Measurement of the coupling between education and regional economic development levels in Western China

Haiyan Wang[a], Yi He[b], Sirong Shang[c]*

[a] 271141489@qq.com,[b] ehyuuki@163.com,[c] sirongshang610@outlook.com.

School of Statistics and Mathematics, Yunnan University of Finance and Economics, Kunming, China

Abstract. The article constructs a coupling and coordination model based on the entropy value method to study the correlation between the level of education and the level of regional economic development in Western China and to make a comprehensive evaluation of the coupling and coordination degree of education and economy in 12 provinces and cities at different times. The results show that the degree of coordination between the education level and the regional economic development level in 12 provinces and cities in western China has maintained an increasing trend year by year, and the degree of coupling is higher in six regions, namely Chongqing, Sichuan, Guangxi, Guizhou, Yunnan and Shaanxi Province. In addition, the article constructs a combined CNNBiLSTM-AM model based on the CNN model and BiLSTM model to achieve the prediction of coupling value in 2024 for western China, and the results show that the model has good robustness and generalization. Finally, we propose feasible recommendations based on the research results.

Keywords: Coupled coordination model; Evaluation index; Educational attainment; Regional economy

1 Introduction

In recent years, although the level of education in China has been well coordinated with the level of regional economic development, there is still a certain disconnect. If the level of education exceeds the level of economic development, there is a situation of "over-education", which means that the level of investment in education exceeds the needs of social and economic development, resulting in inefficient allocation of social resources.

The social development of the western region of China faces the dilemma of an insufficient supply of human resources and an unreasonable structure of supply and demand in the market, in addition to the low contribution of scientific and technological progress to economic growth, which also hurts the sustainable development of the western region. Therefore, promoting the coordinated development of education and the regional economy is of great significance in enhancing the overall cultural literacy of the Western region and achieving poverty alleviation in the regional economy.

ICEKIM 2023, May 26-28, Nanjing, People's Republic of China

DOI 10.4108/eai.26-5-2023.2337300

2 Literature Review

Higher education development should follow its rules and meet the objective requirements of economic and social development. Fan Hua used the Fiducial model to measure the impact of higher education on economic growth in Jiangsu province from 1985 to 2003, and the results show that the degree of the role of higher education varies when the level of economic development varies[1]. To achieve comprehensive socioeconomic development, paying attention to the coordinated development of higher education and the economy is necessary. Lu Zhuo examined the relationship and transmission mechanism between higher education hierarchy and high-quality economic development with the help of a mediating effect model[2]. The study found that the development of higher education in China has a positive effect on both scientific and technological progress and high-quality economic development, and scientific and technological progress plays an important moderating role in it. Chi Jingming et al used regression analysis to explore the interaction between different levels of higher education in China, the social economy, and the mechanism of influence[3]. The continuous development and progress of higher education in China are also in line a lot with the needs of social and economic development. Guo Liqiang et al measured the coupling and coordination between higher education and economic development in 31 provincial administrative units from 2005 to 2015, and pointed out that the coordination and coupling between higher education and regional economy has a significant "Matthew effect"[4]. The synergistic effect of the scale, hierarchical structure, and quality of higher education in promoting economic growth has been increasing with the development of the times.

Many scholars also include science and technology innovation as a part of the mutual influence of higher education and economic development and believe that there is a mutually reinforcing and mutually constraining relationship between higher education, science and technology innovation, and economic development. Yu Hongjiao argued that higher education, economy, and science and technology in Henan Province have all achieved greater development, but the results of regression analysis show that there are certain uncoordinated aspects in the development process of higher education and economy and science and technology[5]. Cai Wenbo et al studied the synergy of the three systems of higher education, science, and technology innovation and economic development in the Yangtze River Economic Belt region and the formation mechanism of their differences using a coupling degree model[6], α convergence test, and a quantile regression model. The results showed that these three systems in the Yangtze River Economic Belt region showed a stable development in general during the period 2007-2018, and the coupling and coordination relationship between the three showed synchronous growth.

Internationally, the annual UNESCO (United Nations Educational, Scientific and Cultural Organization) publication Global Digest of Education Statistics uses 16 statistical tables to present hundreds of statistical indicators for the development of education in individual countries around the world[7]. Bertoletti combines traditional econometric methods with random forests to study the impact of the characteristics of the higher education systems of 29 European countries on the development of the regions in which they are located[8]. The analysis found that the increase in the size of higher education, the increase in the internationalization of students, and the increase in research output had a positive effect on the economic development of the region[9]. Tommaso et al collected data on 284 European regions from 2000-2017 to study the impact of regional higher education systems on economic growth[10]. The results show that increasing the number of higher education institutions in a region can contribute to the economic growth of the region, with the improvement of the quality of teaching and learning in higher education institutions being the main driver of its positive impact on regional economic development.

3 Research Design

3.1 Indicator selection

Based on the experience of scholars such as Zhang Nanxing et al[11], Sun Jihong et al[12], and Xie Chengxing[13] on the selection of indicators for the evaluation of China's higher education development level, and with due consideration to the availability of complex data and the scientific nature of the indicator system, this paper selects six indicators from the four dimensions of higher education teaching scale, faculty strength, education funding, and educational achievement to constitute the higher education evaluation system. Eight indicators representing the level of economic development are selected from the four dimensions of production level, consumption level, trade level, and investment level. The final composition of the indicator system is shown in Table 1 below.

Table 1. Comprehensive Evaluation Indicator System for Education and Economy

System layer	Tier 1 indicators	Tier 2 indicators	Units	Indicator attributes
Economic development	Production	GDP	Billions	+
		General public budget revenue	Billions	+
	Consumption	Total retail sales of consumer goods	Billions	+
		Urban disposable income per capita	RMB/person	+
		Rural disposable income per capita	RMB/person	+
	Trade	Total imports and exports	million (USD)	+
		Total investment in foreign-invested enterprises	million (USD)	+
	Investment	Balance of domestic and foreign currency loans to financial institutions	Billions	+
Higher Education	Faculty	Number of full-time teachers in general higher education	Person	+
		The student-teacher ratio in general higher education	-	+
	Scale of teaching	Number of general higher education schools	House	+
		Number of undergraduate and tertiary students in general education	million people	+
	Funding	General public budget expenditure on education	RMB	+

	Achieve-ments	Effective volume of patents in general higher education	Item	+

3.2 Data processing

The data used for the study were all obtained from the 2017-2021 National Statistical Bulletin on the Implementation of Education Expenditure, provincial and municipal statistical yearbooks, the Public Service Platform for Patent Big Data of Chinese Universities, the statistical database of WIEGO and the People's Bank of China. To improve the accuracy of the calculation results, the raw data of each indicator was standardized using the formula (1).

$$x_{ij}^{'} = \begin{cases} \dfrac{x_{ij} - x_{j\min}}{x_{j\max} - x_{j\min}}, \textit{Positive indicators} \\ \dfrac{x_{j\max} - x_{ij}}{x_{j\max} - x_{j\min}}, \textit{Negative indicators} \end{cases} \tag{1}$$

Based on the entropy method for a comprehensive evaluation, the raw data of 14 indicators collected since 2017-2021 from a total of 12 provinces and cities in the western region were first pre-processed, with a total of 840 samples (including 360 samples of higher education and 480 samples of economic development).

Firstly, the original data matrix is constructed as follows (2).

$$X = \left(x_{ij}\right)_{m \times n}, \left(i=1,2,\cdots,m;\, j=1,2,\cdots,n\right) \tag{2}$$

Next, entropy weights were calculated for the normalized data.

The weight of the indicator value of the jth item under the ith indicator is $P_{ij} = \frac{x_{ij}^{'}}{\sum_{i=1}^{m} x_{ij}^{'}}$. In this way, the entropy value of the ith indicator is calculated as $e_j = -\frac{1}{\ln(m)} \sum_{i=1}^{m} P_{ij} \ln\left(P_{ij}\right)$. Then the variation index d_j of the jth indicator satisfies $d_j = 1 - e_j$. Finally, the entropy weight of the first indicator is calculated as $w_j = \frac{d_j}{\sum_{j=1}^{n} d_j}$.

The entropy weighting formula is used to obtain the weighting percentages for each indicator as shown in Table 2.

Table 2. The weighting of each indicator

System layer	Tier 1 indicators	Tier 2 indicators	Weights
Economic development	Production	GDP	11.256
		General public budget revenue	10.638
	Consumption	Total retail sales of consumer goods	14.107

		Urban disposable income per capita	5.294
		Rural disposable income per capita	4.762
		Total imports and exports	20.272
	Trade	Total investment in foreign-invested enterprises	23.138
	Investment	Balance of domestic and foreign currency loans to financial institutions	10.534
Higher Education	Faculty	Number of full-time teachers in general higher education	13.124
		The student-teacher ratio in general higher education	4.824
	Scale of teaching	Number of general higher education schools	10.745
		Number of undergraduate and tertiary students in general education	13.497
	Funding	General public budget expenditure on education	24.501
	Achievements	Effective volume of patents in general higher education	33.308

Further, we can calculate the combined score for each sample by using equation (3).

$$Z_i = \sum_{j=1}^{n} w_j x_{ij}^{'} \tag{3}$$

The formula was used to obtain the composite evaluation value of higher education and regional economic development of 12 provinces and cities in the western region from 2017 to 2021, with larger values indicating higher representative ratings of the indicators. The article takes 2021 as an example and calculates the composite score of higher education and regional economic development for each province and city in western China as shown in Table 3.

Table 3. Combined scores for provinces and cities in the Western Region

Year	Province	Regional economy	Higher education
2021	Inner Mongolia	0.283641477	0.1819846
2021	Guangxi	0.611750918	0.366210587
2021	Chongqing	0.540183357	0.350743282
2021	Sichuan	0.81159837	0.619165321
2021	Guizhou	0.265790065	0.295930438
2021	Yunnan	0.385963617	0.31715787
2021	Xizang	0.099406888	0.17682815
2021	Shaanxi	0.452235229	0.637969517
2021	Gansu	0.145449507	0.182982886
2021	Qinghai	0.069587295	0.143093693
2021	Ningxia	0.097561286	0.099224845
2021	Xinjiang	0.221753753	0.177808886

The results in Table 3 show that in 2021, all provinces and cities in western China show good growth trends in higher education and regional economic development, but the level of development of higher education in some provinces and cities is not in harmony with the level of economic development.

4 Coupling coordination model

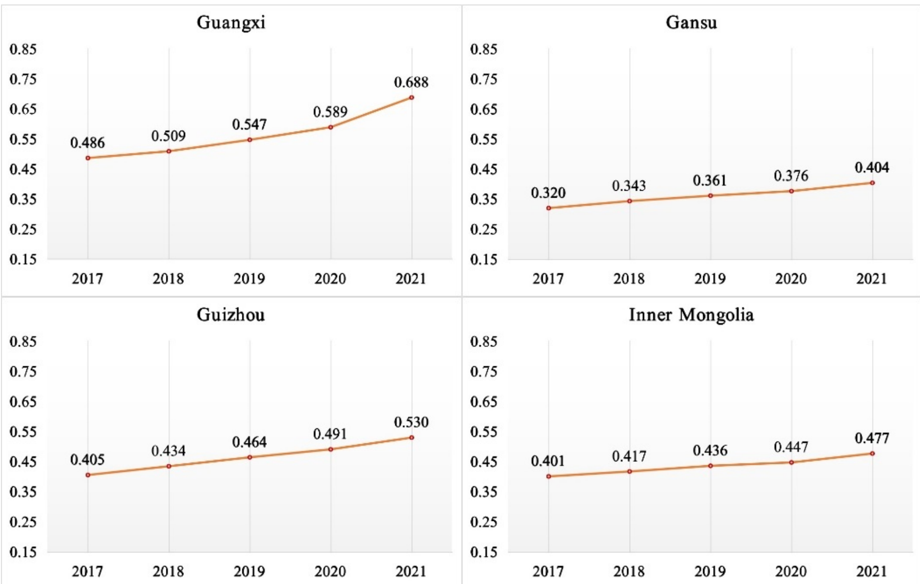
To further analyze the coupling coordination between the two dimensions of the degree of higher education and the level of regional economic development, the article applies the coupling coordination model for index measurement, as shown in the following equations (4)-(6).

$$C = 2\sqrt{f(x)g(y)} / [f(x) + g(y)] \tag{4}$$

$$T = \alpha f(x) + \beta g(y) \tag{5}$$

$$D = \sqrt{C \times T} \tag{6}$$

where equation (4) is the formula for calculating the coupling degree. C is the system coupling degree, and are the comprehensive evaluation indices. Equation (5) is the formula for the comprehensive coordination index. α and β are coefficients to be determined, considering that the level of higher education and economic development are equally important systems so $\alpha = \beta = 0.5$. Equation (6) is then used to calculate the coupling coordination degree. The coupling degree index calculated is finally used to plot the trend of the change of coupling coordination degree of higher education degree and regional development level in 12 provinces and cities in the western region from 2017 to 2021. The details are shown in Figure 1.



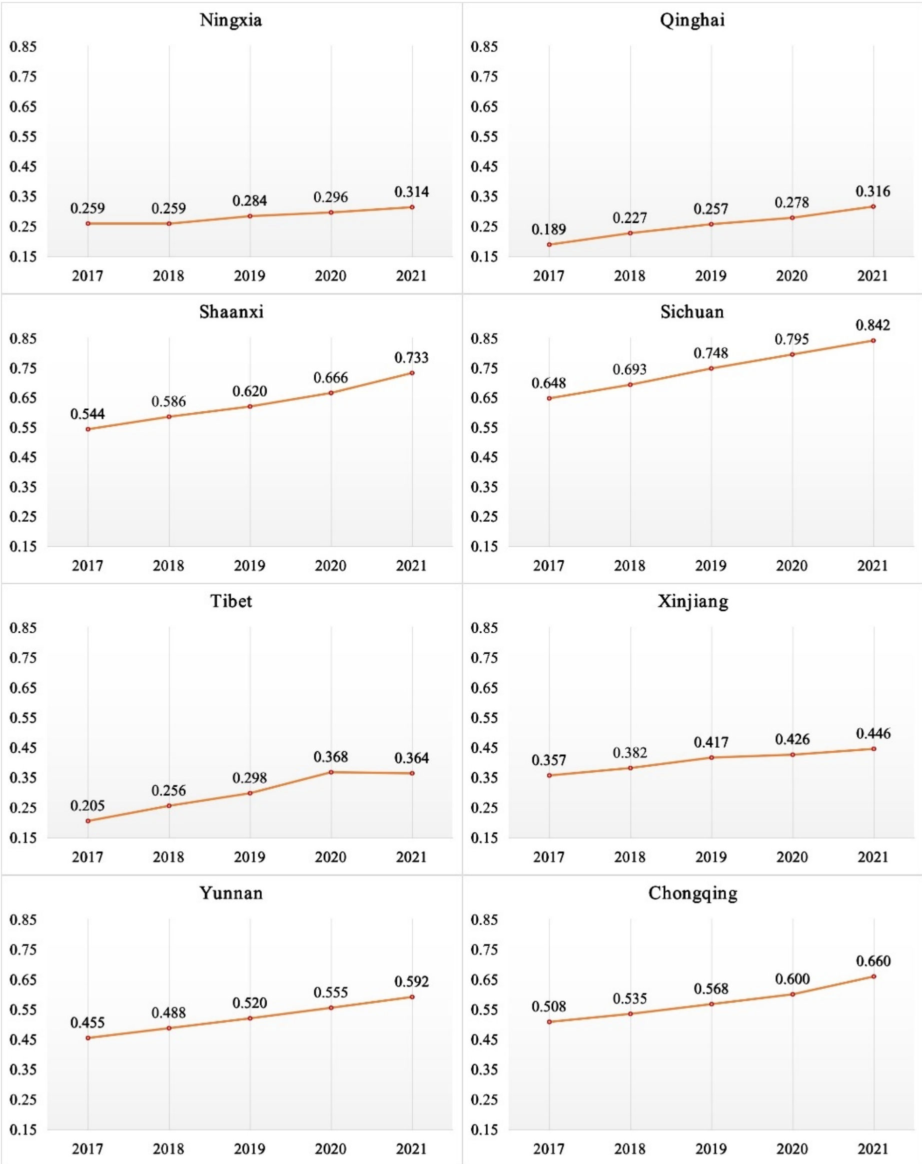


Figure 1. Trends in education and regional economic coupling in the Western Region, 2017-2021

Table 4. Ranking of the coupling index

D-value	Coherence level	D-value	Coherence level
0.00-0.09	Extreme disorder	0.50-0.59	Barely coordination
0.10-0.19	Severe disorder	0.60-0.69	Primary coordination
0.20-0.29	Moderate disorder	0.70-0.79	Intermediate coordination
0.30-0.39	Mild disorder	0.80-0.89	Good coordination
0.40-0.49	Nearly disorder	0.90-1.00	Quality coordination

The changes in the coupling and coordination degree between the degree of higher education and the regional economic development level of 12 provinces and cities in the western region from 2017 to 2021 show that the degree of coordinated development of higher education and the regional economy shows an increasing trend year by year.

It is possible to measure the coupling co-ordination index according to Table 4, as measured by the index in 2021, the coupling and coordination degree of higher education degree and regional economic development in 12 provinces and cities in western China all exceeded 0.3, among which, the coupling degree of six provinces, namely Chongqing, Sichuan, Guangxi, Guizhou, Yunnan and Shaanxi, reached a barely coordinated level. The results of the coupling coherence index measurements for 2021 are detailed in Figure 2.

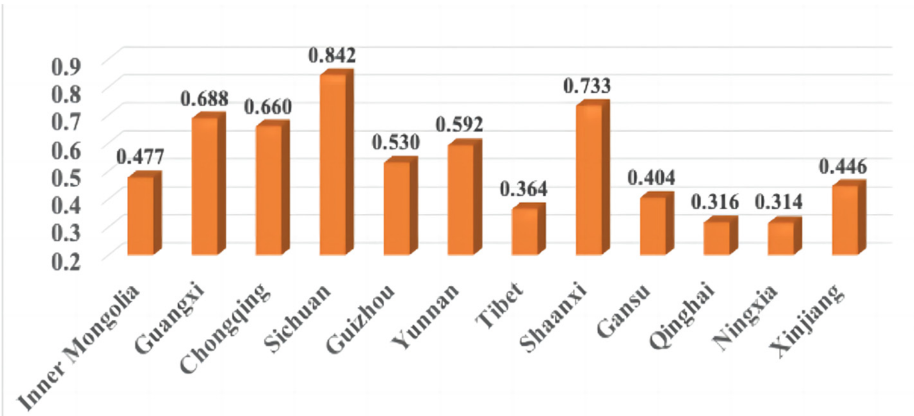


Figure 2. Coupling between education and the regional economy in the Western region in 2021

5 Numerical Prediction Based on Machine Learning MethodsConclusions and Recommendations

As a common prediction model in the field of machine learning and deep learning, deep neural networks can achieve feature extraction of multidimensional complex information by building multidimensional complex network information structures. The article constructs a combined CNNBiLSTM-AM model based on the spatial characteristics of CNN models and LSTM models. Compared with the traditional CNN models and LSTM models, the combined CNNBiLSTM-AM model can better handle the task of time series by non-linear processing of data with multiple implicit layers, while reducing the dependence on long-term preservation of inputs as well as solving the gradient disappearance problem due to back-propagation during training can effectively preserve important information to deal with environmental and network problems with deeper layers.

The computational principle of the CNN model is shown in equation (7) below:

$$Y_m^n = f(n\sum_{i\in O_m} x_i^{n-1}u_{i,m}^n + b_m^n) \tag{7}$$

To further enhance the prediction model's ability to extract and perceive key data information, the output of the BiLSTM model is used as the input layer of the attention model to further enhance the weight value of the predicted data by applying the point multiplication attention calculation principle. The point multiplicative attention calculation principle is used; its formula is shown below.

$$softmax(z_i) = \frac{e^{z_i}}{\sum_j e^{z_i}},\ j = 1, 2, \cdots, K \tag{8}$$

$$Attention = (Q, K, V) = Softmax(Q \cdot K^T)V \tag{9}$$

Where z_i is a k-dimensional vector that enables the elements of the model to take values between (0, 1), Q, K and V refer to Query, Key, and Value respectively, and are mainly used to implement a linear transformation of the input matrix.

After the attention mechanism, the Flatten layer is added to achieve the transition to the fully connected layer, and finally, the prediction results of the data are output through the end neural network of the fully connected layer.

The combined CNNBiLSTM-AM model constructed in this article consists of an input layer, a one-dimensional convolutional CNN layer, an output layer, a BiLSTM layer, an information extraction layer, an attention mechanism, a Flatten layer, and a fully connected layer. However, there are many differences from the traditional neural network model. The BiLSTM network retains a memory storage unit in the implicit layer, through which information from the implicit layer at the previous moment can be transferred to the implicit layer neurons at the current moment, thus explaining the period correlation between different sequences. The framework about the CNNBiLSTM-AM model is shown in Figure 3.

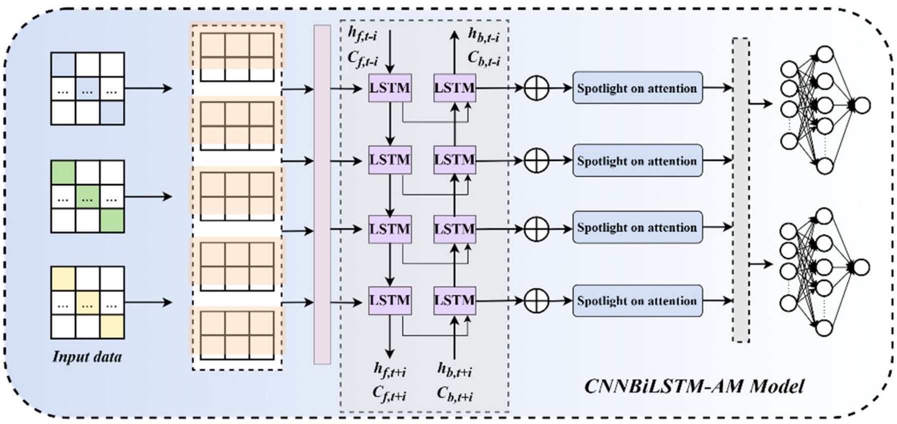


Figure 3. CNNBiLSTM-AM combined model framework

For predicting the coupled values of economic development and education levels in Western China, the model proposed in the article is implemented through the following steps:

(1) Constructing a CNNBiLSTM-AM algorithm model by taking 14 variables in four dimensions - production level, consumption level, trade level, and investment level - as the original variables in the input layer and selecting the coupled values of economy and education as the output variables, and setting the initial parameters of the model;

(2) For the processed time-series data $X=\{x_1,x_2,\cdots,x_t\}$, where $x_i \in R^{m\times n}$, $i\in(1,t)$, m denotes the length of the time window corresponding to each data, and n is the dimensionality of the feature information of the data, we then perform data feature extraction using a one-dimensional CNN model, where we stipulate that the dimensionality of the length of the original data time window remains unchanged after feature extraction of the data information;

(3) For the extracted data features $C=\{c_1,c_2,\cdots,c_t\}$, where $c_i \in R^{m\times h}$, $i\in(1,t)$, h are the filter specification parameters in the convolutional network. The Dropout layer is introduced to ensure that the randomness of the partially hidden layer neurons is shielded, which in turn enhances the training capability of the model.

(4) Adam's gradient descent method was used to evaluate the test performance of the model, and the model was evaluated by observing the "Root Mean Square Error", "Goodness of Fit", " MAE", "MAPE", "MBE" and "MSE" to judge the model's merits and demerits.

The article uses MATLAB software to implement the modeling process of the combined CNNBiLSTM-AM model, and the specific algorithm flow is shown in Table 5.

Table 5. Prediction algorithm based on CNNBiLSTM-AM model

Algorithm: Prediction algorithm based on CNNBiLSTM-AM model

Input: *The 14 variables in the four dimensions of production, consumption, trade, and investment are used as the original variables in the input layer.*
Output: *Coupling values*
Initialize ***Numsize*** to 0.7 and ***Outdim*** to 1;
Select ***the number of samples*** in the ***training set*** and the ***dimensionality*** of ***the input layer features***;
for *i* = 1 : *M* ***do***
    for *j* = 1 : *N* ***do***
        Lp_train{*i*, 1} = *p_train*(:, :, 1, *i*): Training set data conversion;
        Lp_test{*i*, 1} = *p_test*(:, :, 1, *i*): Test set data conversion;
    end
end
Establishing a blank ***network*** structure;
Create *input layer, folding layer, convolutional layer, activation layer, BiLSTM layer, fully connected layer*;
Output *input layer, folding layer, convolutional layer, activation layer, BiLSTM layer, fully connected layer* and connected *multiplication layer*;
Using the ***Adam*** gradient descent ***algorithm***
Setting: *Maximum number of iterations, initial learning rate, descent factor*
Set the *learning rate* value of the model after 700 training sessions
Model *training* and *prediction*
Inverse normalisation of sample data
Calculate the *Root Mean Square Error, Goodness of Fit, MAE, MAPE, MBE, MSE* for the data set

Showing the *network* structure of the *CNNBiLSTM-AM model*
If the values of the *Root Mean Square Error, Goodness of Fit, MAE, MAPE, MBE* and *MSE* coefficients of the model are taken as reasonable ***then***
  Build up the final *price forecast model*
  ifelse
    Adjusting the *parameters* of the model
  end
end
Exporting *data feature information* and Drawing *model construction*

Based on the combined model and the computer process algorithm to achieve the prediction of the coupled economic and educational values for 14 regions in western China in 2024. The prediction results are shown in Table 6 and Figure 4 below.

Table 6. Prediction results based on combined models

Year	Province	Coupling values	Goodness of fit	Prediction error
2024	Inner Mongolia	0.294987136	0.844026998	0.001884915
2024	Guangxi	0.636220954	0.838717073	0.003224437
2024	Chongqing	0.561790691	0.871974271	0.004683447
2024	Sichuan	0.844062304	0.818255647	0.008813402
2024	Guizhou	0.276421667	0.838637171	0.00291284
2024	Yunnan	0.401402161	0.85423036	0.004056821
2024	Tibet	0.103383164	0.84223541	6.573E-05
2024	Shaanxi	0.470324639	0.860424188	0.004080126
2024	Gansu	0.151267487	0.842215422	0.001310458
2024	Qinghai	0.072370787	0.850614421	0.000140195
2024	Ningxia	0.101463738	0.803450009	0.00033195
2024	Xinjiang	0.230623903	0.829064753	0.001275033

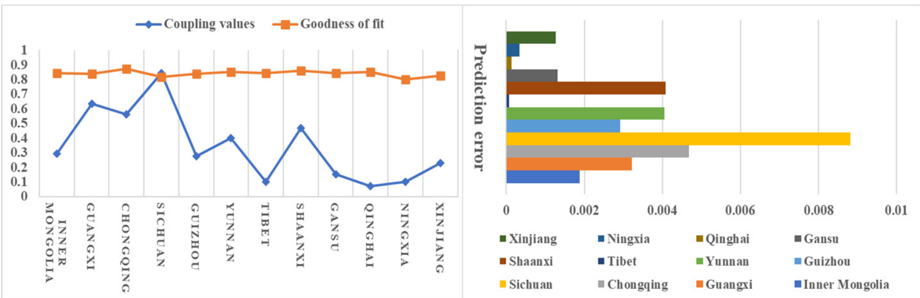


Figure. 4. Distribution of prediction results based on the combined model

In Table 6, column 3 shows the prediction values based on the combined model, and column 4 shows the degree of fit of the model predictions. We can know that the fit values for each region are all higher than 0.80, which indicates that the model has a good fit. In addition, the prediction errors of the models are all lower than 0.01, which also reflects that the combined CNNBiLSTM-AM model constructed in the article has good robustness and generalization.

6 Conclusions and Recommendations

6.1 Conclusions

The article analyses the correlation between educational attainment and regional economic development in 12 provinces and cities in western China by constructing an entropy value method + coupling coordination degree model. The results found that the degree of coupling coordination between educational attainment and regional economic development levels gradually increased over time in 12 provinces and cities in western China, but the degree of higher education and economic development levels were not coordinated in some regions. In addition, the coupling coordination index of these 12 provincial and municipal regions all exceeded 0.3, with Chongqing, Sichuan, Guangxi, Guizhou, Yunnan, and Shaanxi provinces showing a high degree of coupling.

Due to the complex topography and harsh natural conditions in western China, as well as the existence of some objective historical and geographical factors, most of the re-regions have a late start in economic development and a relatively weak foundation for higher education. The lack of high-level talents has also become an important factor affecting the economic development of the Western region, which to a certain extent limits the economic development of the region and thus affects sustainable economic development.

The article constructs a combined CNNBiLSTM-AM model containing a convolutional layer, a BiLSTM layer, and an attention mechanism layer based on the traditional CNN model and LSTM model to realize the coupling value prediction of 12 regions in western China in 2024, among which the coupling value of Sichuan Province is 0.8441, which is the highest value in the prediction period of 2024, indicating that the western region of China Among the 12 regions, Sichuan Province has the best coordination between education and economic development; in addition, the article also introduced four indicators, MAE, MAPE, MBE, and MSE, to evaluate the prediction performance of the model. The prediction error of the model was found to be less than 0.01 for all 12 regions, and the fit of the model was greater than 0.80. This indicates that the model has good robustness and generalization ability.

6.2 Recommendations

Synthesizing the findings from the study, the article makes the following three recommendations.

First, increase capital investment in higher education and science and technology innovation. Gradually increase the capital investment in higher education. The overall level of higher education and science and technology innovation development in the western provinces and cities is still low, and capital investment in the region should be increased, while local governments and universities should establish a cooperation mechanism to build a bridge for two-way talent training. To this end, the government can conduct performance appraisals on the research and innovation aspects of universities, allocate funds based on the appraisals, increase the rewards for teachers with outstanding research achievements on top of the previous ones, increase the funding for teachers' research incentives, and increase the support for students in difficulty.

Secondly, always pay attention to the degree of development of higher education and the quality of teaching. We should take into account the special industries of the western provinces and cities as well as the needs of each industry in terms of specialization, and train the scientific and technological talents needed by society by the actual situation of social development. Of course, it is also necessary for the relevant government departments to give policy incentives to universities to develop higher education to promote the increase of scientific and technological achievements and to transform technological achievements into endogenous power for economic growth.

Thirdly, we should correctly recognize the imbalance of regional development and realize the leap of strong regions leading to weak regions. Establish a regional cooperation mechanism to strengthen the exchange and cooperation of advantageous resources between regions, to narrow the development differences between regions, and realize the sharing of educational resources and coordination of economic development between regions.

References

[1] Hua Fan. An empirical and comparative study on the impact of higher education on economic growth[J]. Liaoning Education Research,2006(02):46-48.

[2] Zhuo Lu. Higher education hierarchy, technological progress, and high-quality economic development[J]. Statistics and Decision Making,2023,39(01):115-119.

[3] Jingming Chi, Xiaofang He, Wen Cheng, Zhu Yan, Xia Li. An empirical study on the relationship between higher education hierarchy and economic development[J]. Education and Economy,2010(01):1-7.

[4] Liqiang Guo, Nainan Zhang. Research on the coordination of higher education and economic development[J]. Heilongjiang Higher Education Research,2018,36(07):43-48.

[5] Hongjiao Yu. An empirical analysis of coordinated development of higher education with economy, science, and technology in Henan Province[J]. Journal of Henan University of Science and Technology,2012(02):40-43.

[6] Wenbo Cai, Niannian Chen. The coupling and coordination effects of higher education, science, and technology innovation and economic growth in the Yangtze River Economic Zone city cluster[J]. Modern Education Management,2022,392(11):33-42.

[7] UIS(UNESCO Institute for Statistics). Global education digest 2009: Comparing education statistics across the world[EB/OL]. [2010].

[8] Bertoletti Alice, Berbegal-Mirabent Jasmina, Agasisti Tommaso. Higher education systems and regional economic development in Europe: A combined approach using econometric and machine learning methods[J]. Socio-Economic Planning Sciences, 2022, 82:1-18.

[9] Tin-Chun Lin. The role of higher education in economic development: an empirical study of Taiwan case[J]. Journal of Asian Economics, 2004: 355-371.

[10] Agasisti T, Bertoletti A. Higher education and economic growth: A longitudinal study of European regions 2000-2017[J]. Socio-Economic Planning Sciences, 2020(1):100940.

[11] Nanxing Zhang, Shu Wang, Jihong Sun. Research on the evaluation of China's comprehensive development level of higher education and regional differences[J]. Education Research,2014,35(05):28-36.

[12] Jihong Sun, Xiaojiang Yang, Rongnan Miao. An empirical analysis of the comprehensive evaluation of regional higher education development[J]. Science and Technology Management,2009,30(12):122-127.

[13] Chengxing Xie, Fengxiao Wang. The dynamic relationship between higher education and economic development in Xinjiang in the context of "One Belt, One Road" based on entropy TOPSIS and VAR model[J]. Frontier Economy and Culture, 2022(04):10-15.