International Education Studies; Vol. 15, No. 6; 2022
ISSN 1913-9020 E-ISSN 1913-9039
Published by Canadian Center of Science and Education

A Comparative Study on the Forecast Models of the Enrollment Proportion of General Education and Vocational Education

Qiongqiong Chen[1]

[1] Vocational and Technical Teachers College, Hubei University of Technology, Hubei, China

Correspondence: Qiongqiong Chen, No. 28, Nanli Road, Hongshan district, Wuhan, Hubei, China

Received: June 3, 2022 Accepted: August 11, 2022 Online Published: November 28, 2022

doi:10.5539/ies.v15n6p109 URL: https://doi.org/10.5539/ies.v15n6p109

Abstract

Predictive research on the enrollment proportion of general education and vocational education is crucial to optimizing the regional talent structure and industrial structure adjustment. The reasonable enrollment proportion of general education and vocational education also plays an important role in the adjustment of the overall employment structure and the development of the regional economy. Therefore, it is imminent to seek a more accurate and reliable prediction model of the enrollment proportion of general education and vocational education. Based on the grey prediction model, exponential smoothing model, ARIMA model and BP neural network, and with the data of the enrollment proportion of all regions in China from 2010 to 2018 as the data sample, the enrollment proportion of each region in 2019 is predicted. By comparing the predicted values with the real values, it is found that the exponential smoothing model has the best accuracy and stability for the enrollment proportion of general education and vocational education forecast. Exponential smoothing model is used to predict the number of high school enrollment and vocational education enrollment, which is of great significance to ensure the reasonable structure of human resources in various regions and promote the coordinated development of the education system.

Keywords: the enrollment proportion, forecast models, the data sample, best accuracy and stability, talent structure

1. Introduction

1.1 Background

In the human capital framework, general education is considered to create “general human capital”, while vocational and technical education creates “specific human capital” (Tilak, 2018), the former cultivates flexible and complex high-quality talents, and the latter cultivates specific job-related technical skills talent, both are important, therefore, many countries will rationally divide the enrollment proportion of general education and vocational education. When designing the enrollment policy, the government fully considers the needs of economic development for skilled operators, and has a certain number of students, thus forming the economic basis for the existence of the “proportion of enrollment in general education and vocational education”. It is crucial to the optimization of the regional talent structure and industrial structure adjustment, and is more conducive to improving the appropriateness of the general job structure and the level of regional economic development. In order to cope with the optimization and upgrading of regional industrial structure and the transformation of economic development mode. Therefore, it is imminent to seek a more accurate and reliable prediction model of the enrollment proportion of general education and vocational education.

1.2 Existing Research and Problems

Regional differences determine the enrolment rates of general education and vocational education in senior high schools. The regional industrial structure, population status and urbanization process are different, and the demand for talent structure is also different. Underdeveloped areas need more technical and skilled talents to promote the rapid development of the local economy. The technical skills that these workers need to master need to be achieved through vocational education, and the scale of vocational education enrollment will expand. The developed regions are in the stage of economic transformation, innovation-driven and technological breakthrough, and the level of economic development is relatively high. Therefore, more innovative talents and senior technical talents are needed, and the structure of talent demand needs to be readjusted. Choi used the probability model and ordinary least squares model, concluded that the employment status and wage level of directly entering the labor market after secondary vocational education are better than those of higher education (Choi, 2021). From the perspective of South Korea's educational path and the influence of the labor market, the development of vocational education is the focus. Hoger believes that the current general occupation has become quite meaningless under the influence of industrial technological change (Hoger, 1990) and its impact on the nature of work. China is a typical country with a divided system. The coordinated development of general education and vocational education in senior high school is related to the increase of gross rate of school attendance, the adjustment of talent structure and the transformation of industrial structure. Now that China is facing complex economic development and regional development, "the proportion of enrollment in general education and vocational education is roughly the same" is increasingly being questioned. It is necessary for the state to make overall planning for regular high school education, but the specific implementation of each region should be based on the differences in the level (Peng, 2007) of regional economic development to formulate the enrollment proportion of general education and vocational education. It has become a research consensus. The prediction of the enrollment proportion of general education and vocational education, in advance can be used as an early warning system to facilitate timely adjustments in various regions.

Since the reform and opening up, China has begun to adjust the structure of secondary education, which also has the related research of the enrollment proportion of general education and vocational education. The trend of the enrollment proportion of general education and vocational education and its accurate prediction will inevitably become an important issue that cannot be avoided in the structure of general and vocational education. Based on the relationship between the types of senior high schools attended by Indonesian youth and their subsequent labor market outcomes, Tuma studied the differences in the treatment of graduates in different education types (Tuma, 1994) and their social rate of return, which provides an important reference for the Indian government to expand the scale of vocational education enrollment. Liu used the linear network model of the MATLAB neural network toolbox to predict the scale of vocational education in Shanghai from 1980 to 2005 (Liu, 2005). The prediction of the scale of vocational education in this study lacks forward-looking, because the paper does not point out the trend of the development scale of vocational education in the future. Hussar *et al.* predicted that the number of high school graduates between 2022-23 will drop by 2% compared to the number of high school graduates in 2009-2010 (Hussar& Bailey, 2007), based on data from the U.S. Department of Education and the National Bureau of Statistics. At this stage, most of the research on the enrollment proportion of general education and vocational education remains in the theoretical explanation, and the prediction results are ambiguous, and the quantitative prediction model has not been really introduced in the research.

The research on the enrollment proportion of general education and vocational education has gradually begun to introduce prediction models, mainly focusing on the prediction of time-series data such as the population of appropriate age, the scale of education, the number of enrollments, the scale of teachers and students, the output of technical and skilled talents, and industrial demand. Hager utilized a predictive model and predicts future enrollment at the Naval Graduate School based on past recruits. The model was applied to predict college student enrollment and included a comparison of predicted results with actual student enrollment (Hager Jr., 1970). Based on data collected from a questionnaire survey of 2,453 public high school seniors in Dade County, Florida, HP Tuckman used ordinary least squares regression (Tuckman, 1971) to analyze the employment and future earnings of high school students. Maxim developed a computer system. This system describes the LIONS (Labor Industrial Occupational Demand System) and is designed to predict New Jersey's demand for various occupational skills based on regional economic development, and this research mainly uses the forecasting model to explore the matching degree between skilled talents and regional economic development (Maxim, 1973). Valentine M A used descriptive analysis to predict the success of students with disabilities in vocational education programs in West Virginia from 1981-82 (Valentine, 1985). Davidson took Hardin-Simons University (HSU) enrollments from 1999 to 2003 as the research object, and used logistic regression models to analyze the impact of regional economy, population, ideas and other variables on Hardin-Simons University enrollments (Davidson, 2005). The prediction methods used by the above researchers are some of the most basic predictions, and the accuracy of the prediction results is poor, and there is no professional prediction model.

Thereafter, the research on the enrollment proportion of general education and vocational education in senior high school has gradually begun to introduce prediction models, mainly focusing on model predictions such as the population of appropriate age, the scale of education, the number of enrollments, the scale of teachers and students, the output of technical and skilled talents, and industrial demand. However, there is a lack of comparison between prediction models. Zhou et al. used the GM(1,1) model to predict and analyze the trend of the enrollment scale of higher vocational education in Hubei Province during the "Twelfth Five-Year Plan" period, and put forward suggestions for realizing the goal of the enrollment scale of higher vocational education (Zhou & Wang, 2012). Fieger used student learning intention, learning satisfaction, employment return and education population as research variables, and used logistic regression model for predictive analysis to analyze the completion rate of students in the vocational education stage. This study analyzes the employment status of vocational education students and its influencing factors (Fieger, 2015). The introduction of the prediction model enriches the research perspective of the enrollment proportion of general education and vocational education, and the scientific and dynamic quantitative model is gradually applied and developed in the scale research of secondary vocational education. Zhang *et al.* analyzed the scale and development of local vocational education and from 2004 to 2013， found out the regular pattern to establish the log linear model， by which the scale of vocational education in Sichuan from 2014 to 2023 can be predicted (Zhang & Zhu, 2015). James, M used a multiple regression model to predict the quality of secondary vocational and technical education according to a variety of influencing factors (James, 2016). Based on the data of school-age population and freshmen population, Fang *et al.* analyzed the development trend of vocational education in China from 2016 to 2030 by establishing a mathematical prediction model of the development scale of vocational education (Fang & Zhang, 2017). At this stage, whether it is a prediction study on the quality of education or a prediction study on the scale of education development, the factors considered in the prediction study are more and more comprehensive and diverse. Based on the background that the government has increased the ratio of vocational high school graduates to regular high school graduates, Wicaksono et al. used the multinomial-Logit model to analyze the types of high schools (Wicaksono&Sparrow, 2018), and used the Logit model and the Probit model to analyze whether students choose vocational education or general education. The innovation of Wicaksono’s forecasting research is the comparison of two forecasting models. Cirelli et al. developed a data analytics model that can be used by universities and colleges to improve student admission and enrollment process (Cirelli, 2018). This data analysis model is mainly applicable to individual school enrollment. Although it is not generalized to the prediction of the entire enrollment field, it has certain reference. McDaniel used the Random Forest (RF) algorithm to predict high school GPA, college courses taken in high school, ACT scores, high school enrollment proportion of low-income students, and high school enrollment proportion of URM students (McDaniel, 2016). In order to make students better adapt to the changes in the post-industrial social and economic environment, Xie et al. conducted research on the quality training of professional talents, and the research results can be applied to the prediction of students’ career prospects by educational institutions and employers, and this kind of prediction research is more dynamic and precise (Xie & Zhu, 2019). The above-mentioned prediction researches on the enrollment proportion of general education and vocational education are based on scientific prediction models, which are innovative to a certain extent. At the same time, their limitations are that only a single model is used for prediction research, and it is impossible to judge the pros and cons of each model’s prediction effect. Lu et al. took China’s 2012-2016 secondary vocational education and general high school enrollment as the research object, and used the method of four intervals to predict and compare the development trend of enrollment (Lu & Polytechnic,2019). This also really begins to introduce the concept of multi-prediction model comparative research. Different prediction models have different applicability to different research objects. Therefore, the prediction of the enrollment proportion of general education and vocational education requires a comparative analysis of multiple prediction models to find the model with the best prediction accuracy and stability.

1.3 Problem Solving and Innovation

To sum up, researches on the enrollment proportion of general education and vocational education mostly focus on the source of students, the size of students in school, the employed population, the teacher-student ratio and other fields. Although some scholars have pointed out the concept of multi-model prediction, no comparative analysis has been conducted among models. On the basis of existing research, this study makes the following innovations. Based on the data of the enrollment proportion of general education and vocational education in 2010-2018 from National Bureau of Statistics of China, grey prediction model, exponential smoothing model, ARIMA model and BP neural network were used to predict the enrollment proportion of general education and vocational education in 2019. The predicted value was compared with the real value of the enrollment proportion of general education and vocational education in 2019, and then the prediction accuracy of each model was obtained. By comparing the prediction accuracy and stability of various prediction models, the optimal model suitable for the enrollment proportion of general education and vocational education prediction is determined. Based on the optimal model, the development of the enrollment proportion of general education and vocational education in the representative cities of the eastern, central and western regions of China in the next five years is predicted, which further illustrates the practicability and effectiveness of this model. This provides an important reference for the selection of the province’s enrollment proportion of general education and vocational education prediction model, and has good practicability and innovation. Predicting the development trend of the enrollment proportion of general education and vocational education in various regions can better grasp the development trend of the structure of general and vocational education in each region, which has important strategic significance for the rationality of the talent structure in each region and its coordination with regional economic development.

2. Method

2.1 Source of Data

The enrollment proportion of general education and vocational education is the ratio of the number of students enrolled in regular high schools to the number of students enrolled in secondary vocational education. In order to predict the development trend of the enrollment proportion of general education and vocational education in various regions of the country, the enrollment data of regular high school and secondary vocational education from 2010 to 2018 published on the website of the National Bureau of Statistics was selected as a sample. Due to the incomplete data on the number of students enrolled in regular high school and secondary vocational education nationwide in 2020, the 2019 data on the enrollment proportion of general education and vocational education in various regions of the country was used as the predicted value, and the gray prediction model, exponential smoothing model, ARIMA model and BP neural network were used to analyze the data respectively. The enrollment proportion of general education and vocational education in 2019 is forecasted and compared with the actual value of the enrollment proportion of general education and vocational education in 2019 released by the Bureau of Statistics. From 2010 to 2019, the number of students enrolled in regular high schools in various regions of the country is shown in Figure 1, and the number of students enrolled in secondary vocational education in various regions of the country from 2010 to 2019 is shown in Figure 2. Figure 3 shows the enrollment proportion of general education and vocational education in various regions of the country from 2010 to 2018.

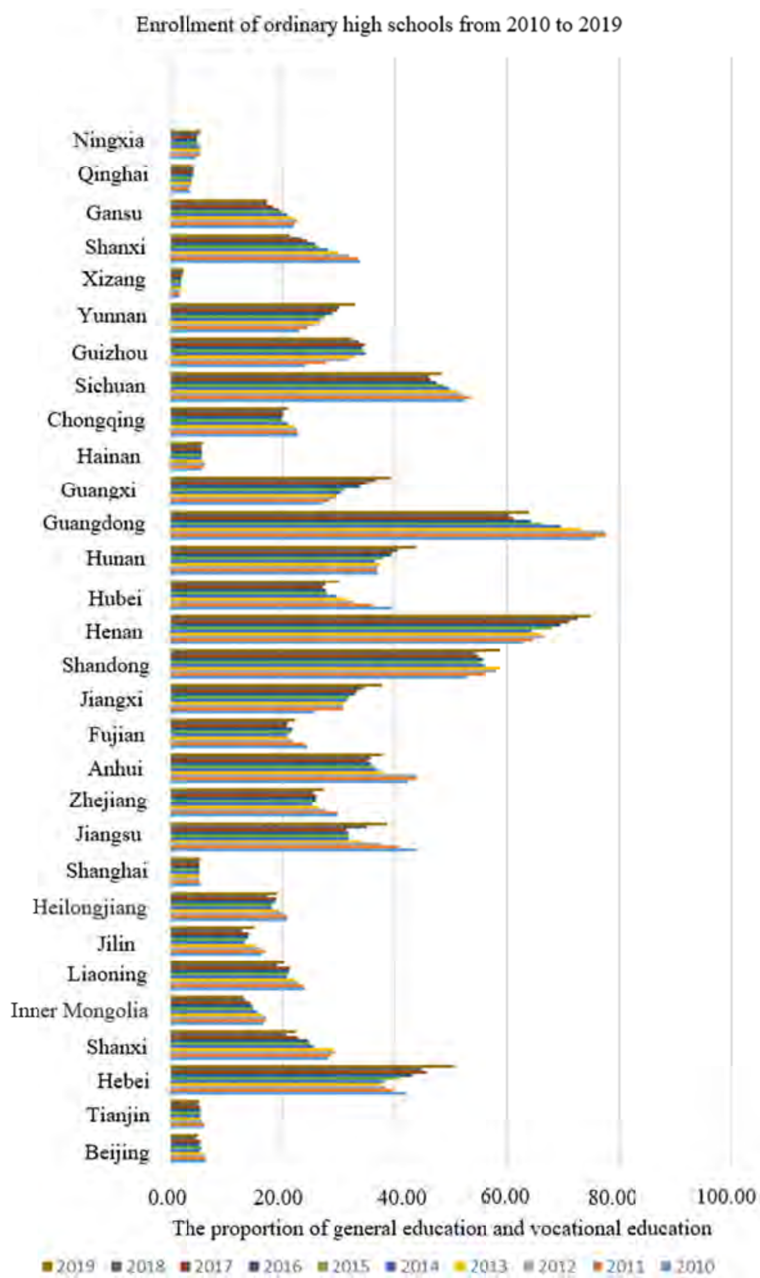

Figure 1. Enrollment of regular high schools in various regions of the country from 2010 to 2019 (*Note*. The data comes from the National Bureau of Statistics)

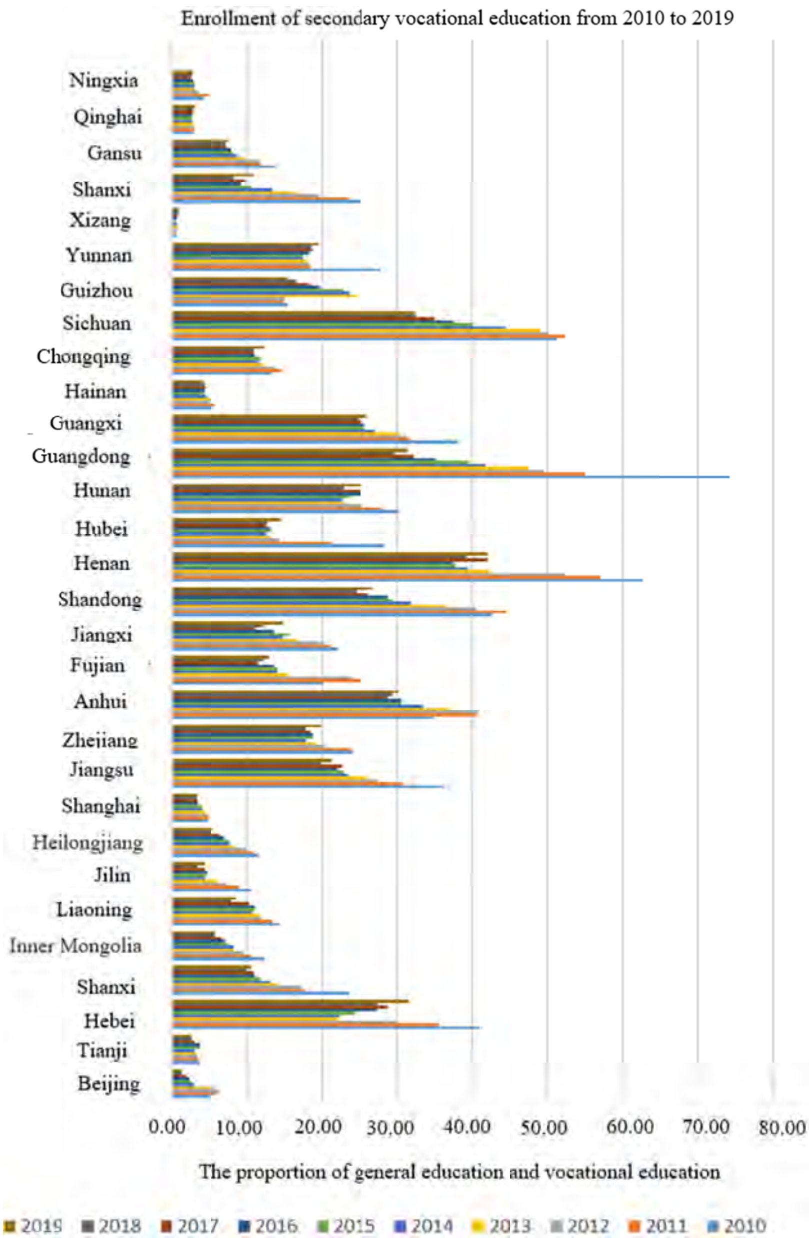

Figure 2. Enrollment of secondary vocational education in various regions of the country from 2010 to 2019 (*Note*. The data comes from the National Bureau of Statistics)

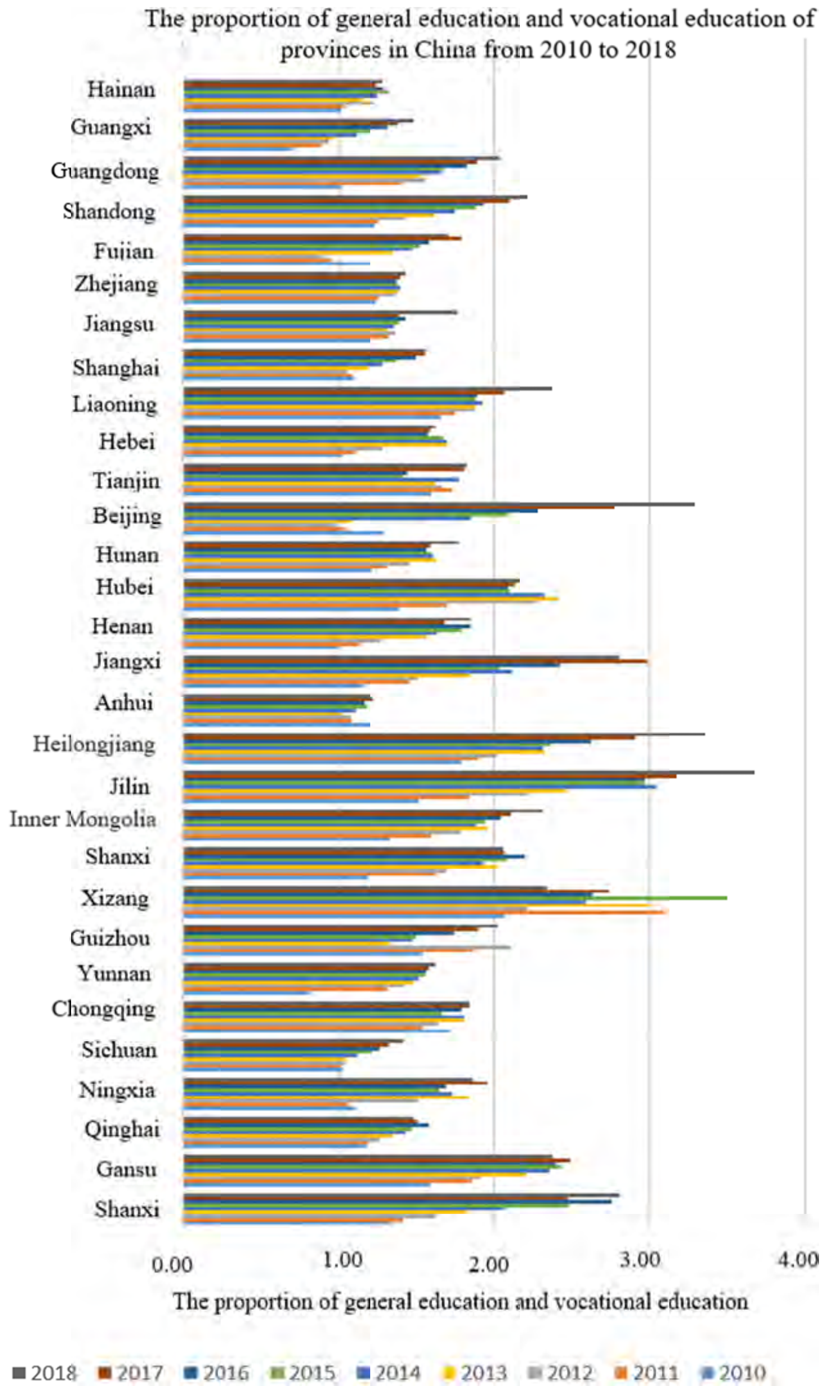

Figure 3. Enrollment of secondary vocational education in various regions of the country from 2010 to 2019

2.2 Model Construction

At present, the main data prediction methods include neural network method, grey prediction method and exponential smoothing, and they have been widely used in various fields. However, for different prediction objects, due to the diversity of influencing factors, there are differences in the prediction effects of each prediction model. Therefore, neural network method, grey prediction method, exponential smoothing and ARIMA model are selected to predict the enrollment proportion of general education and vocational education, and the optimal prediction model is selected by comparing the prediction effect.

2.2.1 Grey Prediction Model

The grey system theory is a novel and efficient method to deal with uncertain problems, which was pioneered by Deng in 1982 (Ju-Long, 1982). It is mainly based on less data and poor information to solve many uncertain problems. "Black" means unknown information, "white" means known information, and "grey" means some information is known and some information is unknown. The theory holds that data is complex, but it is ordered and functional. The generation of grey numbers is to find rules from clutter. Grey system theory is the process of modeling, analysis, control and decision-making. Grey prediction has been widely used in many fields, such as risk management (Li, 2016), energy consumption (Zeng & Zhou, 2017) and road transportation (Balochian & Baloochian, 2021). At present, it has been widely used in practical problems such as enrollment forecast, teacher-student ratio forecast, and gross enrollment rate forecast in the field of education. The most widely used grey forecasting model is the GM (1,1) model consisting of a single-variable first-order differential equation. The model is mainly used for the fitting and prediction of the eigenvalues of a certain dominant factor in a complex system to reveal the changing law of the dominant factor and the future development trend. By identifying the dissimilarity of development trends between systems, the correlation analysis is carried out, and the original data are accumulated or reduced to form a new data series. The prediction model is established by fitting the differential equation of the generated new data series, and then the enrollment proportion of general education and vocational education is predicted. Compared with the traditional prediction methods, the model can obtain the short-term prediction with better accuracy based on a small amount of data. The enrollment proportion of general education and vocational education is affected by various uncertain factors such as national policy, student source, social population and economic development, so it is uncertain and complex and suitable for the prediction of grey prediction model. The specific steps of model construction are as follows:

In the construction of the original sequence, $X^{(0)}=\{X^{(0)}(i),\ i=1,2,\ldots,n\}$, the accumulated AGO (Acumulated Generating Operation) on the sequence $X^{(0)}$ can obtain a set of data sequences with an exponential growth law, which become the cumulative generated sequence $X^{(1)}$, that is

$$\mathbf{X}^{(1)}(i)=\sum_{k=1}^{i}\mathbf{X}^{(0)}(k) \tag{1}$$

Based on the grey system theory, the whitening differential equation about time t is established for the above univariate sequence, and the G M (1,1) model is adopted, which can be expressed by the first-order differential equation:

$$\frac{\mathrm{d}x^{(1)}}{\mathrm{dt}}+a\cdot x^{(1)}=b \tag{2}$$

where a and b are undetermined parameters, denoted as $A=(a,b)^{\mathrm{T}}$. The least squares approximate solution $\hat{\boldsymbol{A}}$ of the equal-step time series A can be derived by discretization and difference.

$$\hat{\boldsymbol{A}}=[a,b]^{\top}=[\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}]^{-1}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{Y}_n \tag{3}$$

Where $B=\begin{bmatrix}-\frac{1}{2}\{x^{(1)}(1)+x^{(1)}(2)\} & 1\\ -\frac{1}{2}\{x^{(1)}(2)+x^{(1)}(3)\} & 1\\ \vdots & \vdots\\ -\frac{1}{2}\{x^{(1)}(n-1)+x^{(1)}(n)\} & 1\end{bmatrix}$, $\hat{X}^{(0)}(k)=\hat{X}^{(1)}(k)-\hat{X}^{(1)}(k-1)$

Finally, determine the function model:

$$\begin{gathered}\hat{X}^{(1)}(k)=\left(X^{(0)}(1)-\frac{u}{a}\right)e^{-a(k-1)}+\frac{u}{a}\\ \hat{X}^{(0)}(k)=\hat{X}^{(1)}(k)-\hat{X}^{(1)}(k-1)\end{gathered} \tag{4}$$

2.2.2 Exponential Smoothing

Smoothing (Exponential Smoothing, ES) is an Exponential method used by Robert G. According to Brown, the situation of time series is stable or regular, so time series can be reasonably extended along the trend. Exponential smoothing is a method of making predictions by taking a weighted average of past observations. The exponential smoothing method is based on the data in the time series and is smoothed to realize the predicted value of the time series. Its advantage is to eliminate the fluctuation caused by random influence, and it has been used in many fields, such as the prediction of inflation rate (Jere & Siyanga, 2016) and energy consumption (de Oliveira& Oliveira, 2018). According to the different times of smoothing, it can be divided into single exponential smoothing, quadratic exponential smoothing and cubic exponential smoothing. Single exponential smoothing is also known as single exponential smoothing. Its calculation formula is as follows:

$$S_{t+1} = \alpha Y_t + (1-\alpha)S_t \tag{5}$$

where Y_t represents the actual observed value in period t; S_t represents the predicted value in period t, S_{t+1} represents the predicted value in period t+1; α represents the smoothing coefficient, also known as the weighting factor, $0 \le \alpha \le 1$. The predicted value in period t+1 is the weighted average of the observed value in period t and the predicted value in period t. Since there is no predicted value S_1 for period 1 at the beginning of the calculation, generally set S_1 as the actual observed value of period 1, that is, $S_1 = Y_1$.

Therefore, the predicted value for period 2 is:

$$S_2 = \alpha Y_1 + (1-\alpha)S_1 = \alpha Y_1 + (1-\alpha)Y_1 = Y_1 \tag{6}$$

The forecast for period 3 is:

$$S_3 = \alpha Y_2 + (1-\alpha)S_2 = \alpha Y_2 + (1-\alpha)Y_1 \tag{7}$$

$$S_t^{(2)} = \alpha S_t^{(1)} + (1-\alpha)S_{t-1}^{(2)} \tag{8}$$

where, $S_t^{(2)}, S_{t-1}^{(2)}$ are the exponential smoothing value of period t and period t-1 respectively, and α is the smoothing coefficient. The prediction model of quadratic exponential smoothing is as follows:

$$\begin{aligned} \hat{Y}_{t+T} &= a_t + b_t \cdot T \\ S_t^{(2)} &= \alpha S_t^{(1)} + (1-\alpha)S_{t-1}^{(2)} \\ a_t &= 2S_t^{(1)} - S_t^{(2)} \\ b_t &= \frac{\alpha}{1-\alpha}\left(S_t^{(1)} - S_t^{(2)}\right) \end{aligned} \tag{9}$$

The calculation formula of cubic exponential smoothing is as follows:

$$S_t^{(3)} = \alpha S_t^{(2)} + (1-\alpha)S_{t-1}^{(3)} \tag{10}$$

The prediction model of cubic exponential smoothing is as follows:

$$\begin{aligned} \hat{Y}_{t+T} &= a_t + b_t T + c_t T^2 \\ a_t &= 3S_t^{(1)} - 3S_t^{(2)} + S_t^{(3)} \\ b_t &= \frac{\alpha}{2(1-\alpha)^2}\left[(6-5\alpha)S_t^{(1)} - 2(5-4\alpha)S_t^{(2)} + (4-3\alpha)S_t^{(3)}\right] \\ c_t &= \frac{\alpha^2}{2(1-\alpha)^2}\left[S_t^{(1)} - 2S_t^{(2)} + S_t^{(3)}\right] \end{aligned} \tag{11}$$

2.2.3 The ARIMA Model

The change of the enrollment proportion of general education and vocational education is often affected by many factors, including demographic factors, industrial structure, employment demand, public opinion, policy orientation, and many other factors are complicated. Therefore, it is difficult to predict the enrollment proportion of general education and vocational education with traditional prediction methods, so ARIMA model can be used for modeling and analysis. The ARIMA (Autoregressive Integrated Moving Average) model is a famous time-series forecasting method proposed by Box and Jenkins in the early 1970s (Box & Jenkins, 1970). This forecasting method is also called box-Jenkins model or Burkes - Jenkins method. This model converts non-stationary time series into stationary series by difference processing and belongs to dynamic digital model. ARIMA (p, d, q) is called differential autoregressive moving average model. In time series analysis, three parameters need to be set, namely the order of the autoregressive model (p), the order of the smooth difference of time series (d) and the order of the moving average (q). The modeling process is as follows:

Convert the series y_t to a stationary series through *d-order* differences, namely:

$$u_t = \Delta^d y_t = (1-B)^d y_t \tag{12}$$

where u_t is a stationary series, $u_t \sim I(0)$, the ARMA model after d-order difference is the ARIMA model, and its formula is as follows:

$$u_t = c + \phi_1 u_{t-1} + \cdots + \phi_p u_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} \ (\phi_p \neq 0, \theta_q \neq 0) \tag{13}$$

2.2.4 BP Neural Network

BP network (Back Propagation), proposed by scientists headed by Rumelhart and McCelland in 1986, is one of the most widely used neural network models, and has achieved extensive research results in classification, prediction, evaluation and so on. BP network (Back-Propagation Artificial Neural Network) is a multi-layer feed-forward neural network, which adopts a back-propagation learning algorithm and uses input and output sample data for training, so that the network can achieve a nonlinear mapping function relationship between input and output. BP network has been used in many fields, such as biomedical science (Liu, 2016) and engineering (Bangalore & Tjernberg, 2015) (Ševo&Avramović, 2016).There are many existing neural network methods, among which the most used neural network model is the back error propagation method. The BP neural network model mainly includes input layer, output layer and hidden layer (as shown in Figure 4), where x_i(i=1~n) is the input of the neuron, and y_i (i=1~n) is the output of the hidden layer, z_i (i=1~n) is the output of the output layer, and w represents the connection weight. Neurons between adjacent layers are connected by weight coefficients, and neurons in the same layer are parallel and unconnected.

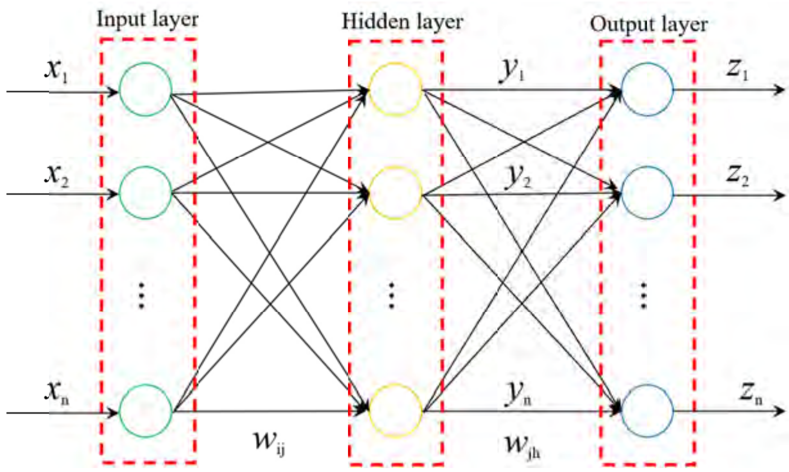

Figure 4. Structure of BP neural network

The specific steps of using BP neural network to predict the enrollment proportion of general education and vocational education are as follows: Firstly, take the data of the enrollment proportion of general education and vocational education from 2010 to 2018 as the input sample, adopt the double hidden layer neural network structure, and normalize the input and target matrix; Then, the training times, training speed and target error are set, the input and output weights and thresholds are initialized, and the outputs of hidden layer and output layer are calculated. The error comparison and sample test are carried out to obtain the trained neural network; finally, the enrollment proportion of general education and vocational education in 2019 is predicted based on the trained network. The specific calculation process is shown in Figure 5.

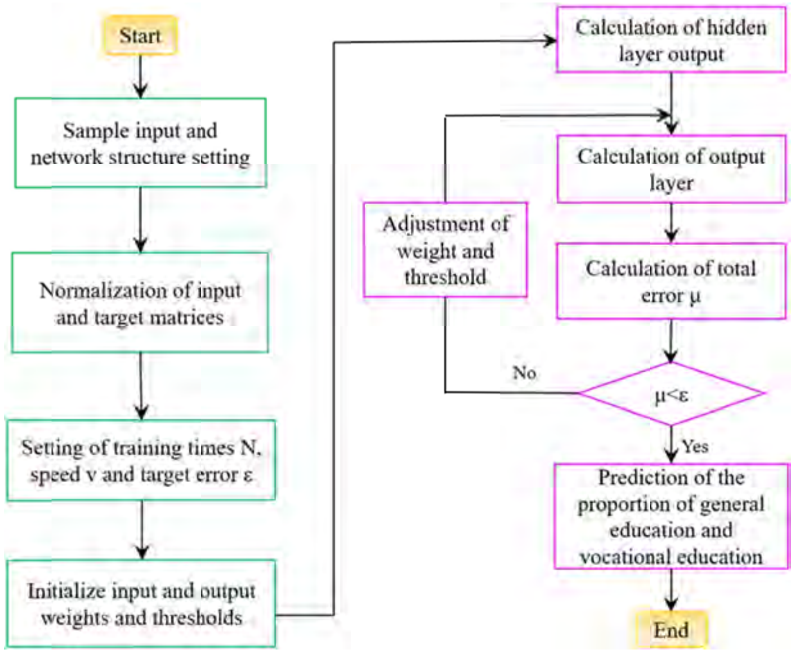

Figure 5. Calculation process of BP neural network

3. Results

Taking the enrollment proportion of general education and vocational education in various regions of the country from 2010 to 2018 as the analysis sample, using the above-established BP neural network model, gray prediction model, exponential smoothing model, and ARIMA model to analyze the enrollment proportion of general education and vocational education in 2019. The prediction results of each model are obtained as shown in Table 1. The comparison between the predictive value of each model and the actual value is shown in Figure 6.

Table 1. The predictive value and the actual value of the enrollment proportion of general education and vocational education in each region in 2019 by each model

Region	Number	Province	Actual value	Grey prediction	Exponential smoothing	ARIMA	BP neural network
Western region	1	Shanxi	1.95	3.215	2.928	2.745	2.702
	2	Gansu	2.26	2.642	2.376	2.271	2.321
	3	Qinghai	1.32	1.626	1.453	1.350	1.429
	4	Ningxia	1.92	2.068	1.874	1.601	1.924
	5	Sichuan	1.49	1.479	1.498	1.477	1.525
	6	Chongqing	1.70	1.899	1.832	1.789	1.847
	7	Yunnan	1.70	1.690	1.683	1.577	1.644
	8	Guizhou	2.10	1.814	1.772	1.723	2.075
	9	Xizang	2.21	2.621	2.607	2.982	2.303
Central region	10	Shanxi	2.10	2.263	2.156	2.008	2.069
	11	Inner Mongolia	2.26	2.358	2.358	2.415	2.580
	12	Jilin	3.41	3.867	3.891	3.879	4.421
	13	Heilongjiang	3.45	3.446	3.735	3.648	3.810
	14	Anhui	1.25	1.259	1.220	1.195	1.215
	15	Jiangxi	2.53	3.214	3.221	2.731	3.122
	16	Henan	1.78	2.044	1.842	1.812	1.797
	17	Hubei	2.08	2.222	2.168	2.176	2.118
	18	Hunan	1.73	1.773	1.770	1.837	2.102
Eastern region	19	Beijing	4.08	3.657	3.725	3.469	3.548
	20	Tianjin	2.09	1.696	1.665	1.659	1.855
	21	Hebei	1.61	1.791	1.628	1.593	1.603

22	Liaoning	2.35	2.263	2.657	2.392	2.585
23	Shanghai	1.58	1.733	1.606	1.533	1.596
24	Jiangsu	1.82	1.622	1.654	1.399	2.656
25	Zhejiang	1.37	1.451	1.438	1.423	1.510
26	Fujian	1.71	1.999	1.723	1.647	1.768
27	Shandong	2.20	2.440	2.339	2.262	2.287
28	Guangdong	2.03	2.088	2.096	1.989	2.141
29	Guangxi	1.52	1.596	1.568	1.528	1.558
30	Henan	1.42	1.358	1.284	1.259	1.280

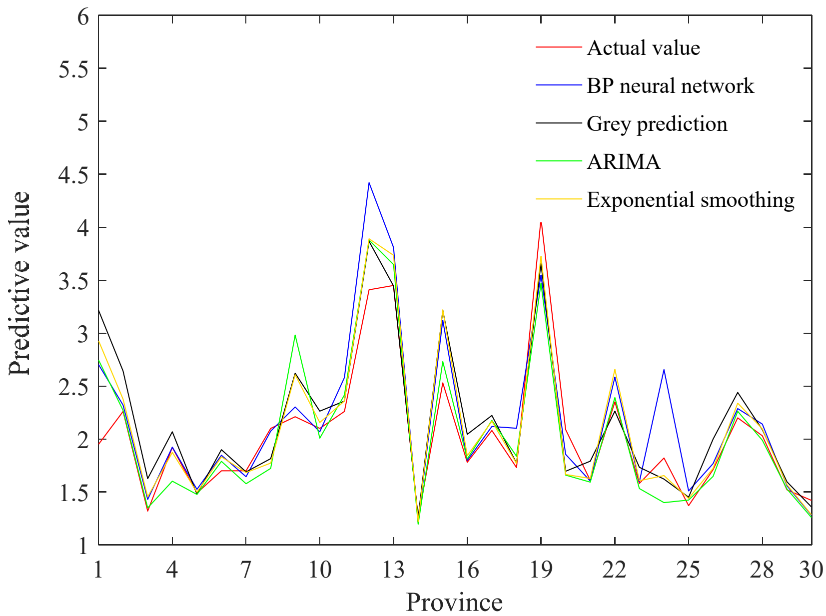

Figure 6. Comparison between the actual value of the enrollment proportion of general education and vocational education in 2019 and the predicted value of each model (*Note*. The numbers on the horizontal axis in the figure represent the numbers of various regions in the country (numbers refer to Table 1))

It can be seen from Figure 6 that the predictive value of the BP neural network for Inner Mongolia, Jilin, Heilongjiang, Jiangsu and Zhejiang provinces is quite different from the actual value. The predictive value of the enrollment proportion of general education and vocational education by exponential smoothing and ARIMA model fluctuates around the actual value.

4. Discussion

4.1 Comparison and Discussion of Accuracy and Stability of Prediction Model

Calculate the prediction results of the enrollment proportion of general education and vocational education in each model in Table 1 and the actual value of the enrollment proportion of general education and vocational education in each region in 2019 to obtain the mean x and root mean square S of absolute error of the predictive value and the actual value of each model to measure the error between the predictive value and the actual value, and then obtain the accuracy and stability of each prediction model.

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{14}$$

$$S = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}} \tag{15}$$

where S is the root mean square, which reflects the stability of the calculation error of the prediction model. The smaller the value, the better the stability of the model prediction; x_i is the absolute value of the calculation error; x is the mean of the absolute value of the error, which reflects the prediction accuracy of the model; n represents the number of provinces.

The absolute value of the error of the prediction results of the gray prediction model, exponential smoothing model, ARIMA model and BP neural network is summarized, as shown in Figure 7.

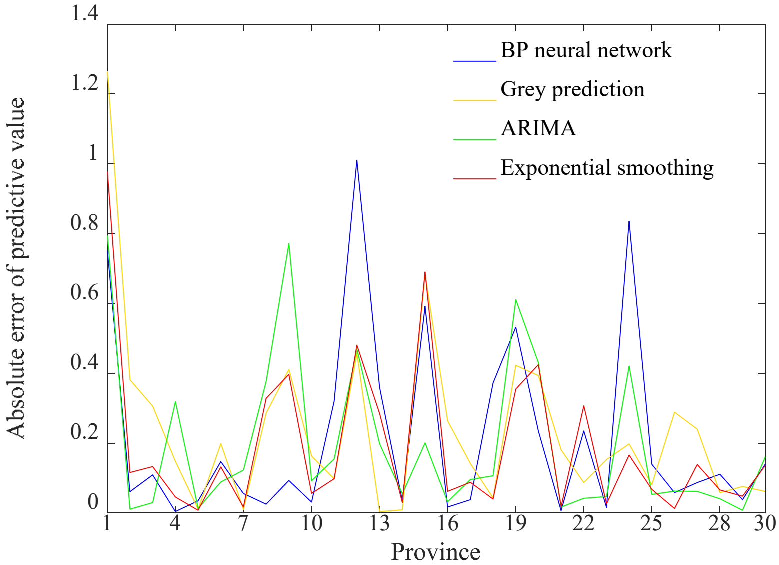

Figure 7. Comparison of absolute values of errors in each model (*Note*. The numbers on the horizontal axis in the figure represent the numbers of various regions in the country (numbers refer to Table 1))

The data shown in Table 2 are obtained by calculating the mean and root mean square of the absolute value of the prediction error of each model.

Table 2. The mean absolute value of the error and the root mean square of the absolute value of the error for each model

Prediction model	The mean of absolute value of error	The root-mean-square of the absolute value of the error
BP neural network	0.216	0.265
Grey prediction	0.237	0.249
ARIMA model	0.196	0.222
Exponential smoothing model	0.192	0.221

It can be obtained from Table 2: (1) The mean absolute value of the error of the exponential smoothing model is 0.192, and the root mean square of the absolute value of the error is 0.221. Both the mean and the root mean square of the absolute error are small, which means that the exponential smoothing model has higher prediction accuracy and better model prediction stability; (2) The calculation accuracy and stability of the ARIMA model are similar to those of the exponential smoothing model, and they are all belong to time series models, which also shows that the time series method has good stability and accuracy in the prediction of the enrollment proportion prediction of general education and vocational education; (3) Although the number of samples required for gray forecasting is small, the model is more suitable for forecasting exponential growth. The enrollment proportion of general education and vocational education in each province is affected by various factors such as economic development, student sources, and national policies. It has great variability, and the grey prediction model is less suitable for the prediction of the enrollment proportion general education and vocational education; (4) The BP neural network has a strong dependence on the training sample data, and the neural network cannot eliminate the contradiction between the prediction ability and the training ability. There is a certain gap between the adaptability of neural network and time series method in the prediction of the enrollment proportion of general education and vocational education.

4.2 Forecast and Trend Analysis

In the prediction analysis of the enrollment proportion of general education and vocational education, the exponential smoothing model is relatively better than the gray prediction model, the ARMIA model and the BP neural network. Therefore, based on the exponential smoothing model, the enrollment proportion of general education and vocational education from 2021 to 2025 was predicted for the randomly selected eastern, central and western regions (Shanghai, Inner Mongolia and Qinghai), and the prediction results are shown in Figure 8.

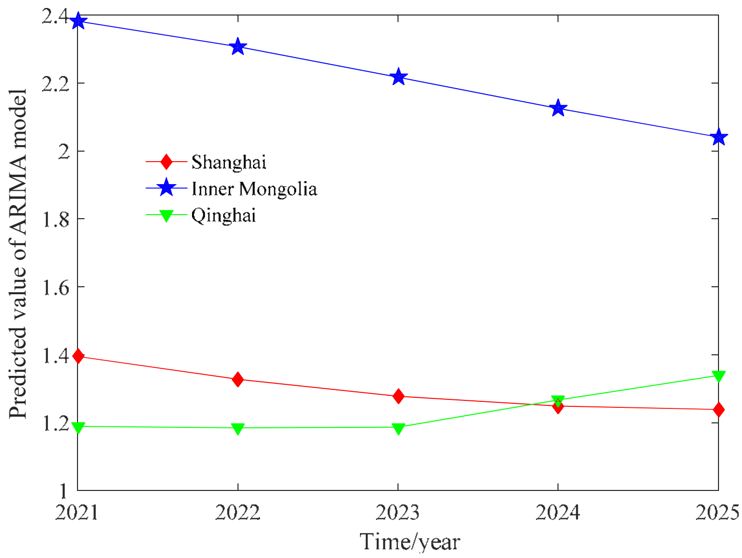

Figure 8. The development trend of the enrollment proportion of general education and vocational education in Shanghai, Inner Mongolia and Qinghai from 2021 to 2025

It can be obtained from Figure 8: (1) the enrollment proportion of general education and vocational education in Inner Mongolia in the next five years will show an upward trend, and will vary between 2.5 and 3, which indicates that the number of regular high school enrollments in Inner Mongolia will continue to increase in the next five years, and the trend is gradually obvious, which is not good for the development of the central underdeveloped areas like Inner Mongolia. Underdeveloped areas need vocational and technical talents to develop regional infrastructure. Therefore, Inner Mongolia should adjust relevant policies to reduce its enrollment proportion of general education and vocational education in the next few years, so as to ensure a reasonable regional talent structure and promote the rapid development of regional economy; (2) The enrollment proportion of general education and vocational education in Shanghai in the next five years will show a relatively stable change trend, fluctuating around 1.7, which means that the number of students enrolled in regular high schools in Shanghai will always be higher than the number of vocational education in the next five years, which is very helpful for the stable development of developed eastern regions like Shanghai. Developed regions need high-level compound talents from regular high schools to develop regional innovation industries and further realize innovation-driven regional economic development. Therefore, according to the development trend of the enrollment proportion of general education and vocational education in Shanghai in the next few years, the current relevant policies of Shanghai are beneficial to the rational development of regional talent structure and the rapid development of regional economy; (3) The enrollment proportion of general education and vocational education in Qinghai will show a relatively slow downward trend in the next five years, changing around 1.3, which shows that its enrollment proportion of general education and vocational education is improving in the direction of optimizing the regional talent structure and steadily improving the economy. It also shows that the current policies on regular high school and vocational education in Qinghai are beneficial to improving the regional talent structure and promoting the development of regional economy.

5. Implications and Conclusion

Through the comparative analysis of the prediction results of BP neural network, grey prediction model, ARIMA model and exponential smoothing model, it can be seen that the exponential smoothing model has better accuracy and stability for the prediction of the enrollment proportion of general education and vocational education. Based on the exponential smoothing model, the prediction of the enrollment proportion of general education and vocational education in Qinghai, Shanghai and Inner Mongolia in the next five years shows that that the development trend of the enrollment proportion of general education and vocational education in Shanghai and Qinghai in the next five years will promote regional economic development. However, the development trend of the enrollment proportion of general education and vocational education in Inner Mongolia in the next few years is not conducive to regional economic development. The province should adopt relevant policies to adjust the enrollment proportion of its general education and vocational education, and make its structure of general education and vocational education develop in a direction that is more conducive to regional economic development. Selecting the optimal and most scientific prediction model for the enrollment proportion of general education and vocational education through comparative analysis can not only provide a solid theoretical basis for the selection of the enrollment proportion prediction method of general education and vocational education in each province, but also improve the regional talent structure through more accurate prediction feedback of the enrollment proportion of general education and vocational education， so as to promote the rapid development of regional economy, which has important practical value and development significance.

6. Advice for Teacher Training

The setting of the enrollment proportion of general education and vocational education should be gradually flexible according to different regions. “The enrollment proportion of general education and vocational education is roughly the same” is a macro policy in a global context, and is a dynamic and flexible concept. As a macro-guidance strategy for the adjustment of the national education structure, the policy of “the enrollment proportion of general education and vocational education is roughly the same” aims to maintain roughly the same nationwide. Changes in the proportion of general education enrollment and vocational education enrollment are affected by a series of factors such as the economic foundation of each region, industrial transformation and upgrading, education popularization, and the status of the school-age population. It is mandatory for the state to maintain a ratio of 1:1 between the number of students enrolled in general education and the number of students enrolled in vocational education in each region, which is neither in line with the actual development of education, nor conducive to promoting regional economic construction and social progress. Therefore, the enrollment proportion of general education and vocational education should not be set uniformly and should not be “one size fits all”. The state should flexibly formulate and adjust the enrollment proportion of general education and vocational education according to local economic development needs and local conditions, allowing the enrollment proportion of general education and vocational education in different regions to fluctuate up and down.

The education department should not mandate the enrollment proportion of general education and vocational education, and should manage this ratio flexibly according to the actual situation of each region, so as to continuously improve the flexibility of the policy adjustment. The education department should flexibly stipulate the enrollment plan according to the actual situation of each region, decentralize the enrollment management power, and allow the provinces and cities to complete the enrollment task according to the actual situation of the region. The education sector should change the way of forced diversion and transition to a natural way of diversion. The diversion of general education and vocational education in China is mainly measured by marks, which results in students having no choice in the type of education, and a large number of students face forced diversion. The implementation of the separation of general education and vocational education should be supplemented by the legal system and scientifically guided. Education departments should respect students’ own needs and embody the management philosophy of democratization, humanity and flexibility. Educational administrators should change the status quo that educational diversion is determined by scores, and construct a selection system based on students’ wishes and endowments, so that candidates can truly choose education that suits them according to their interests, preferences, thinking advantages, and abilities.

The adjustment of the enrollment proportion of general education and vocational education comprehensively considers factors such as the economic development level of a certain region and the status of the appropriate-age population. According to the exponential smoothing model, the enrollment proportion of general education and vocational education in a certain area is predicted, and the result is used as the reference data for the adjustment of the proportion. When the exponential smoothing model predicts that the number of enrollments in general education is significantly higher than that in vocational education, and regions rely on the real economy for support and require a large number of professional and technical personnel, the enrollment scale of ordinary high schools should be appropriately reduced. Education leaders in the region should improve the quality of vocational education and cultivate the “internal skills” of secondary vocational education. All localities should strengthen the construction of demonstration secondary vocational colleges, promote the in-depth integration of industrial development and education, standardize professional settings, and innovate talent training models. Vocational teacher training units should strengthen the construction of teaching staff and focus on cultivating teachers with solid theoretical knowledge and professional skills. Enterprises in various regions should fully participate in the training process of professional talents, and make students into skilled talents that meet the development needs of enterprises. At the social market level, the employment access system should be improved, the employment channels for technical and skilled talents should be broadened, and the social status and economic income of skilled talents should be improved. Only in this way can vocational education attract more students. Industrial transformation and upgrading, and profound changes in the economic field will inevitably lead to a high shift in the demand for talents. Secondary vocational and technical talents can no longer fully adapt to the changes in the new economy and new fields. For areas with a relatively high level of economic foundation and the transformation of industrial structure, the education department should appropriately increase the proportion of the number of students enrolled in general education to vocational education, and reduce the scale of secondary vocational education. Therefore, if relevant researchers predict that the number of enrollment in general education is lower than that in vocational education through the exponential smoothing model, the education department in the region should readjust the critical score line for the separation of general education and vocational education, and the scale of enrollment in secondary vocational education should be readjusted, appropriate reduction. The national level should implement the quality supervision mechanism of general high school education, and revise and improve the quality standards of general high school education. In order to adapt to the adjustment of the industrial structure of the regional economic development industry, ordinary high schools should cultivate high-quality and compound talents. The Municipal Education Bureau should increase financial investment in ordinary high schools and raise the financial subsidy standard for public funds per student.

References

Balochian, S., & Baloochian, H. (2021). Improving grey prediction model and its application in predicting the number of users of a public road transportation system. *Journal of Intelligent Systems, 30*(1), 104-114. https://doi.org/10.1515/jisys-2019-0082

Bangalore, P., & Tjernberg, L. B. (2017). An artificial neural network approach for early fault detection of gearbox bearings. *IEEE Transactions on Smart Grid, 6*(2), 980-987. https://doi.org/10.1109/TSG.2014.2386305

Bo, Z., Meng, Z., & Jun, Z. (2017). Forecasting the energy consumption of china's manufacturing using a homologous grey prediction model. *Sustainability, 9*(11), 1975. https://doi.org/10.3390/su9111975

Box, G. E. P., & Jenkins, G. M. (1970). Time series analysis forecasting and control. *Journal of Time Series Analysis, 3*(3228).

Brown, J. M., & Chang, G. (1982). The Predictive Validity of the Minnesota Reading Assessment for Students Inpostsecondary Vocational Education Programs. *Educational and Psychological Measurement, 42*(1), 345-352. https://doi.org/10.1177/0013164482421042

Choi, S. (2021). The impact of education levels and paths on labor market outcomes in South Korea: Focusing on vocational high school graduates. *Social Sciences & Humanities Open, 4*(1), 100152. https://doi.org/10.1016/j.ssaho.2021.100152

Cirelli, J., Konkol, A. M., Aqlan, F., & Nwokeji, J. C. (2018, January). Predictive analytics models for student admission and enrollment. In *Proceedings of the International Conference on Industrial Engineering and Operations Management* (Vol. 2018, No. SEP, pp. 1395-1403). Retrieved from https://www.researchgate.net/publication/329998695_Predictive_Analytics_Models_for_Student_Admission_and_Enrollment

Davidson, J. S. (2005). *Enrollment Prediction at a Texas Baptist University*. Baylor University, Waco, TX 76706，USA

De Oliveira, E. M., & Oliveira, F. L. C. (2018). Forecasting mid-long term electric energy consumption through bagging ARIMA and exponential smoothing methods. *Energy, 144*, 776-788. https://doi.org/10.1016/j.energy.2017.12.049

Fang, F. W., & Zhang, X. C. (2017). *Analysis of vocational education trend in china from 2016 to 2030*. Vocational Education Research.

Fieger, P. (2015). Determinants of course completions in vocational education and training: Evidence from Australia. *Empirical Research in Vocational Education and Training, 7*(1), 1-14. https://doi.org/10.1186/s40461-015-0025-5

Hager Jr. R. D. (1970). *Analysis of a Cohort Prediction Model with Applications to Student Enrollment Forecasting*.

Hoger, P. (1990). Vocational education/general education: a faise dichotomy?. *Studies in Continuing Education, 12*(1), 13-23. https://doi.org/10.1080/0158037900120102

Hussar, W. J., & Bailey, T. M. (2007). *Projections of Education Statistics to 2016*. NCES 2008-060. National Center for Education Statistics.

Jere, S., & Siyanga, M. (2016). Forecasting inflation rate of Zambia using Holt's exponential smoothing. *Open journal of Statistics, 6*(2), 363-372. https://doi.org/10.4236/ojs.2016.62031

Ju-Long, D. (1982). Control problems of grey systems. *Systems & control letters, 1*(5), 288-294. https://doi.org/10.1016/S0167-6911(82)80025-X

Li, C., Qin, J., Li, J., & Hou, Q. (2016). The accident early warning system for iron and steel enterprises based on combination weighting and Grey Prediction Model GM (1, 1). *Safety science, 89*, 19-27. https://doi.org/10.1016/j.ssci.2016.05.015

Liu, X., Liang, W., Wang, Y., Li, S., & Pei, M. (2016, September). 3D head pose estimation with convolutional neural network trained on synthetic images. In *2016 IEEE international conference on image processing* (ICIP) (pp. 1289-1293). IEEE. https://doi.org/10.1109/ICIP.2016.7532566

Liu, Y. (2005). *The vocational education quantity prediction using neural network*. System Simulation Technology.

Lu, L. I., & Polytechnic, E. (2019). The ratio's four interval analysis framework of secondary vocational education scale and general high school education scale and its application. *Journal of Xiangyang Polytechnic*.

Maxim, D. (1973). *Job Demand Forecasting Program. A Vocational Education Planning System for Local School Districts* (Volume VII).

McDaniel, T. (2016). *Utah's pathways to higher education: a critical, quantitative analysis* (Doctoral dissertation, The University of Utah). Retrieved from https://collections.lib.utah.edu/ark:/87278/s6cp0d6g

Peng, Y. (2007). Optimization of vocational and technical education structure based on industry cluster talent demand. *Research in Higher Education of Engineering, 10*(1), 2006.

Ševo, I., & Avramović, A. (2016). Convolutional neural network based automatic object detection on aerial images. *IEEE geoscience and remote sensing letters, 13*(5), 740-744. https://doi.org/10.1109/LGRS.2016.2542358

Tilak, J. B. (2003). Vocational education and training in Asia. In *International handbook of educational research in the Asia-Pacific Region* (pp. 673-686). Springer, Dordrecht. https://doi.org/10.1007/978-94-017-3368-7_46

Tuckman, H. P. (1971). *The Use of Predictive Models in Forecasting Student Choice*. Paper presented at the Annual Meeting of the American Educational Research Association, New York, New York, February 1971. Retrieved from https://eric.ed.gov/?id=ED050169

Tuma, J. (1994). Measuring Enrollment and Participation in Secondary Vocational Education with High School Transcript Records. *Journal of Vocational Education Research, 19*(3), 3-22. Retrieved from https://eric.ed.gov/?id=EJ507742

Valentine, M. A. (1985). *Predicting Special Education Students' Success in Vocational Education Programs*. Retrieved from https://eric.ed.gov/?id=ED256090.

Wicaksono, W., Sparrow, R. A., & van Bergeijk, P. (2018). The impact of parents' education and attending vocational high school to college entrance. *Jurnal Pendidikan Vokasi, 8*(1), 12-23. https://doi.org/10.21831/jpv.v8i1.17938

Xie, L. C., Zhu, X. D., & Cao, C. Y . (2019). *Research on the Construction of Education Trajectory and Career Planning Based on Individual Characteristics in Higher Vocational Colleges*. Retrieved from https://xueshu.baidu.com/usercenter/paper/show?paperid=1n1k0e207b680tn0a3470xh0pm240322

Zhang, J., & Zhu, Q. J. (2015). Scale prediction of higher vocational education from 2014-2023 in sichuan. *Journal of Chengdu Normal University*.

Zhou, Z. X., Wang, Z. F., & Zhou, F. (2012). Study on prediction of the enrollment development trends of hubei higher vocational education. *Journal of Xiangfan Vocational and Technical College*.