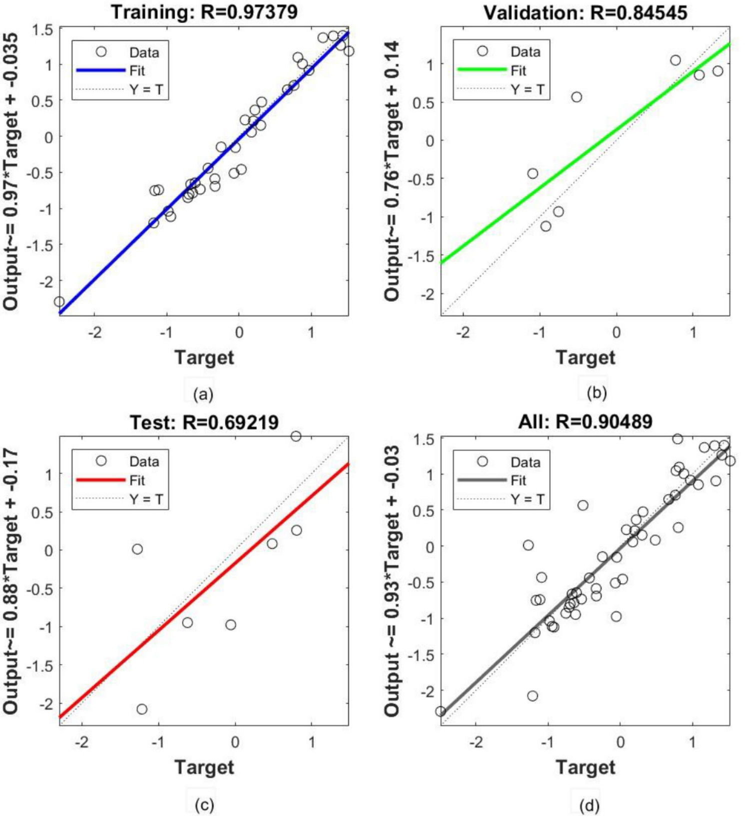


**Fig. 3: Goodness of fit of the BP neural network model after optimization**

Source: own

allows the model to better generalize, capturing complex patterns during training while still providing a reliable test set for evaluation. Therefore, this study also adopts this ratio, using data from the first 49 months (August 2018–August 2022) as the training set and the remaining 12 months (September 2022–August 2023) as the test set. This ensures that both the training and test sets reflect the impact of the pandemic, thereby supporting the model's generalization ability.

In order to apply the data, this paper calculates the principal component data by using the component matrix obtained from the principal component analysis, and standardizes the principal component data to the interval (−1, 1).

This paper predicts the trade volume under economic uncertainty on the supply chain with three principal components as input variables and the trade volume as output neurons. A hidden layer is set to predict the trade volume. The following are the settings for the neural network prediction model's parameters:

i) Input layer nodes: the number of principal components retained after PCA (e.g., numComponents);
ii) Output layer nodes: 1 node (since this is a regression task);
iii) Tolerance: TolFun = 1e-6;
iv) Maximum iterations: MaxIter = 100;
v) Display progress: Display = "iter."

Fig. 3 illustrates the performance of the PCA-SA-BPNN model across different datasets, showcasing its regression fitting capabilities. To ensure methodological transparency, we clarify that the validation set was derived from the initial training data through an 80:20 split, serving exclusively for hyperparameter tuning during model development. The test set (20% of the full dataset) remained strictly independent and was reserved solely for final evaluation.

The model achieves an excellent fit on the training set with an *R*-value of 0.97379, indicating a strong linear relationship between the predicted and target values and suggesting effective learning of underlying patterns. On the validation set, the *R*-value of 0.84545

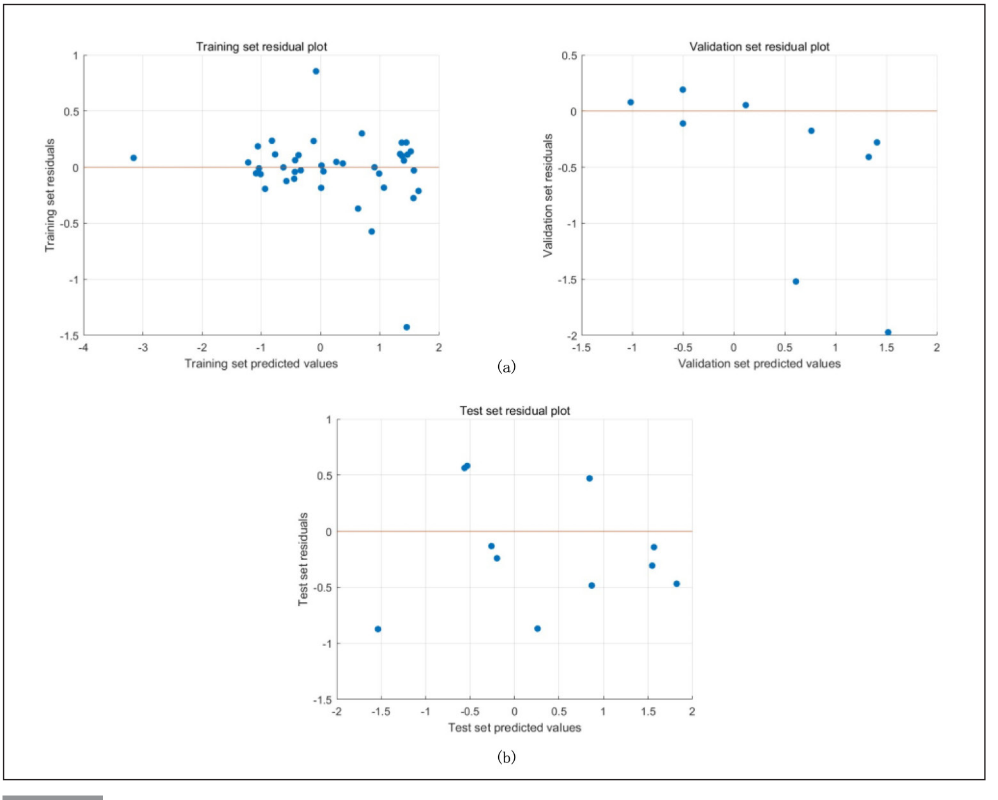


**Fig. 4:** **Residual plot of the training and validation sets**

Source: own

reflects robust generalization on data unseen during training, while the test set yields an *R*-value of 0.69219, demonstrating moderate yet practical predictive capability for real-world scenarios. The overall *R*-value across all data is 0.90489, signifying relatively high predictive consistency. Although test set performance is slightly lower, residual analysis confirms strong generalization, as errors exhibit no systematic bias (Feng et al., 2020). Residual plots for the training and validation sets (Fig. 4) further validate the model's robustness, showing evenly distributed residuals around zero without clear patterns indicative of overfitting or underfitting. Similarly, the residual plot for the test set (Fig. 4) exhibits a comparable residual distribution, with points scattered randomly around the zero line, further indicating the absence of systematic bias or noticeable deviation. Integrating findings across all residual plots, it is clear that the model maintains a balanced and reliable predictive performance without signs of significant overfitting or underfitting, thereby supporting its applicability and robustness for practical use.

Prediction error is used to describe the potential discrepancy between the actual and projected values. An important phase in the forecasting process is forecasting error analysis. We can determine the model's prediction accuracy by computing the prediction error. The stronger the model, the higher the prediction accuracy and the closer the predicted outcomes are to the actual results.

To evaluate the performance of the PCA-SA-BPNN model, this paper forecasts the trade volume from September 2022 to August 2023 using the established PCA-SA-BPNN framework. A comparative analysis is also provided, highlighting the differences in predictive accuracy between the PCA-SA-BPNN model and other models, including the BPNN, PCA-BPNN, PCA-ADE-BPNN, and XGBoost models.

The forecasted results from the five predictive models, alongside the actual trade volume, are presented in Fig. 5. This figure compares

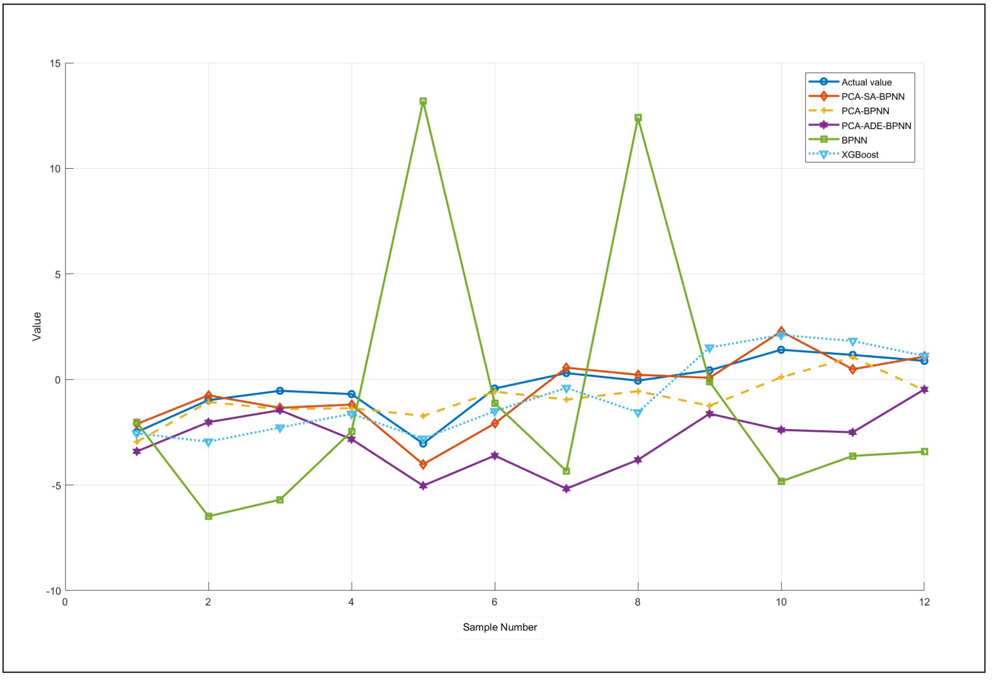


**Fig. 5: Comparison results of actual trade volume with those of PCA-SA-BPNN, PCA-BPNN, PCA-ADE-BPNN, BPNN, XGBoost**

Source: own

predictions of various models (PCA-SA-BPNN, PCA-BPNN, PCA-ADE-BPNN, BPNN, and XGBoost) with actual values. The PCA-SA-BPNN model aligns closely with actual data, showing the best fit. PCA-BPNN performs stably with minor deviations, while XGBoost captures general trends despite some errors. BPNN is the least stable, with notable fluctuations, and PCA-ADE-BPNN has moderate accuracy, showing larger errors in some cases. Overall, PCA-SA-BPNN performs the best among all models.

In order to further evaluate the quality of PCA-SA-BPNN model, this paper uses MAE, MSE, MAPE, RMSE and MAPE to evaluate the prediction accuracy of five models. They are calculated as shown in Equations (21–24). The prediction performance and error comparison results of models are shown in Tab. 5.

$$MSE = \frac{\sum_{i=1}^{i=n}(y_i - y)^2}{n} \tag{21}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{i=n}(y_i - y)^2}{n}} \tag{22}$$

$$MAE = \frac{\sum_{i=1}^{i=n}|y_i - y|}{n} \tag{23}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{Y_{actual,i} - Y_{pred,i}}{Y_{actual,i}}\right| \times 100\% \tag{24}$$

Tab. 5 indicates that the PCA-SA-BPNN model achieves the best performance across all evaluation metrics. It yields the lowest MSE, RMSE, MAE, and MAPE, demonstrating minimal prediction errors. Its $R^2$ value of 0.818 indicates high accuracy and strong predictive power. PCA-SA-BPNN effectively captures main trends, leading in overall prediction performance. While XGBoost performs well, it trails behind PCA-SA-BPNN, with PCA-BPNN close behind. PCA-ADE-BPNN and BPNN perform the worst, with BPNN showing high errors and a very low $R^2$, underscoring PCA-SA-BPNN's advantage in complex predictions. Negative $R^2$ values for models PCA-ADE-BPNN and BPNN indicate that their predictive accuracy was inferior to a simple mean predictor under extreme economic shocks. This serves as an empirical warning against applying linear assumptions to non-linear crisis scenarios.

**Tab. 5: Model evaluation results**

| Model | MSE | RMSE | MAE | $R^2$ | MAPE (%) |
|---|---|---|---|---|---|
| **PCA-SA-BPNN** | 0.573 | 0.757 | 0.624 | 0.818 | 9.494 |
| **PCA-BPNN** | 1.345 | 1.160 | 0.882 | 0.572 | 14.550 |
| **PCA-ADE-BPNN** | 7.036 | 2.653 | 2.118 | −1.239 | 20.820 |
| **BPNN** | 52.245 | 7.228 | 5.274 | −15.626 | 33.719 |
| **XGBoost** | 0.956 | 0.978 | 0.874 | 0.696 | 13.165 |

Source: own

## 3.2 Discussion

This paper develops a PCA-SA-BPNN-based model to predict trade volume under economic uncertainty in the supply chain, validated by a numerical example. Integrating search engine data, especially the Baidu index, enhances real-time accuracy by reflecting public trends and sentiment shifts, which can directly influence trade volumes.

The strong predictive performance of the model can be attributed to the effectiveness of PCA in reducing dimensionality, eliminating redundant information while preserving key features. This not only mitigates the risk of overfitting in BPNN but also enhances computational efficiency. Li et al. (2024) demonstrated that integrating PCA with BPNN improved the model's goodness-of-fit ($R^2$) from 0.85 to 0.986 after dimensionality reduction, underscoring PCA's ability to enhance sensitivity to critical features. Simultaneously, SA optimizes the initial weights of BPNN through global search, preventing the model from being trapped in local optima (Liu et al., 2020). Liu et al.

(2020) also showed that the SA-optimized BPNN model achieved higher prediction accuracy than the standard BPNN model, making it particularly effective in the complex, uncertain environment of trade volume forecasting.

Although BPNN has been around for a long time, it still holds significant advantages over other models in certain scenarios. For instance, while LSTM excels at sequential modeling, it requires large datasets for effective training and imposes higher computational costs (You et al., 2019). Transformer models, on the other hand, are prone to overfitting and often necessitate data augmentation (Shorten & Khoshgoftaar, 2019). In the context of this study, the dataset consists of monthly trade volumes, forming a short time series with high dimensionality. LSTM's tendency to underfit when confronted with limited sequence lengths would limit its applicability here. For medium-sized, high-dimensional datasets where temporal dependencies are not dominant, the PCA-SA-BPNN framework proves highly efficient by leveraging dimensionality reduction and global optimization to achieve a balanced trade-off between accuracy and computational efficiency.

In comparison, models like XGBoost, which excel with structured data, face challenges in handling the high-dimensional sparse nature of the data in this study, particularly after PCA reduces dimensionality. While XGBoost performs effectively in certain scenarios, especially with tabular datasets, its tree-based structure may struggle to capture the complex, non-linear relationships present in our feature set, which includes 429 keywords from search engine data. The discrete nature of XGBoost's hyperparameters, such as tree depth and leaf node count, presents additional challenges when attempting to combine it with optimization techniques like simulated annealing (SA), which is better suited for continuous parameter spaces (Geng et al., 2024; Lee & Lee, 2014). As a result, while XGBoost demonstrates strong performance, the PCA-SA-BPNN model provides superior results by effectively integrating dimensionality reduction and global optimization to address the unique characteristics of our trade volume forecasting task.

In practical trade volume forecasting, there is typically a certain tolerance for error. The acceptable margin of error for trade volume predictions in supply chain management generally depends on the specific business context and the forecasting time horizon. In most cases, an error within 10% is considered acceptable (Yang, 2024). However, for forecasts with longer time horizons, due to greater market volatility and increased data uncertainty, decision-makers may need to adjust the acceptable error tolerance based on specific business requirements. For example, in supply chain management, small discrepancies in trade volume forecasts can have significant implications for inventory management, production scheduling, and demand planning. An acceptable margin of error allows businesses to maintain a buffer for disruptions while ensuring optimal resource allocation. Given the residual analysis in the current model, the observed errors fall within this acceptable range, suggesting that the model's predictions are robust enough for practical use in supply chain management. This is especially important when anticipating supply chain disruptions such as those arising from geopolitical factors, like Sino-US trade tensions, which could affect trade volumes, inventory levels, and resource planning. By forecasting potential disruptions, the model provides valuable insights that allow businesses to mitigate risks and maintain operational efficiency. The robustness of the PCA-SA-BPNN framework, particularly in managing high-dimensional, sparse data, and its ability to avoid local optima, makes it a reliable tool for navigating complex supply chain environments where uncertainty is high and accurate forecasting is critical for maintaining competitive advantage.

## Conclusions

The results clearly demonstrate this model (PCA-SA-BPNN) provides good forecasting performance in forecasting trade volume under economic uncertainty within supply chains. With consistently lower MSE, RMSE, MAE, and MAPE values compared to other models, along with an $R^2$ closer to 1, the model provides strong evidence of its predictive accuracy. This performance is particularly significant because the PCA-SA-BPNN effectively balances the trade-off between accuracy and computational efficiency. The integration of principal component analysis (PCA) enhances the model by reducing dimensionality, discarding irrelevant data, and preserving key features that are essential for accurate prediction. This not only mitigates the risk of overfitting but also streamlines the computational process, making the model more efficient in handling large, high-dimensional datasets.

Moreover, the use of simulated annealing (SA) plays a critical role in the model's optimization. By applying global search techniques, SA ensures that the model avoids local optima, improving its ability to navigate complex and uncertain environments. This ability to find global optima contributes to the model's overall predictive strength, especially in scenarios with high variability, such as those found in global trade forecasting.

This model offers substantial practical value, as it enables managers to allocate resources, such as production capacity, logistics, and workforce planning – more effectively by comparing predicted trade volumes with actual data. Importantly, its ≤10% MAPE aligns with the acceptable error threshold for supply chain forecasts, providing operational reliability. The insights generated can also inform strategic decisions, including investment planning, market expansion, and partnership selection, particularly in light of external challenges like the Sino-US trade conflict. By offering a reliable prediction tool, PCA-SA-BPNN empowers businesses to respond proactively to market changes, optimizing decision-making processes and helping mitigate risks.

Furthermore, the incorporation of real-time data from search engines, such as the Baidu index, enhances the model's accuracy by capturing public sentiment and trend shifts, which have direct implications for trade volumes. These innovations not only improve the model's predictive capabilities but also support the broader digital transformation of businesses, driving innovation in research, development, and technology across the logistics and supply chain sectors (Zhou et al., 2022).

Despite the model's strengths, some limitations remain. The use of Pearson correlation for keyword selection primarily captures linear relationships, which may overlook non-linear dynamics such as cyclical fluctuations or delayed effects between search volume and trade volume. Additionally, the dataset's temporal scope may limit the model's application to long-term economic cycles. Finally, while demonstrating robustness during the study period (2018–2023), the current results may be specific to the post-COVID economic context and cannot guarantee generalizability across other economic regimes. Therefore, extrapolation beyond the observed timeframe remains a key limitation for future research to address. Future work could explore non-linear feature selection methods, extended temporal coverage, and rolling window validation to enhance forecasting stability across diverse economic conditions.

***Acknowledgments:*** *Supported by Zhejiang Province Philosophy and Social Science Project (24NDJC139YB); Special Fund for Basic Research Expenses of Zhejiang Provincial Universities (XT202207); China Social Science Key Fund (21&ZD154); China Scholarship Council (202009545007) and Zhejiang Gongshang University Graduate Research Innovation Fund Annual General Project (YBXM2024084).*

## References

Alam, T. (2019). Forecasting exports and imports through artificial neural network and autoregressive integrated moving average. *Decision Science Letters*, *8*(3), 249–260. https://doi.org/10.5267/j.dsl.2019.2.001

Bahrami, S., & Doulati Ardejani, F. (2016). Prediction of pyrite oxidation in a coal washing waste pile using a hybrid method, coupling artificial neural networks and simulated annealing (ANN/SA). *Journal of Cleaner Production*, *137*, 1129–1137. https://doi.org/10.1016/j.jclepro.2016.08.005

Bi, J.-W., Liu, Y., & Li, H. (2020). Daily tourism volume forecasting for tourist attractions. *Annals of Tourism Research*, *83*, 102923. https://doi.org/10.1016/j.annals.2020.102923

Bichri, H., Chergui, A., & Hain, M. (2024). Investigating the impact of train/test split ratio on the performance of pre-trained models with custom datasets. *International Journal of Advanced Computer Science and Applications*, *15*(2), 331–339. https://doi.org/10.14569/IJACSA.2024.0150235

Bigné, E., Oltra, E., & Andreu, L. (2019). Harnessing stakeholder input on Twitter: A case study of short breaks in Spanish tourist cities. *Tourism Management*, *71*, 490–503. https://doi.org/10.1016/j.tourman.2018.10.013

Chaâbane, N. (2014). A hybrid ARFIMA and neural network model for electricity price prediction. *International Journal of Electrical Power & Energy Systems*, *55*, 187–194. https://doi.org/10.1016/j.ijepes.2013.09.004

Chen, W.-C., Fu, G.-L., Tai, P.-H., & Deng, W.-J. (2009). Process parameter optimization for MIMO plastic injection molding via soft computing. *Expert Systems with*

*Applications*, *36*(2), 1114–1122. https://doi.org/10.1016/j.eswa.2007.10.020

Cheng, S., Hua, X., & Wang, Q. (2023). Corporate culture and firm resilience in China: Evidence from the Sino-US trade war. *Pacific-Basin Finance Journal*, *79*, 102039. https://doi.org/10.1016/j.pacfin.2023.102039

Claveria, O., & Torra, S. (2014). Forecasting tourism demand to Catalonia: Neural networks vs. time series models. *Economic Modelling*, *36*, 220–228. https://doi.org/10.1016/j.econmod.2013.09.024

Feng, C., Li, L., & Sadeghpour, A. (2020). A comparison of residual diagnosis tools for diagnosing regression models for count data. *BMC Medical Research Methodology*, *20*(1), 175. https://doi.org/10.1186/s12874-020-01055-2

Geng, X., Wu, S., Yan, Q., Sun, J., Xia, Z., & Zhang, Z. (2024). An optimized XGBoost model for predicting tunneling-induced ground settlement. *Geotechnical and Geological Engineering*, *42*(2), 1297–1311. https://doi.org/10.1007/s10706-023-02619-x

Jiancheng, F., & Sheng, Y. (2011). Study on innovation adaptive EKF for in-flight alignment of airborne POS. *IEEE Transactions on Instrumentation and Measurement*, *60*(4), 1378–1388. https://doi.org/10.1109/TIM.2010.2084710

Kazançoğlu, Y., Ekinci, E., Mangla, S., Sezer, M. D., & Ozbiltekin-Pala, M. (2022). Impact of epidemic outbreaks (COVID-19) on global supply chains: A case of trade between Turkey and China. *Socio-Economic Planning Sciences*, *85*, 101494. https://doi.org/10.1016/j.seps.2022.101494

Kratzert, F., Klotz, D., Brenner, C., Schulz, K., & Herrnegger, M. (2018). Rainfall-runoff modelling using long short-term memory (LSTM) networks. *Hydrology and Earth System Sciences*, *22*(11), 6005–6022. https://doi.org/10.5194/hess-22-6005-2018

Kulkarni, G., Kannan, P. K., & Moe, W. (2012). Using online search data to forecast new product sales. *Decision Support Systems*, *52*(3), 604–611. https://doi.org/10.1016/j.dss.2011.10.017

Lazer, D., Kennedy, R., King, G., & Vespignani, A. (2014). The parable of Google Flu: Traps in big data analysis. *Science*, *343*(6176), 1203–1205. https://doi.org/10.1126/science.1248506

Lee, C.-Y., & Lee, D. (2014). Determination of initial temperature in fast simulated annealing. *Computational Optimization and Applications*, *58*(2), 503–522. https://doi.org/10.1007/s10589-013-9631-y

Li, G., Teng, Y., Ding, S., & Hou, X. (2024). Complex mechanical system safety prediction based on multidimensional indexes: An MBSA-PCA-BPNN method. *Engineering Failure Analysis*, *159*, 108130. https://doi.org/10.1016/j.engfailanal.2024.108130

Li, L., & Zhao, J. (2015). Comprehensive evaluation of parallel mechanism and robot performance based on principal component analysis and kernel principal component analysis. In *Proceedings of International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. 39th Mechanisms and Robotics Conference* (Vol. 5B). ASME. https://doi.org/10.1115/DETC2015-47032

Li, S., Chen, T., Wang, L., & Ming, C. (2018). Effective tourist volume forecasting supported by PCA and improved BPNN using Baidu index. *Tourism Management*, *68*, 116–126. https://doi.org/10.1016/j.tourman.2018.03.006

Li, X., Li, H., Pan, B., & Law, R. (2021). Machine learning in internet search query selection for tourism forecasting. *Journal of Travel Research*, *60*(6), 1213–1231. https://doi.org/10.1177/0047287520934871

Lin, B., Yuan, L., & Lu, B. (2023). How does trade policy uncertainty affect supply chain efficiency: A case study of listed companies of Chinese port industry. *Sustainability*, *15*(9), 7140. https://doi.org/10.3390/su15097140

Liu, B., Wang, R., Zhao, G., Guo, X., Wang, Y., Li, J., & Wang, S. (2020). Prediction of rock mass parameters in the TBM tunnel based on BP neural network integrated simulated annealing algorithm. *Tunnelling and Underground Space Technology*, *95*, 103103. https://doi.org/10.1016/j.tust.2019.103103

Livieris, I. E., Pintelas, E., & Pintelas, P. (2020). A CNN-LSTM model for gold price time-series forecasting. *Neural Computing and Applications*, *32*(23), 17351–17360. https://doi.org/10.1007/s00521-020-04867-x

López-Lineros, M., Estévez, J., Giráldez, J. V., & Madueño, A. (2014). A new quality control procedure based on non-linear autoregressive neural network for validating raw river stage data. *Journal of Hydrology*, *510*, 103–109. https://doi.org/10.1016/j.jhydrol.2013.12.026

Lyu, J., & Zhang, J. (2019). BP neural network prediction model for suicide attempt among Chinese rural residents. *Journal of Affective Disorders*, *246*, 465–473. https://doi.org/10.1016/j.jad.2018.12.111

Marukatat, S. (2023). Tutorial on PCA and approximate PCA and approximate kernel PCA. *Artificial Intelligence Review*, *56*(6), 5445–5477. https://doi.org/10.1007/s10462-022-10297-z

Mkadem, F., & Boumaiza, S. (2009). Extended Hammerstein behavioral model using artificial neural networks. *IEEE Transactions on Microwave Theory and Techniques*, *57*(4), 745–751. https://doi.org/10.1109/TMTT.2009.2015092

Munir, K., & Ameer, A. (2018). Effect of economic growth, trade openness, urbanization, and technology on environment of Asian emerging economies. *Management of Environmental Quality: An International Journal*, *29*(6), 1123–1134. https://doi.org/10.1108/MEQ-05-2018-0087

Muzaffar, S., & Afshari, A. J. E. P. (2019). Short-term load forecasts using LSTM networks. *Energy Procedia*, *158*, 2922–2927. https://doi.org/10.1016/j.egypro.2019.01.952

Panigrahi, S. S., Bahinipati, B., & Jain, V. (2018). Sustainable supply chain management: A review of literature and implications for future research. *Management of Environmental Quality: An International Journal*, *30*(5), 1001–1049. https://doi.org/10.4018/IJSDA.2019070102

Pei, L., & Liu, J. (2021). A predictive analysis of the business environment of economies along the belt and road using the fractional-order grey model. *Journal of Mathematics*, *2021*, 3153731. https://doi.org/10.1155/2021/3153731

Rajath Kumar, M. P., Keerthi Sravan, R., & Aishwarya, K. M. (2015). Artificial neural networks for face recognition using PCA and BPNN. In *Proceedings of TENCON 2015–2015 IEEE Region 10 Conference*. IEEE. https://doi.org/10.1109/TENCON.2015.7373165

Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, *6*(1), 60. https://doi.org/10.1186/s40537-019-0197-0

Sokolov-Mladenović, S., Milovančević, M., Mladenović, I., & Alizamir, M. (2016). Economic growth forecasting by artificial neural network with extreme learning machine based on trade, import and export parameters. *Computers in Human Behavior*, *65*, 43–45. https://doi.org/10.1016/j.chb.2016.08.014

Sun, S., Wei, Y., Tsui, K.-L., & Wang, S. (2019). Forecasting tourist arrivals with machine learning and internet search index. *Tourism Management*, *70*, 1–10. https://doi.org/10.1016/j.tourman.2018.07.010

Tian, X., & Sarkis, J. (2023). Towards greener trade and global supply chain environmental accounting. An embodied environmental resources blockchain design. *International Journal of Production Research*, *62*(8), 2705–2724. https://doi.org/10.1080/00207543.2023.2232890

Tsai, J.-T., & Yang, P.-Y. (2015). PCA-LBG-based algorithms for VQ codebook generation. *International Journal of Electronics*, *102*(4), 529–547. https://doi.org/10.1080/00207217.2014.905994

Wang, L., Zeng, Y., & Chen, T. (2015). Back propagation neural network with adaptive differential evolution algorithm for time series forecasting. *Expert Systems with Applications*, *42*(2), 855–863. https://doi.org/10.1016/j.eswa.2014.08.018

Wen, L., Liu, C., & Song, H. (2019). Forecasting tourism demand using search query data: A hybrid modelling approach. *Tourism Economics*, *25*(3), 309–329. https://doi.org/10.1177/1354816618768317

Wu, L., Liu, S., Yao, L., Yan, S., & Liu, D. (2013). Grey system model with the fractional order accumulation. *Communications in Nonlinear Science and Numerical Simulation*, *18*(7), 1775–1785. https://doi.org/10.1016/j.cnsns.2012.11.017

Yang, X., Pan, B., Evans, J. A., & Lv, B. (2015). Forecasting Chinese tourist volume with search engine data. *Tourism Management*, *46*, 386–397. https://doi.org/10.1016/j.tourman.2014.07.019

Yang, Y. (2024). The improved BP neural network supply chain performance evaluation model under the cross-border e-commerce marketing chain. *Journal of Computational Methods in Sciences and Engineering*, *24*(3), 1727–1740. https://doi.org/10.3233/jcm-247290

Yang, Y., Pan, B., & Song, H. (2014). Predicting hotel demand using destination marketing organization's web traffic data. *Journal of Travel Research*, *53*(4), 433–447. https://doi.org/10.1177/0047287513500391

You, Y., Hseu, J., Ying, C., Demmel, J., Keutzer, K., & Hsieh, C.-J. (2019). *Large-batch training for LSTM and beyond* [arXiv:1901.08256]. ArXiv. https://doi.org/10.48550/arXiv.1901.08256

Yu, L., Zhao, Y., Tang, L., & Yang, Z. (2019). Online big data-driven oil consumption forecasting with Google trends. *International Journal of Forecasting*, *35*(1), 213–223. https://doi.org/10.1016/j.ijforecast.2017.11.005

Zhang, C., Tian, Y.-X., & Fan, Z.-P. (2022). Forecasting sales using online review and search engine data: A method based on PCA-DSFOA-BPNN. *International Journal of Forecasting*, *38*(3), 1005–1024. https://doi.org/10.1016/j.ijforecast.2021.07.010

Zhou, G., & Wang, S. (2025). Enhancing credit risk decision-making in supply chain finance with interpretable machine learning model. *IEEE Access*, *13*, 14239–14251. https://doi.org/10.1109/access.2025.3530433

Zhou, G., Xu, Y., & Zhang, F. (2024). Measurement of innovation efficiency in logistic enterprises: Evidence from China based on the three-stage DEA-Malmquist index model approach. *The American Journal of Economics and Sociology*, *83*(2), 331–381. https://doi.org/10.1111/ajes.12545

Zhou, G., Zhang, F., & Fei, Y. (2022). Research on the development efficiency of educational science and technology in China: An approach from three-stage DEA Malmquist model. *Technology Analysis & Strategic Management*, *36*(6), 1067–1082. https://doi.org/10.1080/09537325.2022.2069556