

# Using Artificial Intelligence in IC Substrate Production Predicting

Zhifang Liu (✉ [liuzhifang710811@163.com](mailto:liuzhifang710811@163.com))

Nanyang Institute of Technology

---

## Research Article

**Keywords:** Artificial intelligence, IC substrate production prediction, Neural network

**Posted Date:** July 19th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-552378/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Using artificial intelligence in IC substrate production predicting

Zhifang Liu<sup>1\*</sup>

<sup>1</sup>School of information engineering, Nanyang Institute of technology, Guangzhou, 510800, China  
E-mail: liuzhifang710811@163.com

\*Corresponding author: Zhifang Liu (liuzhifang710811@163.com)

**Abstract:** Today's technology products are changing with each day, the purpose is to bring more convenience to people, but also the competition among the technology industries is more competitive. In such environment, whether the company's decision-making is correct or not will directly affect the future development of an enterprise. Therefore, how an enterprise can formulate and construct a set of appropriate decision-making systems to accurately predict the future market will be the first important issue for enterprises. This research proposed an artificial intelligence predicting system to estimate manufacturing capacities and client demands, and providing it to manufacturing managers as a reference for inventory arrangements so that inventory can be adjusted appropriately to avoid excessive inventory levels. In recent years, neural networks have been widely and effectively applied to many predicting problems. The main reason is that most of the predicting problems are nonlinear models. And the backward neural network has the ability to construct nonlinear models. In this study, a predicting model combining grey correlation and neural network will be used to establish a high-accuracy prediction system for the production predict of IC product. First, grey correlation analysis will be used to screen out the most relevant factors among many factors. And then put these factors into the neural network prediction model for training and prediction. The results show that the training prediction error and the empirical error value are about 14%. This value indicates that the prediction ability is better, so the proposed prediction model can be applied to the prediction of IC substrate production. It provided a predictive reference material and provide decision making with a more accurate, convenient and a fast tool to enhance the company's competitiveness.

**Keywords:** Artificial intelligence; IC substrate production prediction; Neural network.

## 1. Introduction

In recent years, in response to changes in market demand, improvements in packaging processes and continuous growth in IC technology, flexible electronic packaging products have developed rapidly and have been widely used in the semiconductor packaging industry. Among them, Chip On Film( COF) is a combination of IC chips, The products of placement technology and flexible substrate carrying technology are mainly used in the IC packaging process of panel displays. The finished products include flat-panel displays, notebook computers, mobile phones and other related electronic products. Since COF is the largest application in display driver IC packaging, the maturity of the technology is directly related to the development and growth of the flat panel display industry. According to the current market, COF driver IC packaging is mainly used higher-end displays have advantages especially in driver IC packages that require thinner circuits, which can avoid the loss of panel scrap due to driver IC bonding errors. In recent years, the display industry as Fig.1has been booming globally and the global market for new COF soft boards. The next few years will be a period of high COF growth.

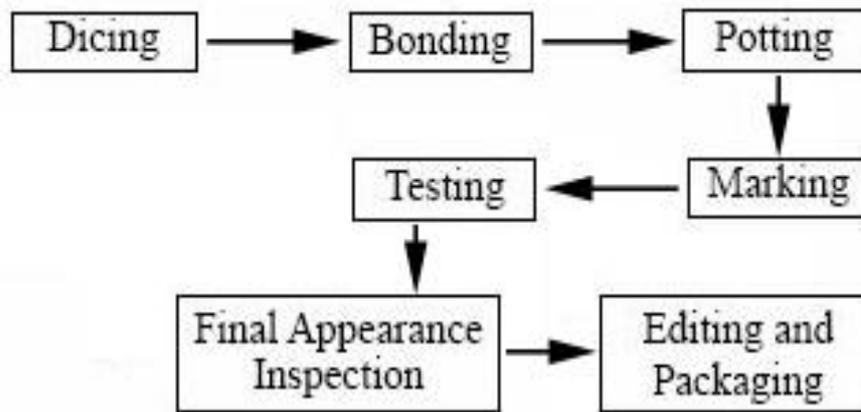


Fig. 1 Tape and reel mass production line structure

Since COF demand is arduous in the flat display panel industry chain, COF is the carrier material of driver ICs and stands in the position of strategic key components. Therefore, the COF industry is worthy of becoming a strategic key component industry. Global panel manufacturers continue to adjust inventory, affecting global panels. The industry's demand for terminal products also affects the prosperity of the electronics industry. Compared with the traditional manufacturing industry, the current production pattern of the electronics industry is more complex, with diverse and small product portfolios, unstable orders, urgently needed delivery and other issues, coupled with factors such as the global economic downturn, rising unemployment and sluggish prosperity. It has a certain impact on the supply and demand of end products. The semiconductor packaging industry is a kind of order-to-order manufacturing service industry. Its production is based on customer orders. The demand for production capacity also varies with market demand and inventory levels from time to time. However, changes have made it impossible to accurately provide capacity strategies as an investment benchmark. Therefore, the semiconductor packaging industry not only pays attention to information integration between upstream and downstream manufacturers, but also pays more attention to the issues of capacity strategy and application. In order to reduce inventory pressure, companies need to accurately predict the volume of orders from customers, so that the company's own competitiveness and profits are better than other competitors. How to establish a production predicting mechanism to become a useful competitive among companies [1-3].

This research takes COF products in the display driver IC packaging industry as the research object, and explores the changes in future market output. Compared with other manufacturing industries, this industry is particularly focused on continuous innovation. How to produce lighter, thinner, and more durable products is a goal that needs to be pursued continuously. Therefore, the establishment of an effective and accurate capacity predicting system can improve predicting performance and reduce the cost backlog caused by overproduction and the loss of orders caused by shortages. After researching and collecting data, it is found that the prediction ability of smart prediction systems is more accurate than traditional prediction models. In particular, neural network research theories are widely used in the field of prediction. Therefore, this research adopts this type of prediction mode and its main research purpose as follows:

1. Aiming at COF products as the main research body, establish a production predict model.
2. This study combine grey relational analysis and neural network predicting methods to predict the output to improve the accuracy of the prediction.
3. Use this predicting model to provide the company with a more objective and accurate predicting method in practice, thereby reducing company costs.

Generally speaking, IC carrier boards are divided into four categories: (1) material properties, (2) packaging form, (3) connection technology between chip and carrier, and (4) number of packaged chips. According to the four categories, multiple categories can be extended according to the four categories. The various data used in this research are the monthly requirements for IC carrier boards, among which the network training data is from January 2018 to December 2019 [4-6].

## 2. Literature review

### 2.1 IC carrier board manufacturing process and application

#### 2.1.1 IC carrier board

At present, IC substrate technology can be divided into two categories, one of which is rigid IC substrate, such as: Ball Array (BGA), Flip Chip (FC), Wire Bonding (WB), Chip Level Packaging (CSP); The other type is soft IC carrier, such as TCP (Tape Carrier Package) tape and reel chip carrier package, COG (Chip on Glass) flip glass, and COF flip chip package (Chip on Film). The soft IC carrier board can be mainly divided into CSP/BGA, TAB, COG and COF. As the screen resolution increases, the number of driver IC pins will also increase, making the IC pin pitch smaller and smaller. According to the statistics of IEK, the main foot pitch of the driver IC will be reduced from the previous 40um-50um to 35-40um. Due to the limitations of materials and manufacturing processes, the TCP carrier board will not be able to compete with the COF carrier board due to the significant decrease in the yield rate when the driver IC pitch is reduced to 35-40um. The COF carrier board has high size, stability and high heat resistance, so it is more suitable for high-density packaging and also has the characteristics of high bending and thin packaging. It can also place some passive components and help drive ICs. The electrical nature. Therefore, COF carrier boards will gradually replace TCP and become the mainstream of driver IC packaging, and will gradually be applied to markets such as LCD TV and PDP TV. Therefore, many new domestic manufacturers have jumped out of TCP and directly switched to the next generation of COF product lines. In order to seek lower component costs, downstream LCD panel manufacturers are bound to support COF substrate manufacturers more actively. Once manufacturers can achieve effective breakthroughs in product yield control, they will greatly expand their production capacity and further divide the driver IC substrate market. There are not many installations, and the production cost is significantly higher than that of hard substrates. In recent years, with the rapid development of IC substrate technology, flexible CSP/BGA substrates have gradually been replaced [5-7].

#### 2.1.2 Classification of IC carrier boards

According to the current industrial structure, the classification of IC substrates can be classified according to materials, electroplating types, product circuit specifications and application areas, as follows:

- **Distinguish by material:**

- Three-layer material**

For this type of product, the incoming PI and copper foil will be pressed and processed first, and the PI and copper foil will be combined before subsequent processes are carried out. This type of production method is called TCP product.

- Flexible Copper Clad Laminate (FCCL)**

This material is a copper foil substrate directly provided by the raw material supplier. After entering the factory, there is no need to press PI and copper foil, and the follow-up process can be directly carried out. IC substrates produced by this type of material are called COF products. Because the thickness of this material is thinner, the precision is higher than that of TCP, and the process cost is lower than that of TCP, it has obviously become the current market trend as Fig. 2.

- **Distinguish by route:**

According to the end customer's application requirements and product design specifications, it can be divided into two types of designs: Gate and Source. Three, according to the electroplating category:

- Tin plating**

The IC carrier board (TCP and COF), in order to protect the circuit from external force or foreign matter damage, and is designed to be plated with a layer of inert metal on the surface of the product. Its function can prevent the circuit from oxidizing and has a protective effect.

- Gilded**

In order to pursue better conductivity and circuit protection, some products require electroplating of gold

on the circuit surface, such as: P.H. products, CSP products, PDP products.

Four, according to the field of application:

- LCD panel driver IC (TFT-Source/Gate): Such products are mainly used in the driver ICs of larger TFT LCD screens, such as: LCD monitors, LCD TVs, digital cameras...etc.
- OLED/STN: These products are used in small-size OLED and STN screen driver ICs, such as mobile phones, mobile photo albums, MP3 players.
- PDP: Such products are used in the driver IC of plasma TV screens.
- CSP: Such products are used in the packaging of mobile phone chips, and products such as consumer electronics and DRAM are gradually being used.
- P.H.: These products are used in the packaging of ink cartridge control chips.

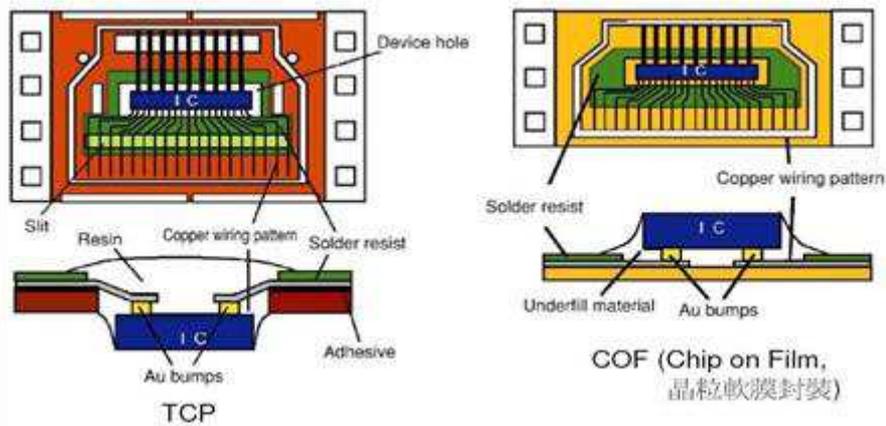


Fig. 2 Product example of IC carrier board

### 2.1.3 The manufacturing process of IC carrier board

The production process of flexible IC substrates will be adjusted according to the needs of end customers, product categories, material form specifications, and there will be different production sequences or process adjustments. The common process as fig. 3.

#### Step 1. Punching

Design the mold according to the customer's specifications and use the punching machine to punch out the hole on the PI raw material.

**Step 2. Photoresist Coating:** The surface of the product is coated with photosensitive emulsion.

**Step 3. Exposure.** The photomask after the circuit design is used as the negative film, and the photosensitive emulsion coated on the product is exposed to the circuit and pattern using the ultraviolet rays of the machine.

**Step 4. Developing:** Use developing syrup to reveal the lines and graphics of the product after exposure.

**Step 5. Etching:** Remove the unused area with etching potion, and the remaining part is the required circuit.

**Step 6. Stripping** clean the remaining photosensitive emulsion on the surface with a potion.

**Step 7. Semi-Inspection:** Semi-finished product line inspection is carried out by automatic optical inspection equipment or manual inspection. If the product is abnormal or the yield is too low, follow-up arrangements can be made immediately ·

**Step 8. Plating:** Plating tin or gold on the wires to prevent oxidation of the product wires.

**Step 9. SR Printing:** The ink is printed on the product circuit area to protect the integrity of the circuit.

**Step 10. Punching:** The mold is designed according to the customer's needs, and the required hole is punched out on the product with a punching machine.

**Step 11. Open/Short Test:** According to the specifications required by the customer, conduct short-circuit test and inspection of the product line ·

**Step 12. Final Inspection:** Perform final circuit and appearance inspections according to the specifications required by the customer.

**Step 13. Packing:** According to the specifications required by the customer, the production batch is

packaged.

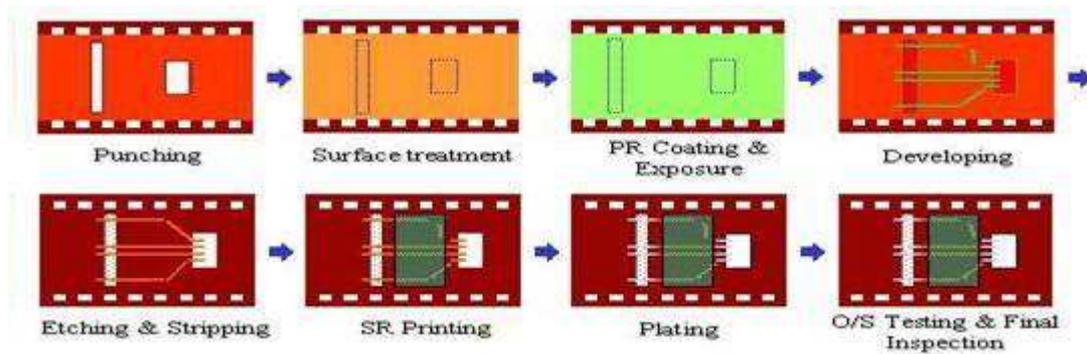


Fig. 3 Process flow chart of IC carrier board

## 2.2 Prediction definition

### 2.2.1. Prediction Theory

With the continuous improvement of industrial technology and the progress of scientific research and development, Predicting has been widely used in various fields, such as: economics, management, education, trade, information, finance, transportation, human resources and many other aspects. This is an important method for making management decisions in the future. The so-called Prediction needs to have the following four characteristics:

- **The continuity of the predict program:**

The environment is dynamic and changing, and its impact on predicting items is also inconsistent. Prediction need to understand these effects first, and appropriately adjust the past predict results according to the current situation.

- **The uncertainty of the predict situation:**

The importance of predicting is based on the uncertainty of the unknown situation in the future. The uncertainty of the future is caused by the impermanence of related factors. Although related factors can sometimes be predicted, the degree of mutual impact is difficult to measure. Because of the inability to control the mutual influence of these factors, the uncertainty of the future situation is still an inevitable result.

- **Continuity of predict items:**

Only the prediction items can appear continuously, and the existing data can be used as the basis of the prediction. These data can be formed into a certain pattern. By understanding and analyzing this pattern, an unknown possible pattern can be derived. If the predict item is an emergency, For example: World War, etc., this matter cannot be predicted

- **The error of the predict results:**

Under normal circumstances, the prediction result cannot be in error with the facts. Even if the data used completely reflects the real event, the method used is perfect and correct. However, due to the uncertainty of the future and the fact that the data reflecting the facts have certain limits, the prediction result. There will still be a certain degree of gap with the truth [8-11].

### 2.2.2. Definition of prediction

The so-called prediction is an art and a science of inferring future events. Prediction is based on collecting historical data and analyzing changes in the real situation through a certain mathematical model to understand the reasons for the changes and the state of their effects, and to predict and infer what is likely to happen in the future. The definition of prediction is as follows: Prediction is an explanation of unobserved events. The so-called unobserved events not only refer to future events, but also events that have occurred. If it involves these two things, it is called prediction in the broad sense; if it involves only future events, it is called predicting in the narrow sense. The accuracy of the prediction is closely related to the length of the prediction. Different predicts have different lead times, so the required time is also

different. It divides the predict time into three categories:

- **Short-Term Predicting:** The coverage period is usually three months, which is suitable for the company's middle and low-level management units, such as material ordering quantity, material ordering timing, production batch size, production timing, operator assignment and machine allocation, etc.
- **Medium-Term Predicting:** The coverage period is from three months to two years, which is suitable for the allocation of various resources within the enterprise, such as production and inventory budgets, seasonal manpower, products, equipment, capital and material production planning, etc.
- **Long-Term Predicting:** Refers to predicts of more than two years, such as the basis of high-level strategic planning, that is, the decision of the site, the expansion of the factory, the development of the product, and the capital planning [12-15].

### 2.2.3 Method of prediction

Prediction is the origin of the overall production plan. Without an accurate predict of production demand, any production plan cannot proceed smoothly. Predicting methods are divided into three categories: 1. Qualitative methods; 2. Time series analysis and projection; 3. Causal models; in addition to the above three categories, Added six categories of statistical predicting methods, artificial intelligence and grey system theory.

#### (1). Traditional predicting methods:

- **Qualitative methods:** Qualitative methods are usually based on the opinions of experts and some special materials that record events, and predict future events based on past experience or special sensory functions. Through the collection, analysis and integration of the opinions of experts and customers, we can explore the problems and make predictions. This method is suitable when historical data is lacking or when available data is uncertain, such as Delphi Method, Market Research, Panel Consensus, Visionary Predicting, and Historical analogy, etc., such as when a new product is to be introduced into the market, can use the above method.
- **Time series analysis and projection:** Time series analysis and projection methods are the opposite of qualitative analysis methods. They pay attention to past patterns and changes in patterns, and use past historical data to predict the future. It is necessary to collect continuous observations that appear in the chronological pattern and use these observations. As a basis, predict the possible value at the next point in time. Such as self-regressive integrated moving average mode (ARIMA), exponential smoothing method (Exponential Smoothing), moving average method (Moving Average), X-11 method and trend projection method. Generally speaking, the longer the prediction period, the greater the error; the shorter the prediction period, the smaller the error. The time series method performs better than other models for trendy and seasonal predicting models.
- **Causal models:** The causal analysis method pays attention to the relationship between the various elements of the system, and emphasizes historical data, transforms historical events into a trend chart of time series data, and distinguishes its characteristics. This analysis method includes regression model and econometric model, Input-Output Model and Intention-to-Buy and Anticipation Surveys. The main purpose of econometric models is to explore the relationship between external economic variables. Statistical methods are used to measure or verify and then provide the basis for analysis. Nowadays, it is not common to consider only econometric models as a tool for analysis or predicting, and most of them will be used in conjunction with other methods.

The traditional predicting methods have quite effective, but with the evolution of the times, information technology has become more and more developed, and more computationally complex models have been development. Compared with the previous predicting models, the demand has been insufficient, and have begun to develop artificial intelligence-based predicting methods, and many artificial intelligence models have emerged as a result. Artificial intelligence refers to computer systems or computer programs with human-like behavior and knowledge, including problem reasoning solving, knowledge storage and learning, and the ability to recognize and interpret human

language [16-19].

## (2). Smart predicting methods

- **Artificial intelligence:** Use computers to simulate the learning process of humans or the organization of knowledge, including expert systems, fuzzy theory, artificial neural network, genetic algorithm.
- **Grey system theory:** The characteristics of this method are that the information is incomplete and the number is to be counted. This mode is quite simple in operation, the modeler does not need to have a deep statistical foundation, and the amount of data required is very small, so it is suitable for use when the data is insufficient or difficult to obtain.
- **Statistical predict method:** Statistical predicting is to conduct surveys and statistics on different objects according to projects and goals, and to find out the predicting model of the predicted development trend of the project.

Use regression and time series to construct a tourism demand predict model, predict the number of foreign tourists to China, and find the most suitable model. The results of the study found that the tourism industry has seasonality and trends, combined with regression models and time series models. Excellent predictive ability. Use the business tax declaration data of the manufacturing and service industries, analyze and filter them as a sample, combine with neural network technology such as data mining, and explore the important factors that affect the predict of tax evasion, so as to construct a business tax evasion predict model. According to the research results, the more important influencing factors are the ratio of tax payable for the current period to total sales, the ratio of tax before deduction to total sales, and the ratio of before deduction. The ratio of tax amount to deductible tax amount, the classification accuracy rate reached 90%. Based on the electricity consumption data and the temperature and humidity data of various places in the Meteorological Bureau, the temperature is first analyzed by regression method, and then the backward transfer neural network analysis is used to calculate and solve according to different learning rates and inertia factors to construct the best the power predicting model allows the power dispatching unit to do the dispatching work in advance. Apply predicts to analyze the current development of the TFT panel industry and future development trends, and explore the relevance of the TFT panel industry's capacity utilization rate to the business performance of the TFT panel industry under the pressure of global market competition, and put forward specific suggestions to provide TFT The panel industry is planning production capacity in the future, the capacity utilization strategy adopted [16-18].

The gray predict analysis model in the gray system theory is used to predict the power supply demand, and the result of comparison with the data is used to analyze and confirm the degree of accuracy. The research results show that the actual power consumption statistics are compared, and the error of the results is maintained within two percentages. The accuracy is quite high, which confirms that this predicting method is suitable for district power demand.

## 2.3. Grey association analysis

Some of the information is known and some of the information is unknown. In control theory, color is a way of naming. The depth of color is often used to represent the researcher's knowledge and understanding of the internal information and the entire system itself; black represents the lack of information, while white represents sufficient information, which is somewhere between white. The ones in the middle of black represent some unknown and some known information, and this type of system is called a gray system. Grey system theory, its theory is mainly aimed at the uncertainty of the system model and the incompleteness of the system data, the relational analysis, the model construction, and the use of prediction and decision methods to analyze and discuss system conditions. The grey theory can be used for the multi-input, not certainty, not enough, discrete data and the discrete data of events. Efficient analysis has a wide range of applications. The application of gray system theory to management and engineering mentioned that gray theory can be roughly summarized as: Grey generating, Grey relational analysis, Grey model, Grey prediction, Grey decision making, Grey control and the six types as follows:

- **Grey generating:**

Gray generation is mainly used in the processing of supplementary information data. In some messy data, it can be used to generate

Means to reduce the randomness in the data, and try to show the concealed rules and characteristics to improve and modify the regularity. The generation of this method belongs to the transformation of the data level, and the purpose of the level transformation is to find out the law among them. The common generation methods in gray theory include the following:

- \* Grey Relational Generating Operation: The data is processed without distortion by referring to the actual situation.
- \* Accumulated Generating Operation: Accumulate data according to the number of times.
- \* Inverse Accumulated Generating Operation: Inverse Accumulated Generating Operation of data.

- **Grey relational analysis:**

Traditional statistical regression stipulates that there must be a "mutual influence" relationship between variables and variables, and it needs to be used in an environment with a large amount of data, a typical distribution of data (such as a normal distribution), and a limited number of changing factors. The gray relational analysis is mainly aimed at the comparison of the evolution of the system development pattern and the quantitative description method. According to the mathematical foundation of space theory, according to the four theorems: "proximity, symmetry, integrity, standardization", to confirm the analysis and comparison. The correlation degree and correlation coefficient between the sequence and the reference sequence. Grey correlation is mainly looking for a quantitative method to measure and evaluate the degree and size of the correlation between various factors, to find out the important factors and main characteristics that affect things or the target value, and then promote the system to be more rapid and effective. The quantitative measurement value in the gray correlation analysis process is called the gray correlation degree. The correlation degree analysis depends on the quantitative comparison analysis of the development trend of the system. In essence, it is the analysis and comparison of the geometric shape between the lines of the variable changes to explore the relationship between the two series of numbers. Degree of relevance. Grey relational analysis is based on the degree of grey relation as a method to measure the degree of correlation between factors. It has the following advantages: the calculation method is simple, the sample size is small, the data does not have to conform to the typical distribution such as the normal distribution, the calculation workload is small, and the qualitative analysis conclusion will not contradict each other. Therefore, it is suitable for dealing with prediction models with unclear and incomplete data.

- **Grey model:**

The model or differential equation of gray theory is established by using the data in the generation process to generate a set of gray differential equations and gray differential equation models, which is called gray modeling. It can be roughly divided into the following categories:

- \* GM (1, 1): represents the first-order differential, an input variable, the simplest and most commonly used, generally used as prediction.
- \* GM (1, N): represents the first-order differential, with N input variables, which is generally used in multi-dimensional correlation analysis.
- \* GM (0, N): stands for zero-order differential, with N input variables, which is generally used in multi-dimensional correlation analysis.

- **Grey prediction:**

Using the GM (1, 1) model as the basis to predict the existing data is actually to find out the future state of each element in a certain series. According to its purpose, it is divided into the following four types:

- \* Data prediction: do a series of predictions on the size of the data, such as numerical predictions.
- \* Anomaly prediction: The prediction of whether an abnormal phenomenon occurs within a certain period of time is often used for weather or disaster prediction.
- \* Graphical predictions: predictions made by constructing graphics for the development of existing data.
- \* System prediction: Combining GM (1, 1) and GM(1, N) models to predict multiple variables in the system, and predict the relationship between each other.

- **Grey decision making:**

When a certain event occurs, it has different effects due to different countermeasures, and the decision made by combining countermeasures with the GM (1, 1) model is called gray decision. In order to adapt to different system environments, it can be divided into several decision-making types. Divided into grey statistical decision-making, grey relational decision-making, grey clustering decision-making, grey situation decision-making, grey-level decision-making, grey planning decision-making, grey model decision-making and grey interval decision-making.

- **Grey control:**

Gray control is to find out the regularity of the development behavior through the system behavior data to predict the future behavior. When the predictive value is obtained, the predictive value is transmitted back to the system for system control law, which belongs to the feedforward control method. , It is similar to artificial intelligence and has the function of self-adjustment. The sales volume and share of the automobile market and the sales volume and share of the mainland automobile market are statistical data, and the GM (1, 1) model in the gray theory is used as a research method to predict the sales volume of the mainland automobile market and each car The share of funds, and compare with time series method, DRI/WEFA, expert predict method and other methods. The empirical results show that the average accuracy of the GM (1, 1) model for predicting and the mainland automobile market is more than 90%, which shows that the gray theory system is also applicable to the automobile industry. The export volume is predicted for the products of the auto parts industry, and the GM (1, 1) model and GM (1, 1)  $\alpha$  model in Grey Theory are used to predict. Take the piston as an example, the data range from January 2002 to December 2007 is the monthly data of export volume. The empirical results show that the MAPE predicted by the GM (1, 1) model and the GM (1, 1)  $\alpha$  model are 10.14% and 10.09%, respectively, indicating that the gray theory system is suitable for predicting and analyzing the export volume of auto parts, but There is still room for improvement.

Applying grey relational analysis combined with neural network training, taking the lead frame and ball gate array (BGA) products in the packaging industry as the research objects, discussing the changes in future market output, and comparing traditional predicting models with intelligent predicting systems, Found that the prediction ability of intelligent prediction system is more accurate, especially the neural network is widely used in prediction. Take the quotation history data of the electronic test company as an example, apply grey relational analysis to select the predictor to be used and combine it with a neural network, test the actual case and compare its differences with each prediction model, and aggregate and analyze the result data , The results show that this predicting model is capable of providing predict data for the average unit price of the test, which serves as a template for the company's future predicts to reduce the cost loss caused by excessive predict errors.

## 2.4 Neural network

The learning algorithm of the hidden layer has enabled the resurgence of neural networks and new breakthroughs. The research and application of neural networks has entered a new era, and the applications of neural networks are gradually no longer limited to speech recognition and handwriting. Information processing issues such as text recognition, speech synthesis, image compression, noise filtering, and adaptive control have been extended to various fields, such as integrating psychological aspects, physiological aspects, and computer technology to evolve into new research directions. The neural network model mentioned in the application and practical composition of the neural network model after years of research, its function is mainly to try to imitate the biological nerve conduction system, because the biological nerve conduction system is both in sound, hearing, image and vision. Has a very good performance, of course, I also hope that these models can have good results in this field. The biological nervous system is composed of many nerve cells. The function of nerve cells is the same or related in series to form a neural network, which is the so-called neural network. The input signal in the neuron is through the synapse (Synapse). After the internal potential changes, it is transmitted from the dendrites to the cell body (Soma), and then transmitted to the dendrites of the next neuron through the axon, and converted into the input signal of the next neuron. Assuming that the input path of the cell is dendrites, the synapses connected between the cell bodies receive the signals from the surrounding cell

bodies, and the output path of the cell bodies is equivalent to the axons. When the nerve cell is stimulated, the received information is transmitted to the brain through the neural network, and the brain recognizes or judges the information, and finally transmits the response to the nerve cell through the neural network, so that the nerve cell performs the desired action [16-17].

The current definition of artificial neural network is a computing system that includes software and hardware, and consists of many non-linear computing units (Neuron) and many connections between these computing units, and these computing units. Generally, calculations are performed in a parallel and decentralized manner, so that a large amount of data can be processed at the same time. This design can be used to solve various applications that require a large amount of data calculation, such as mechanical vehicle engine diagnosis, electronic circuit diagnosis, Satellite communication broadcasting fault diagnosis, etc. As the neural network has the ability to deal with nonlinear operations and learning, it has excellent performance in certain fields of diagnosis, so it is often widely used in diagnosis and classification problems, and the application level is more extensive. Neural networks have the following characteristics:

- **The characteristics of parallel processing:** In the early days, artificial intelligence mainly focused on small-scale parallel processing research, but recently it has turned to super large-scale research. Artificial neural networks are designed with reference to biological tree-like neural networks as a blueprint. In the early days, parallel technology was not yet mature and there was no way to conduct in-depth research. Nowadays, due to the maturity of ultra-large parallel processing and the continuous improvement and progress of related theories, parallel processing has become among the most active research areas.
- **Fault tolerance features:** With a high degree of tolerance in operational operations, the entire neural network will participate in the operation to solve problems at the same time. If the input data is interfered by some impurity signals, it will not affect the accuracy of the calculation. And even if 10% of the neural network fails, it can still operate normally.
- **Associative memory:** It can also be called Content Addressable Memory, which can memorize the previously trained input pattern and correspond to the ideal output value. Just provide a part of the data, you can accept all the data and can tolerate errors, just as human beings want to evoke all the images and only need to look at a small part of the images to remember it. This is called the combined memory effect.
- **Optimization:** It can be applied to solve the problem of non-algorithm processing, or the algorithm processing will be very labor-consuming and time-consuming.
- **Storage capacity:** The traditional artificial intelligence storage method is to store all the rules and databases in the computer, while the neural network disperses the rules and databases in the weight of each neuron connection, which greatly reduces the data access time and reduces the data storage space.
- **Inductive ability:** The data input that has never been seen or is not yet complete can be classified and summarized according to the original network structure, and it is not necessary to have a clear output for reference.

Artificial intelligence combined with neural networks can learn the characteristics of knowledge and the ability of the control system to recognize unknown system environments, establish groundwater flow patterns to simulate the groundwater level changes in observation wells, and conduct research on groundwater geological parameters. The method is to establish a closed-loop control system in which the information fed back from the neural network training process is input into the control system detection device to correct the groundwater flow continuity method. On-site hydrogeological parameters in the program. According to the research results, the training error of the closed-loop control system device made by the similar neural network can be lower than  $10^{-3}$  power during simulation, and the error of data test and model prediction is between  $5.1 \times 10^{-4}$  to  $9.2 \times 10^{-3}$ . Compared with the previous results obtained by only using a neural network to predict the groundwater level or using MODFLOW software to adjust the hydraulic parameters at a fixed rate by manual trial and error, this method has greatly improved the accuracy of the model simulation. Neural network models are divided into supervised, unsupervised, associative, and optimized models.

There are many types of network modes included in the class of neural networks. According to different problem types, there are different applicable network modes. The most representative learning mode in the class of neural networks is the backward neural network, and the scope is also the most common mode. It belongs to a supervised learning network, so it is especially suitable for diagnosis and prediction applications. Discuss the performance evaluation of the backward neural network in production predicting, and mix the genetic algorithm and Taguchi experiment method to obtain the best predicting effect. Experimental results show that the hybrid artificial intelligence method is significantly better than the two traditionally commonly used predicting methods, grey predicting and regression analysis, when evaluating predict errors and accuracy. Combining particle swarm algorithm and neural network prediction methods, and using gray correlation analysis to filter out the factors with higher correlation, then use the intelligent prediction mode to predict the production of IC substrates. The research results show that the prediction model constructed by particle swarm calculation combined with neural network has a higher prediction error than the prediction model of backward neural network, but its prediction speed is significantly better than other prediction models. Therefore, it is recommended that the case company use particle swarm Algorithms and backward pass neural compound methods for prediction. Construct MIMO process prediction and control mode with backward neural network. The two neural networks are used to establish the process output prediction mode and the process adjustment mode. At the same time, the backward neural network is used to establish the CMP process input and output relationship model, which is used to verify the prediction of the neural network and the benefits of the controller. Evaluate the object, and construct the prediction and control model of the process to compensate for the error caused by the interference. The research results show that the backward-propagation neural network prediction and controller can effectively control the interference output and reduce the changes caused by the process interference. Future research can consider the control cost factors to optimize the process control.

Aiming at the elevation change speed of 653 first-class leveling points in the area, artificial neural networks (ANNs) are used to construct the prediction of the stratum subsidence within the area, and the nodes of another 37 first-class leveling networks are selected as independent Detection points are used to measure the accuracy of the prediction model. The test results show that the "Feed-Forward Neural Networks (FNNs)" built using 20 neurons, the root mean square error (RMSE) predicted by the results is about  $\pm 5.21$  mm/yr can effectively predict the degree of subsidence of the southwest coastal strata., Applying Adaptive Fuzzy Neural Network System (ANFIS) to estimate the wind predict of the wind system in winter and summer for two consecutive years. Based on the measured data of the wind power system and the wind speed predict data of the Central Meteorological Administration, an adaptive fuzzy neural network prediction model is established. Then compare the power predicted by the adaptive fuzzy neural network with the actual measured power. By comparing the error iteration, adjust the network architecture and parameters until convergence. According to the research results, no matter what kind of attribution function (triangular, trapezoidal, bell and Gaussian) is used to perform wind predicting, the difference of the error obtained is not high, and the error quickly converges to a stable value. In the future, the established model can be used to accurately estimate the power generation capacity of wind turbines at any site, which can then be used as an assessment of the power generation cost and economic benefits of the wind power system [18-19].

Combining the empirical mode decomposition method and neural network, the data is decomposed into several intrinsic mode functions by the empirical mode decomposition method, and the decomposed signal is The method of predicting one day a day is combined with Neural Network (NN) to analyze the time series data of stock prices and propose a new method to analyze signals to evaluate the results of predictions, which can be used as a basis or indicator for future stock price estimation. The research results show that the stock price is better than the results of a single BPN prediction model in terms of the five evaluation indicators MAPE, RMSE, MAD, DS, and CD used in the combination of EMD and BPN prediction models. Therefore, EMD removes noise from the original data and after analyzing the phase characteristics, it can effectively improve the accuracy of the BPN model prediction.

### **3. Research methods**

### 3.1. The demand of industries

The IC substrate industry is related to the IC manufacturing industry. Therefore, in this study, the five manufacturing and related industries are selected as the following indicators.

#### 3.1.1. Electronic manufacturing sales volume index

This index is a measure of the relative change in product sales volume in a certain period and the base period. The calculation formula 1 is as follows:

$$\begin{aligned} S_{oi} &= \frac{\sum \frac{Q_i}{Q_o} * Q_o P_o}{\sum Q_o P_o} * 100 \\ &= \sum \frac{Q_i}{Q_o} * W_o * 100. \end{aligned} \quad (1)$$

Among them,  $W_o = \frac{Q_o P_o}{\sum Q_i P_i}$ .

$Q_i$ : production volume in the calculation period,  $Q_0$ : production volume in the base period,  
 $W_0$ : base period sales value weight,  $P_0$ : base period sales unit price.

### 3.1.2. Electronic manufacturing production index

This index is a measure of the relative changes in the production volume of the overall manufacturing sector at a certain time and base period. Formula 2 as follows:

$$I_{oi} = \frac{\sum Q_i P_o}{\sum Q_o P_o} * 100 \quad (2)$$

Among them,  $Q_i$ : production volume in the calculation period,  $Q_0$ : Base period production volume,  $P_0$ : Unit price of net production value in the base period.

### **3.1.3. Computer communication and audio-visual electronic product index:**

Since computer communication and audio-visual electronic products account for a high proportion of the final products of IC products. This index is an index that measures the relative changes in the production of computer communications and audio-visual electronic products at a certain time and base period, and its formula is the same formula 2.

### **3.1.4. The production index of electronic components:**

Electronic components must be used in all kinds of electronic products, and IC downstream products also need to use various electronic components. Therefore, the relationship between the electronic component industry and IC products is also considered in this research. And the relevant indicators of the electronic components industry, the electronic component production index is used, and the calculation formula 2.

### **3.1.5. The output of packaged IC:**

Substrate is the chip that provides load-bearing and heat-dissipating functions during the assembly process. In response to the assembly requirements of high pin count, high performance and heat dissipation, the carrier board carries IC chips and signal connections are currently the most costly Effective way, and the number of pins is increased from peripheral pins to area array pins, which can greatly increase the number of pins and reduce the volume of the overall package. Therefore, the output of packaged products also needs to be considered.

### **3.1.6. The downstream demand**

In recent years, consumer electronics products are quite developed, which has driven the demand for optoelectronics and semiconductors. According to a survey, more than 70% of the downstream products of IC substrates in the region are used in computers and peripheral equipment, and about 20% are used in communication products. Among the key components that make up such electronic products, IC products are the primary essential component. From this data, it is known that half of the output of IC products comes from information hardware products, and its main products are notebook computers, desktop computers, motherboards, servers, picture tube monitors, LCD monitors, mobile phones, The total output value of digital cameras and other products has accounted for nearly 90% of the total output value of the information hardware industry. With the popularization of mobile communication systems and the development of mobile products, the digital age has driven the consumer market for audio-visual entertainment products, and the increase in demand for related equipment has led to the development of related manufacturing industries.

- **TV monitor sales volume:**

The largest application of the soft IC carrier is in the packaging of display driver ICs, so the maturity of the IC carrier is relatively directly related to the growth of the flat-panel display industry. In general, the flexible IC carrier board driver IC package is used in higher-end displays, especially in the thin-circuit driver IC package, which has advantages. Due to the gradual development of the display industry, and the world's top three large-size displays product supplier countries therefore have strong demand for IC substrates, so the sales volume of TV monitors is included as one of the factors that affect downstream demand.

- **Sales volume of notebook computers:**

In recent years, due to the advancement of science and technology, the types of electronic products used by consumers have been changed, resulting in the current consumer electronic products becoming more powerful, thinner and lighter, and more user-friendly and convenient in use. Modern notes functions, computing power and price of notebook personal computers (NB PCs) have been compared with those of desktop computers.

In addition, the notebook computer has wireless transmission and the product characteristics of light

and easy to carry, which brings a lot of convenience to the user, making the notebook computer gradually replaced desktop computers. Whether it is a portable computer or a desktop computer, it is necessary to apply the related products of the IC substrate. Therefore, the sales volume of the notebook computer is included as one of the factors that affect the downstream demand.

- **Mobile phone sales:**

At present, smart phones are the most popular among consumers in the world. Judging from the trend of the smart phone market in recent years, with the demand of the consumer market, the display quality of mobile phones has gradually improved. Functional requirements are increasing. The design of mobile phones needs to consider the diverse requirements of lightness and thinness, integration of multimedia, and multi-function. In contrast, the number of pins on the design of the driver IC will increase significantly for functionality, and the line spacing will be slow. Slowly shrink to meet market demand for products. Therefore; high-density flexible IC carrier board bonding technology will be the mainstream of the IC package in the future, so the sales volume of mobile phones will be included as one of the influencing factors.

- **Tablet sales:**

Since tablet PCs need to use related products with flexible IC substrates, this study will use the sales volume of tablet PCs as one of the factors affecting IC products.

- **Sales volume of flat panel displays**

At present, the technology of flat-panel displays is becoming more mature, and the functions are becoming more diversified, which has changed the concept and habits of modern consumers in the use of traditional displays in the past. Because such products must also be equipped with a soft IC carrier board to be able to signal send to the screen, so the sales volume of flat-panel displays is included as one of the influencing factors.

### 3.2 Grey relational method

The main function of gray correlation analysis is to measure the degree of correlation between different sequences. Gray correlation analysis is a quantitative analysis of the dynamic development process of a system. It is based on the degree of similarity or difference in the development situation between factors. To measure the degree of correlation between factors, the closer the development trend, the greater the correlation between factors. When the amount of data is insufficient or cannot meet the specific allocation, the gray correlation analysis is a quite practical analysis method to analyze the relationship between factors. It is more practical than the regression model or econometric model commonly used in traditional statistics. This model as Fig. 4 has the following Features: (1) Grey relational analysis does not require a large amount of data (2) It does not need to assume that the data of the series of comparisons confirm to a specific distribution (3) The established model is a non-functional sequence model (4) The calculation method is simple.

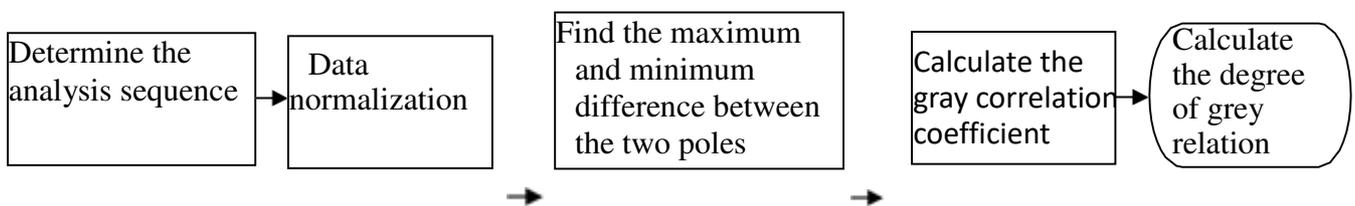


Fig. 4 Grey relational calculation flow chart

This study will use the grey relational analysis model to screen the most suitable factors to predict the production. The main steps are as follows:

**Step 1: Determine the analysis sequence**

Suppose the original sequence is:

$$X_i(k) = [X_i(1), X_i(2), \dots, X_i(n)], \quad (3)$$

Among them:  $i=1,2,\dots,m \in \mathbb{N}$ , representing a total of  $m$  sets of numbers (influence factors),  $k=1,2,\dots,n \in \mathbb{N}$ , representing that each series contains  $n$  factors (the number of data). Therefore, the expansion of Equation 3.5 can be written as:

$$\begin{aligned} X_1(k) &= [X_1(1), X_1(2), \dots, X_1(n)] \\ X_2(k) &= [X_2(1), X_2(2), \dots, X_2(n)] \\ &\dots \\ X_m(k) &= [X_m(1), X_m(2), \dots, X_m(n)] \end{aligned}$$

Define the group number sequence as the reference number sequence as  $X_0(k)=[X_0(1), X_0(2), \dots, X_0(n)]$ , the purpose is the  $m$  group number sequence associated objects, so  $m$  sets of sequence are called comparison sequence.

### Step 2: Data pre-processing

Since the units used by each factor in the data are not the same, in order to avoid the phenomenon of data extremes, the data must be normalized first, that is, all the data must be converted into the same interval. Generally speaking, the action of normalization is to convert all data into values between (0,1), which can be divided into the following three ways:

- The type of hopefulness: when I hope that the bigger the goal, the better.

$$X_i^*(k) = \frac{X_i^{(0)}(k) - \min[X_i^{(0)}(k)]}{\max[X_i^{(0)}(k)] - \min[X_i^{(0)}(k)]} \quad (4)$$

- The type of small hope: when the goal is as small as possible.

$$X_i^*(k) = \frac{\max[X_i^{(0)}(k)] - X_i^{(0)}(k)}{\max[X_i^{(0)}(k)] - \min[X_i^{(0)}(k)]} \quad (5)$$

- The type of vision: hope that the goal is between the maximum and the minimum.

$$X^*(k) = 1 - \frac{|X_i^{(0)}(k) - OB|}{\max[X_i^{(0)}(k)] - \min[X_i^{(0)}(k)]} \quad (6)$$

$$\max \left\{ \max[X^{(0)}(k)] - OB, OB \min[X^{(0)}(k)] \right\}$$

among them,

$X_i^*(k)$ : the value after the gray correlation is generated

$\max[X_i^0(k)]$ : the maximum value of the factor in the sequence  $\min[X_i^0(k)]$ : the minimum value of the

factor in the sequence OB: the selected value in  $X_i^0(k)$

### Step 3: Calculate the gray correlation coefficient

After the pre-processing of  $m$  sets of numbers, the  $m$  sets of numbers ( $X_i(k)$ ) can be converted into an  $m$ -dimensional array  $A$ , and the array each column in  $A$  subtracts the reference number sequence ( $X_0(k)$ ) and takes the absolute value, and the array  $A$  can be converted into an array  $\Delta$ . Each element in the array  $\Delta$  is denoted as  $\Delta_{oi}$ , where the largest element is  $\Delta_{\max}$  and the smallest element is  $\Delta_{\min}$ . And define the gray correlation resolution coefficient as  $\rho$ , the main function is the comparison between the background value and the object to be tested, and the value range is  $[0, 1]$ . The gray correlation coefficient  $\xi$  is defined as:

$$\xi_i(k) = \frac{\Delta_{\min} + \rho \Delta_{\max}}{\Delta_{oi} + \rho \Delta_{\max}} \cdot \forall i=1,2,\dots,m \cdot k=1,2,\dots,n. \quad (7)$$

$X_0(k)$  is the reference sequence,  $X_i(k)$  is the comparison sequence,

$\Delta_{oi} = X_0(k) - |X_i(k)|$  represents the absolute difference between  $X_0(k)$  and  $X_i(k)$ .

$\Delta_{\min} = \min \min \Delta_{oi}(k)$  called the two-level minimum difference  $i, k$ .

$\Delta_{\max} = \max \max \Delta_{oi}(k)$  called the maximum difference between the two levels  $i, k$ .

#### Step 4: Grey Relation

The gray correlation degree can express the correlation coefficient value between the comparison series and the reference series, but if there are  $n$  factors, there will be  $n$  gray correlation coefficient results, which will cause information to be scattered and will not be conducive to evaluation and comparison. Therefore, the gray correlation degree of each time of each comparison sequence must be concentrated to a point, and this point is called the gray correlation degree. The gray correlation degree can be calculated according to the difference of the weight, and there are two calculation methods, namely the equal weight correlation degree and the weighted correlation degree. But generally speaking, it is mainly based on the use of equal rights relevance.

### 3.3. Neural network method

Artificial Neural Network is a parallel calculation system, which includes hardware and software. It is composed of multiple artificial nerve cells also called artificial neuron, connects a large number of the artificial neurons form a network, imitating the ability of biological neural networks. In the current modern intelligent control field, artificial neural network has become the mainstream of modern intelligent control. Artificial neural network is a data processing system that imitates biological neural network. This computing system is composed of many components of the nodes (neurons) of the layers are generally divided into three layers: "input layer, hidden layer, and output layer". The input is called the Input layer, this layer has no neurons and only input values. The output is called the output layer. If there are only output values without neurons, then three neurons are included. The neurons between the input and output belong to the hidden layer, and the hidden layer is optional.

#### 3.3.1. The architecture of neural network

The neural network architecture type is a neural network that imitates biology. The entire network can be divided into three elements, namely neurons (processing units), layers, and networks. The description is as follows:

- **Processing Element:**

The processing unit or artificial neuron is the unit that basically composes the neural network. It mainly converts the relationship between the output value and the input value through a mathematical formula, and then uses this signal as the output result or becomes Input of other operands. There are three kinds of functions for the operation element to process data. The brief introduction is as follows:

- Summation Function:**

The function is to integrate the input variable data or the output of the previous layer of neurons with the connection weight value to become the input of the operand. Commonly used functions include weighted product sum and Euclidean distance.

- Activity Function:**

The integrated function value and the neuron state are integrated, usually directly using the integrated function output. The purpose of the conversion function is to calculate the output value of the integrated function into the output value of the processing unit through a conversion equation.

- Transfer Function:**

The output of the action function is converted into the output value of the processing unit, and the output value of the processing unit is calculated through a conversion equation. Commonly used functions include hard limit functions, linear functions, and non-linear number functions.

#### 3.3.2, backward neural network

The backward neural network is one of the more commonly used networks at present. The network architecture of the backward neural network is the Multilayer Perceptron (MLP), and the commonly used learning algorithm is the error backward algorithm. (Error Back Propagation, EBP), referred to as BP (Back Propagation) algorithm, combine this (MLP+EBP) is called Backpropagation Neural Network (BPN). Its architecture is shown as Fig.5:

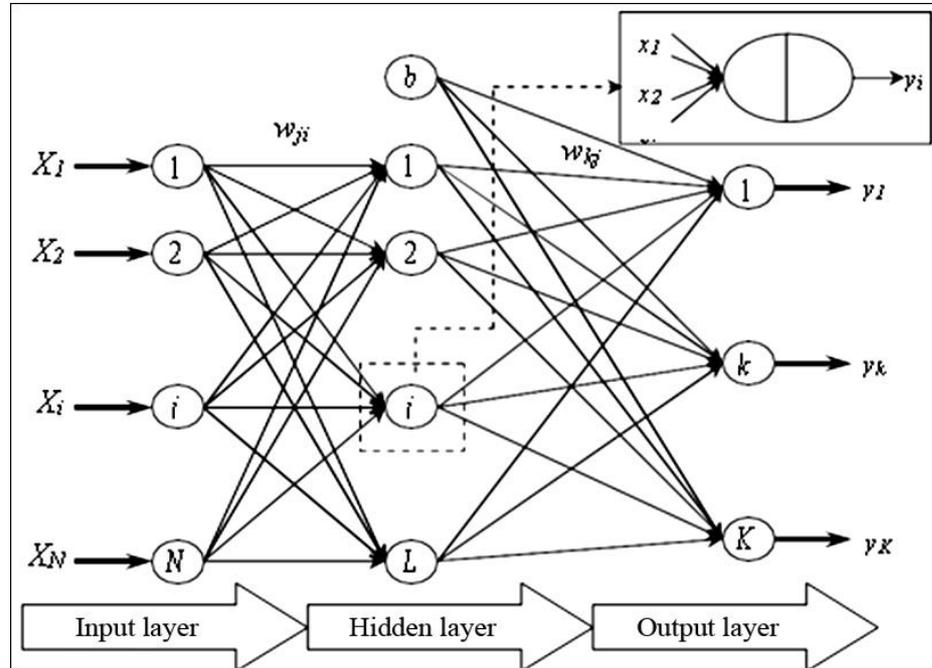


Fig. 5 Backpropagation neural network architecture

The basic principle model of the backward neural network as Fig. 6 is to use the concept of the Gradient Steepest Descent Method to minimize the error function. Generally, the learning process will proceed one training example at a time, until all learning and training examples are completed, that is, a learning round. A network can learn training examples repeatedly until the network learning reaches the effect of convergence.

The standard architecture of the backward neural network can be divided into two parts:

- **Data forwarding:** It means that the output value of the input layer is passed into the neurons of the hidden layer after the integrated function operation, and the neurons of the hidden layer are converted into output values through the conversion function through the value obtained in the integrated function. This output value pass the integration function to the output layer, and the output layer uses the value obtained in the integration function to convert it into an output value through the conversion function.
- **Backward propagation of errors:** The so-called error back propagation is to compare the value obtained from the output layer with the actual value, and then adjust the weight from the output layer to the hidden layer and the hidden layer to the input layer according to this error value.

The calculation steps of the backward neural network are as follows:

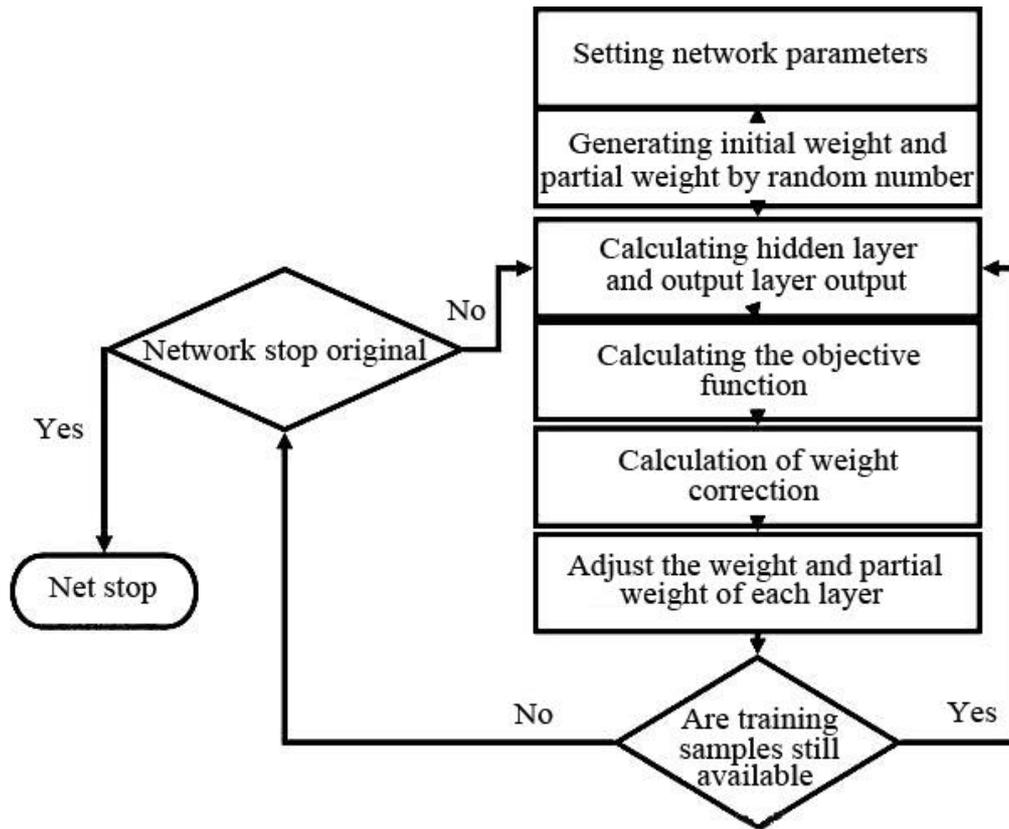


Fig. 6 Flow chart of backpropagation neural network

### Step 1: Set network parameter values and conversion functions

Because the range of input variables is very different, it may cause some input variables with smaller range to lose their effect, which makes the weight error of the network during training become larger. Therefore, it is necessary to normalize each input variable to change the value of the variable. The minimum and maximum values are mapped to the expected minimum and maximum values. This method is called interval response method. The formula is as follows:

$$X_{new} = \frac{X_{old} - \min(x_1, x_2, \dots, x_n)}{\max(x_1, x_2, \dots, x_n) - \min(x_1, x_2, \dots, x_n)} \times (D_{max} - D_{min}) + D_{min} \quad (8)$$

Among them,

- $X_{new}$ : normalized parameter data value,
- $X_{old}$ : the original parameter data value,
- $D_{max}$ : the expected maximum value,
- $D_{min}$ : the minimum value expected.

### Step 2: Initialize weights and partial weights

Before learning, it is necessary to initialize the weight of each neuron connected in the network, and set the initialized weight value to a very small random value. A too large weight can easily lead to the phenomenon of unit saturation and make the network error higher. Therefore, a small initial weight value can make the network easier to converge.

### Step 3: Calculate the hidden layer output

For each hidden layer neuron  $j$ , calculate the sum of the products from each  $X_i$  and the corresponding weight  $V_{ij}$   $Z_{in}$  in  $j$ , and transform  $Z_{in}$  in  $j$  into  $Z_j$  through the conversion function.

$$Z_{in_j} = V_{oj} + \sum_{i=0}^p V_{ij} * X_i$$

$$Z_j = f(Z_{in_j}) \quad (9)$$

**Step 4: Calculate the output of the output layer**

For each neuron  $k$  in the output layer, calculate the sum  $Y_{in\ j}$  from the product of each  $Z_j$  and the corresponding weight  $W_{jk}$ , and convert it into  $Y_k$  through a conversion function.

$$Y_{in\ k} = W_{ok} + \sum_{j=0} W_{jk} * Z_j$$
$$Y_k = f(Y_{in\ k}) \quad (10)$$

**Step 5: Calculate the weight correction value between the hidden layer and the output layer**

Calculate the difference  $\delta_k$  between each output value  $Y_k$  and the actual value  $T_k$ , and calculate the weight correction value  $\Delta W_{jk}$  between the hidden layer and the output layer and the weight correction value  $\Delta W_{ok}$  of the hidden layer partial weight according to  $\delta_k$ .

$$\begin{aligned} \delta_k &= (T_k - Y_k) f'(Y_k) \\ \Delta W_{jk} &= a \delta_k Z \end{aligned} \quad (11)$$

**Step 6: Calculate the weight correction value between the input layer and the hidden layer**

For each hidden layer neuron  $j$ , calculate the total error  $\delta_{in_j}$  value from the output layer, and calculate the error value  $\delta_j$  of each neuron according to  $\delta_{in_j}$ . Using  $\delta_j$ , the weight correction value  $\Delta V_{ij}$  between the input layer and the hidden layer and the weight correction value  $\Delta V_{oj}$  of the bias weight of the input layer can be calculated. Among them,  $\alpha$  is the learning rate of the network.

$$\delta_{in_j} = \sum_{k=1} \delta_k W_{jk}$$

$$\delta_j = \delta_{in_j} f'(Z_{in_j})$$

$$\Delta W_{ok} = \alpha \delta_k \quad (12)$$

**Step 7: Adjust and update the weights and partial weights of each layer**

$$\Delta V_{ij} = \alpha \delta_j X_i$$

$$\Delta V_{oj} = \alpha \delta_j \quad (13)$$

$$W_{jk}(new) = W_{jk}(old) + \Delta W_{jk} \quad (14)$$

$$W_{ok}(new) = W_{ok}(old) + \Delta W_{ok} \quad (15)$$

$$V_{ij}(new) = V_{ij}(old) + \Delta V_{ij} \quad (16)$$

$$V_{oj}(new) = V_{oj}(old) + \Delta V_{oj}$$

**Step 8: Test whether the network stop condition has been reached.** If the network stop condition is not met, go to step 2. If the network stop condition is met, the network learning is ended. Among them, the termination conditions reached by the network can be divided into four methods:

**1. Learning times:**

Specifying the network to complete the preset number of learning times can be used as one of the conditions for the network to end learning. This study uses this method as the network stop condition.

**2. Gradient method:**

The learning of the inverted transfer network is to change to the direction of the maximum slope. When the gradient is unchanged or the gradient change is small, the weight will not change at this time, and the learning will stop.

**3. Root mean square:**

When the root mean square error value of the network is less than a certain set convergence value, it means that the network has reached a certain degree of convergence, and then stop learning.

**4. Cross-Validation:**

The samples are divided into training and test data. One set is used as training samples and one set is used for testing. If the error level of training and testing is less than a certain set value at the same time, you can stop learning. If the training is good but the test is not good, it means over-learning; if the training is not good but the test is good, it means under-learning.

The data from January 2018 to December 2019 is used as a training example of the backward neural network to verify the accuracy of the prediction model.

**4. Experiment analysis**

**4.1 Evaluation Index**

In order to evaluate the prediction accuracy and prediction error performance of each model, this study adopted the following two evaluation indicators, namely the mean absolute percentage error (MAPE) and the mean absolute deviation (MAD). :

**1. Mean Absolute Percentage Error (MAPE)**

$$MAPE = \frac{1}{n} \sum_{t=1}^n \frac{|F_t - A_t|}{A_t}, \quad (17)$$

F<sub>t</sub>: predict value of period t  
A<sub>t</sub>: actual value of period  
t n: number of periods

## 2. Mean Absolute Deviation (MAD)

$$MAD = \frac{1}{n} \sum_{t=1}^n |F_t - A_t| \quad (18)$$

Among them,

$F_t$ : the predicted value of period  $t$

$A_t$ : the actual value of the  $t$  period

$n$ : number of periods

The evaluation method of the two indicators mentioned above is that the smaller the value is the better; the smaller the value will the higher the agreement between the estimated results in the predict mode and the historical data. The classification of the evaluation criteria can be roughly divided into four items, as shown in Table 1. In this study, in order to achieve an objective and fair evaluation standard, MAPE was chosen as the accuracy measure of the evaluation and prediction results.

Table 1 Classification of evaluation criteria

MAPE	Effect
<10%	Excellent prediction ability
10%~20%	Good prediction ability
20%~50%	The prediction ability is reasonable
>50%	Incorrect prediction

## 2 Application of grey relational analysis

The 18 factors related to COF packaging products are divided into three levels for discussion. However, due to the difference in the degree of impact on packaging production between the factors considered, it will affect the prediction results of the inverted neural network. To reduce the error of the prediction results, the gray correlation analysis will be performed on the various factors of the past historical data, and the factors with a high degree of correlation with the prediction standard will be selected as the basis for the selection of the input factors of the backward neural network. To improve the accuracy of the prediction. First, collect the gray correlation degree between 18 factors and the COF packaging production volume of the predict target, set the selection factor to be more than 0.7 correlation degree, and select a total of 9 factors as shown in Table 2. The 9 factors are unemployment rate (global), Export value, average domestic production, consumer price index, electronics manufacturing sales volume index, electronics manufacturing production index, electronic component production index, package IC production, notebook computer sales.

Table 2 The table of grey relation

Three levels	Impact factor	IC	
		Grey relational degree	Grey relational degree >0.7
Overall economy	unemployment rate (%)	0.66	
	Unemployment rate (global)%	0.76	△
	National economic growth rate %	0.58	
	Import value	0.68	
	Export value	0.73	△
	Average gross domestic product	0.75	△
	Consumer Price Index	0.73	△
	Domestic boom indicator indicator	0.59	
Industrial manufacturing	Electronics Manufacturing Sales Volume Index	0.74	△
	Electronics Manufacturing Production Index	0.73	△
	Computer communications and audio-visual electronic products	0.67	
	Electronic components production index	0.75	△
	Package IC output	0.73	△
Downstream demand	TV monitor sales	0.66	
	Laptop sales	0.73	△
	Mobile phone sales	0.64	
	Tablet PC sales	0.67	
	Flat panel display sales	0.65	

According to the gray correlation analysis of the above COF products, there are three levels of factors selected. Among them, there are four factors that have a relevance greater than 0.7 at the overall economic level and industrial manufacturing level, while there is only one at the downstream demand level. It can be seen that the production volume of COF products has a greater correlation with changes in the overall economic and industrial manufacturing levels. The downstream demand level will be affected by the time of sales inventory processing, so it appears initially that the correlation is low.

### 3 Prediction of backward neural network

The BNN model has different parameters (number of hidden layer units, learning rate, inertia factor) are set to distinguish different modes, and then separate the data is input to the BNN to start learning and training. The method of correcting the weight is the steepest descent method. After the training is completed, the training samples are used to test the prediction performance of the network as Table 3.

### 1. Parameter setting

Table 3 Types of neural training parameter setting table

Input layer	X1: Unemployment rate (global). X2: Export amount. X3: Average domestic production. X4: Consumer price index. X5: Electronic Manufacturing Sales Volume Index X6: Electronic Manufacturing Production. Index X7: Electronic Components Production Index. X8: Integrated IC Production. X9: notebook computer sales. A total of nine input units.
Output layer	Y1: IC production, a total of one output unit.
Hidden layer	There is one hidden layer, and the number of neurons in the hidden layer is 5, 10, and 15 respectively.
The initial learning rate	0.1, 0.5, 0.9
The inertia factors	0.1, 0.5, 0.9
The number of iteration cycles	10,000 times; and 30,000 times

### 2. Analysis of the results of BPN training

Use the data from January 2018 to December 2019 as the training sample for training, use the data from January to December 2017 as the test sample, and then perform 10,000 and 30,000 iteration cycles for the parameters of the same conditions. After comparing the training results of various setting parameters, the error results of the test example as Fig. 7 and the training example as Fig. 8 show that the error value of training 30,000 times is lower than the error value of training 10,000 times.

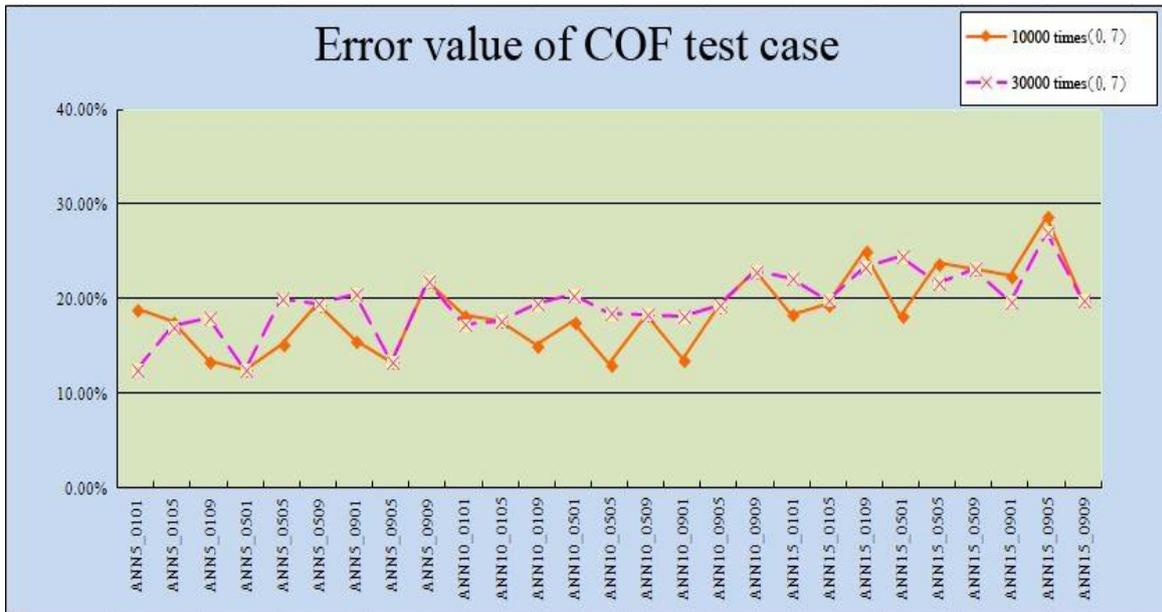


Fig. 7 Test example error

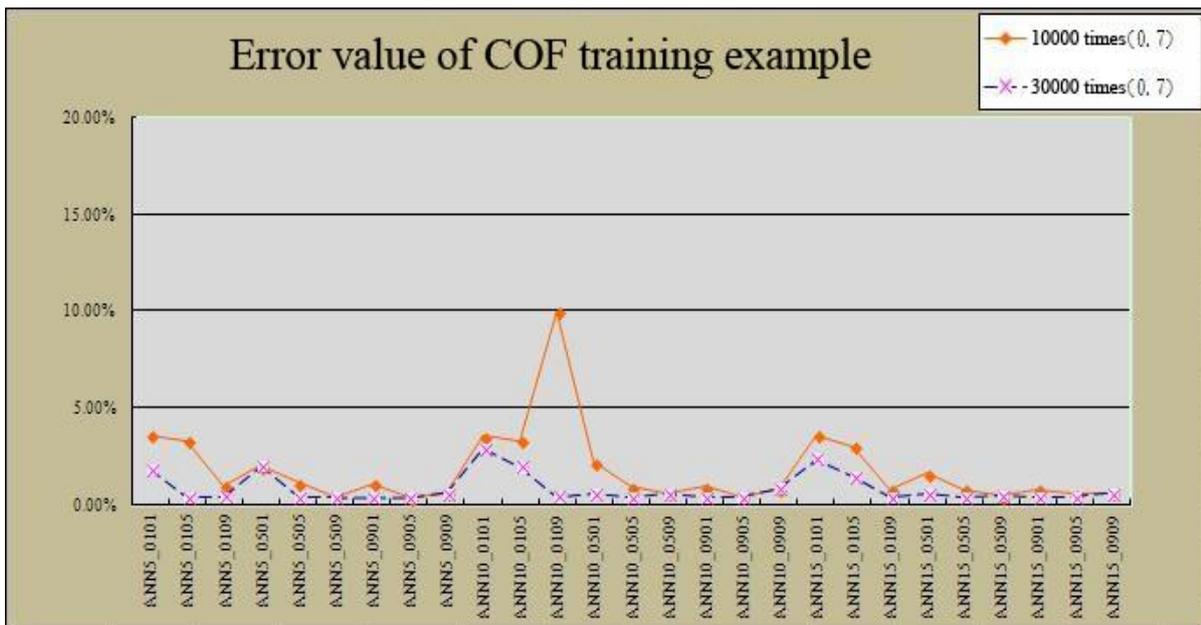


Fig. 8 Trend chart of training example error

Fig. 9 is a comparison between the actual value of the training example and the predicted value of BPN. It can be found that the predicted result of BPN is very close to the actual value, and the difference is not high. Fig. 10 is a comparison of the predicted value and actual value of the backward neural network of the test example

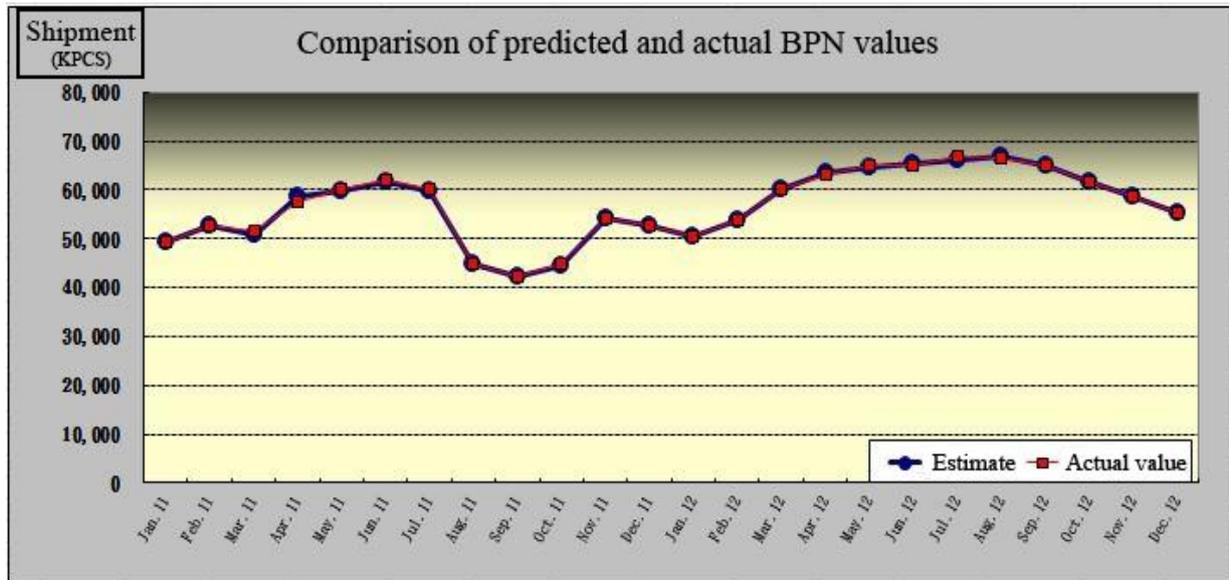


Fig. 9 Comparison of the predicted value and actual value of the backward neural network of the training example.

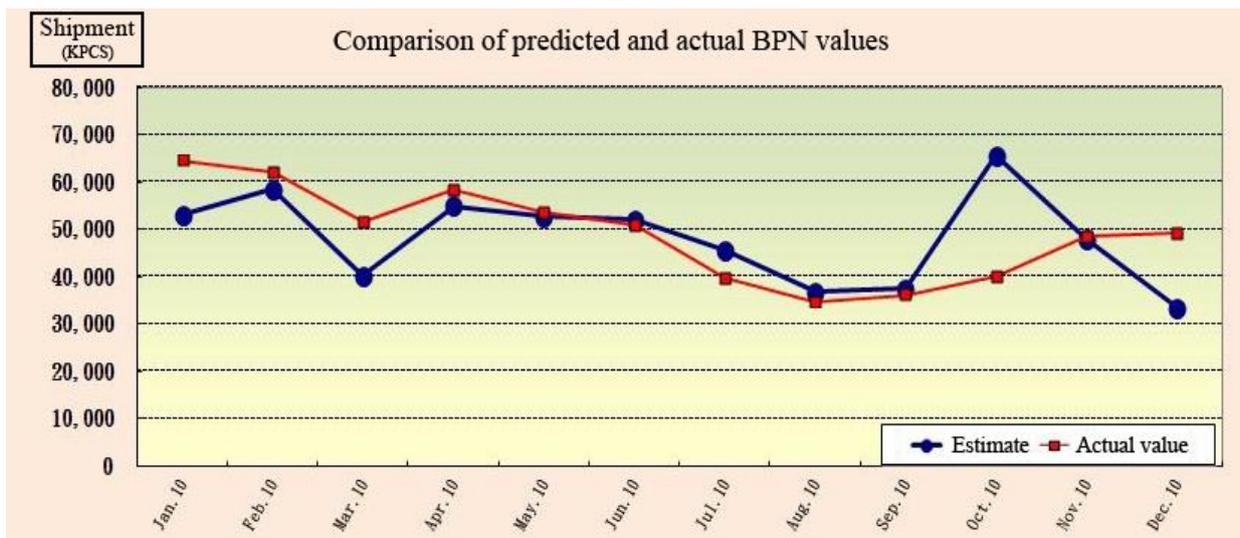


Fig. 10 Comparison of the predicted value and actual value of the backward neural network of the test example

The next step is to compare MSE (Mean Squared Error). It can be found that the result of training 30,000 times is more convergent than the result of training 10,000 times, as shown in the following table 4. The training error results show that training 30,000 times, hidden layer 5, Learning rate 0.9, inertia factor 0.5 is the best result, MSE finally converges to 0.00001, and the training error is the lowest 0.37%, the test error is the second lowest 13.33%, so this group of parameters is selected for subsequent analysis and research.

Table 4 The training results of IC output BPN

IC						
Training condition	10000(0.7)			30000(0.7)		
	MSE	Training example error	Test example error	MSE	Training example error	Test example error
ANN5_0101	0.00061	3.6%	18.9%	0.00035	1.8%	12.4%
ANN5_0105	0.00052	3.3%	17.6%	0.00011	0.4%	17.0%
ANN5_0109	0.00005	1.0%	13.4%	0.00001	0.4%	17.9%
ANN5_0501	0.00014	2.0%	12.4%	0.00001	2.0%	12.4%
ANN5_0505	0.00005	1.1%	15.1%	0.00001	0.4%	19.9%
ANN5_0509	0.00001	0.4%	19.4%	0.00001	0.4%	19.4%
ANN5_0901	0.00005	1.1%	15.6%	0.00001	0.4%	20.4%
ANN5_0905	0.00001	0.4%	13.3%	0.00001	0.4%	13.3%
ANN5_0909	0.00001	0.6%	21.9%	0.00002	0.6%	21.9%
ANN10_0101	0.00055	3.5%	18.2%	0.00034	2.9%	17.2%
ANN10_0105	0.00047	3.3%	17.6%	0.00012	1.9%	17.7%
ANN10_0109	0.00005	9.9%	15.1%	0.00001	0.4%	19.4%
ANN10_0501	0.00015	2.2%	17.5%	0.00002	0.6%	20.4%
ANN10_0505	0.00005	0.9%	13.1%	0.00001	0.4%	18.4%
ANN10_0509	0.00001	0.6%	18.3%	0.00001	0.6%	18.3%
ANN10_0901	0.00005	0.9%	13.6%	0.00001	0.4%	18.1%
ANN10_0905	0.00001	0.4%	19.23%	0.00001	0.42%	19.2%
ANN10_0909	0.00001	0.8%	22.88%	0.00003	0.86%	22.9%
ANN15_0101	0.00054	3.6%	18.32%	0.00020	2.38%	22.1%
ANN15_0105	0.00040	3.0%	19.40%	0.00007	1.46%	19.7%
ANN15_0109	0.00003	0.81%	25.01%	0.00001	0.43%	23.27%

Based on the training example and prediction values of the selected parameters (number of hidden layer neurons=5, learning rate=0.9, inertia factor=0.5), the MAPE shown in Table 5 is calculated as 1.1%, refer to MAPE for the evaluation criteria, the value of less than 10% indicates that the predictive ability is excellent. It can be seen that the predictive model of this group of parameters of the inverted neural network is suitable used to predict the production volume of IC products. The MAPE shown in Table 6 is calculated as 12.4%, refer to MAPE

for the evaluation criteria.

Table 5 The COF output training predict value and actual value MAPE

<b>MONTH</b>	<b>MAPE</b>	<b>MAD</b>
Jan-11	0.1%	46,700
Feb-11	0.1%	40,800
Mar-11	0.4%	168,800
Apr-11	0.8%	340,300
May-11	0.7%	125,400
Jun-11	0.7%	68,500
Jul-11	0.6%	46,400
Aug-11	0.6%	12,400
Sep-11	0.5%	2,300
Oct-11	0.5%	11,700
Nov-11	0.5%	19,800
Dec-11	0.4%	3,000
Jan-12	5.4%	68,500
Feb-12	2.8%	64,900
Mar-12	1.9%	43,900
Apr-12	1.5%	67,600
May-12	1.3%	78,200
Jun-12	1.2%	65,700
Jul-12	1.2%	119,700
Aug-12	1.1%	61,700
Sep-12	1.0%	13,500
Oct-12	0.9%	3,800
Nov-12	0.8%	2,900
Dec-12	0.7%	2,400
Add up	26%	1,479,300
Average predict error	1.1%	61,600
Standard deviation	1.1%	73,500
Maximum	5.4%	340,300
Minimum	0.1%	2,300

Table 6 The COF production test predict value and actual value MAPE

<b>Month</b>	<b>MAPE</b>	<b>MAD</b>
Jan-10	17.7%	11,417,690

Feb-10	11.7%	7,527,400
Mar-10	15.3%	6,361,200
Apr-10	12.9%	2,422,900
May-10	10.6%	614,900
Jun-10	9.2%	295,900
Jul-10	10.0%	891,500
Aug-10	9.5%	384,600
Sep-10	8.9%	201,700
Oct-10	14.4%	2,585,900
Nov-10	13.2%	287,700
Dec-10	14.8%	1,351,800
Add up	148%	34,343,200
Average predict error	12.4%	2,861,900
Standard deviation	2.8%	3,632,200
Maximum	17.7%	11,417,700
Minimum	8.9%	201,700

#### 4.4 Comparison of predicted and actual values of COF production

Then use the training weights of the best parameters selected by the above training results (number of hidden layer neurons = 20, learning rate = 0.5, momentum = 0.9) as the verification condition, and apply it. The actual production data from January to April is used as a verification example, and the calculation results are as follows: The calculated prediction result has a MAPE value of 23.9% as shown in Table 7. Although the value is higher than the result of the training example from 2018 to 2019, it is between 20% and 50% according to the evaluation standard of MAPE. Time, indicating that the predicting ability is still within a reasonable range.

Table 7 The comparison of MAPE for COF production

Month	MAPE	MAD
Jan-13	41.8%	23,770,200
Feb-13	22.5%	12,733,900
Mar-13	17.4%	5,699,300

Apr-13	14.0%	2,018,700
<b>Add up</b>	96%	44,222,200
<b>Average predict error</b>	23.9%	11,055,500
<b>Standard deviation</b>	12.4%	9,571,400
<b>Maximum</b>	41.8%	23,770,200
<b>Minimum</b>	14.0%	2,018,700

The issue December 2019 annual chassis point settlement of orders in advance has affected the order volume in January 2020. Therefore, the January data is shaved and recalculated. The calculation result is 17.9%. Refer to the assessment of MAPE the standard, whose value is between 10% and 20%, indicates that the predictive ability is excellent. It can be seen that, except for the large difference in January, the error value from February to April is about 18%. The predict value is quite close to the actual output value, as Fig.11, this result be seen that the prediction results made by the application of the backward neural network can be used as a reference basis for the production of COF products in the future.

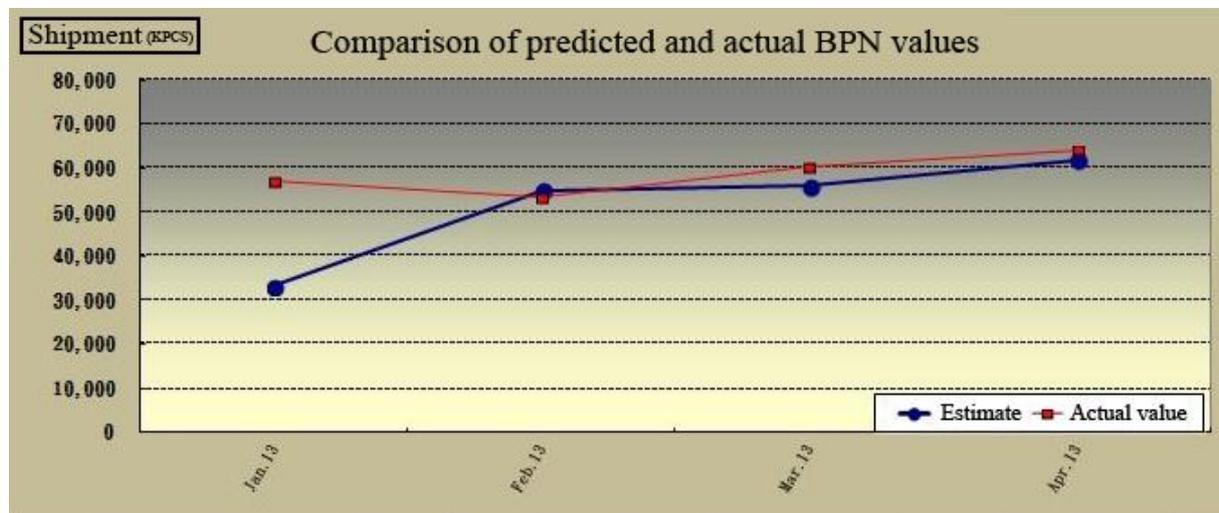


Fig. 11 Comparison of predicted and actual values of COF production verification results

## 5. Conclusions

This paper proposed the prediction of IC output with a BNN architecture. Whether an accurate and simple prediction model can be constructed to provide enterprises with a scientific and modern management model for output prediction. The production manager can plan and execute material preparation operations in advance, and then help companies reduce

costs, increase more profits and increase their competitiveness. In the research process, different parameters and iteration times were set for the prediction model, and the output was predicted and analyzed. Then the average absolute error was used to select and evaluate which is the more effective prediction model. The results showed that the training prediction error was less than 11%, the average error for three months can also be less than 13%, which represents the prediction model of the backward neural network, which is applicable in the predict of IC product production. The results and contributions are as the followings:

1. The data collected is based on the factors of IC industries, and the different indicators are divided into three levels. First, the gray correlation analysis is used to screen the factors with higher correlation, and the factors with lower correlation are deleted, and the prediction result can improve the accuracy of the prediction model.

2. When establishing the backward neural network prediction model, set different parameters to verify and perform two different iteration experiments with different parameters to get the best value of MSE, and the result shows when the number of iterations reaches 30,000 times, the MSE value is more convergent, and the prediction result is more accurate than 10,000 times. Future work, we can prepare appropriate materials in advance to reduce the company's inventory of raw materials. It can provide customers with accurate and fast delivery times, and can also provide more competitive delivery times than competitors in the industry.

#### **Ethical approval**

No need ethical approval.

#### **Funding details**

No funding.

#### **Conflict of interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### **Informed Consent**

No need informed consent.

#### **Authorship contributions**

Zhifang Liu: Conceptualization, Methodology, Validation, Investigation, Writing, Funding acquisition, Formal analysis, Software, Resources, Visualization, Funding acquisition.  
Methodology, Validation, Writing -Review & Editing, Supervision.

#### **Reference:**

1. Zhang Z, Ni G, Xu Y (2020) Trajectory prediction based on AIS and BP neural network. 2020 IEEE 9th Joint International Information Technology and Artificial Intelligence Conference (ITAIC).
2. Chang W, Wu S, Lin B (2018) Intelligent Prediction Platform of Lathe Machine Based on

- Back Propagation Neural Network. 2018 IEEE International Conference on Advanced Manufacturing (ICAM).
3. Li J, Li X, Yu L (2018) Ship traffic flow prediction based on AIS data mining. 2018 33rd Youth Academic Annual Conference of Chinese Association of Automation (YAC).
  4. Abdellah A, Mahmood O, Paramonov A, Koucheryavy A (2019) IoT traffic prediction using multi-step ahead prediction with neural network. 2019 11th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT).
  5. Budhani N, Jha C, Budhani S (2014) Prediction of stock market using artificial neural network. 2014 International Conference of Soft Computing Techniques for Engineering and Technology (ICSCTET).
  6. Tu Y, Ai G, Tao X, Fang W (2011) Application of neural network in predication model of flotation indicators. 2011 3rd International Conference on Computer Research and Development.
  7. Peng C, Wang S, Liu S, Yang Y, Liao B (2020) Artificial Neural Network Application to the Stroke Prediction. 2020 IEEE 2nd Eurasia Conference on Biomedical Engineering, Healthcare and Sustainability (ECBIOS).
  8. Zou J, Chen Z, Cheng Z, Liao J (2020) Research on Collision Avoidance of Navigation Marks Based on Neural Network. 2020 7th International Forum on Electrical Engineering and Automation (IFEEA).
  9. Zhu X (2010) Application of Composite Grey BP Neural Network Predicting Model to Motor Vehicle Fatality Risk, 2010 Second International Conference on Computer Modeling and Simulation, 2010, Volume: 2
  10. Guo N, Li L, Xiang L, Bose S, Shen G (2020) What if AI Fails: Protection against Failure of AI-Based QoT Prediction. 2020 Optical Fiber Communications Conference and Exhibition (OFC).
  11. Naidu G, Govinda K (2018) Bankruptcy prediction using neural networks. 2018 2nd International Conference on Inventive Systems and Control (ICISC).
  12. Lee K, Lee Y, oon H (2019) Development of Edge-based Deep Learning Prediction Model for Defect Prediction in Manufacturing Process. 2019 International Conference on Information and Communication Technology Convergence (ICTC).
  13. Chharia A, Gupta M, Gupta S, Gupta A, Arya V (2012) A Comparative Study of AI Techniques for Failure Risk Prediction in Lightning Surge Protection. 2012 Fifth International Conference on Emerging Trends in Engineering and Technology.
  14. Ding H, Gao R, Isaksson A, Landers R, Parisini T, Yuan Y (2020) State of AI-Based Monitoring in Smart Manufacturing and Introduction to Focused Section. IEEE/ASME Transactions on Mechatron. Volume: 25, Issue: 5.
  15. Lin Y, Liu X, Su Li S, Lin Q (2020) Research on the Back Propagation Neural Network Haze Prediction Model Based on Particle Swarm optimization. 2020 International

Conference on Computer Engineering and Application (ICCEA).

16. Zong P, Wang Y, Xie F (2018) Embedded Software Fault Prediction Based on Back Propagation Neural Network. 2018 IEEE International Conference on Software Quality, Reliability and Security Companion (QRS-C).
17. Joshi S, Borse M (2016) Detection and Prediction of Diabetes Mellitus Using Back-Propagation Neural Network. 2016 International Conference on Micro-Electronics and Telecommunication Engineering (ICMETE).
18. Zhang X, Ye Z, Yan L, Wang C, Wang R (2018) Security Situation Prediction based on Hybrid Rice Optimization Algorithm and Back Propagation Neural Network. 2018 IEEE 4th International Symposium on Wireless Systems within the International Conferences on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS-SWS).
19. Zhao Y, Zou Z, Wang S (2015) A Back Propagation Neural Network Model based on kalman filter for water quality prediction. 2015 11th International Conference on Natural Computation (ICNC).