# Construction and Application of an MR-BP Neural Network Combined Forecasting Model Based on Adaptive Weighted Fusion

**Yao Wang, Zhixiang Yang***
*School of Mathematics and Data Science, Changji College, Changji, Xinjiang, China*
**Corresponding Author*

**Abstract: In view of the deficiency of traditional multiple regression model in dealing with nonlinear relations, and the poor interpretability and high training cost of BP neural network model, this paper proposes a combined forecasting model combining multiple regression and BP neural network, which takes into account the advantages of linear trend catching and nonlinear relationship mining. Based on the data of GDP per capita and related macroeconomic indicators of China from 2005 to 2023, this paper constructs a fusion framework to co-optimize the analytical ability of multiple regression and the fitting ability of neural network. The results show that the combination model proposed in this paper can fit the complex trend of GDP per capita more accurately, and is significantly better than other comparative models in prediction accuracy and stability, and effectively reduces the prediction error.**

**Keywords: Multiple Regression; BP Neural Network; Combined Model; GDP Forecasting**

## 1. Introduction

As a core indicator for measuring the economic development level and residents' well- being of a country or region, gross domestic product (GDP) per capita has crucial practical significance for formulating macroeconomic policies, optimizing industrial layouts, and guiding corporate strategic investment. Especially against the background of profound changes in the current global economic structure and increasing uncertainty, achieving reliable forecasts of per capita GDP has become a focus of academic attention [1]. In order to achieve accurate prediction of economic indicators, many researchers have put forward a variety of models and methods. Traditionally, multiple linear regression model is widely used in the field of economic forecasting because of its clear principle and clear economic meaning, which can effectively reveal the linear structural relationship between variables [2].

However, the model is essentially based on linear assumptions, and it is difficult to capture the complex nonlinear dynamic characteristics in the economic system, which leads to the bottleneck of its prediction accuracy when dealing with high-dimensional and nonlinear data. On the other hand, the machine learning method represented by BP neural network, with its powerful adaptive learning and nonlinear mapping ability, shows significant advantages in complex pattern recognition and prediction [3].However, this kind of "black box" model often lacks economic interpretability, and relies heavily on data quantity and quality, and the training process is easy to fall into local optimum, which limits its application in rigorous macroeconomic analysis to some extent. In order to give consideration to the prediction accuracy and interpretability of the model, researchers began to explore a hybrid modeling path that combines traditional econometric models with machine learning algorithms. For example, some scholars try to use lasso-logistic regression analysis to predict shoulder-hand syndrome in elderly stroke patients, or use statistical analysis to analyze the adiabatic temperature rise model of concrete [4-5].The research on deep coupling of multiple linear regression and BP neural network, especially through dynamic weight optimization mechanism, is still in the exploratory stage, and its theoretical framework and application efficiency need to be further explored [6].

In recent years, researchers have begun to explore the combination of traditional statistical models and machine learning methods to give consideration to the prediction accuracy and interpretability of the models. Some achievements have been made in the combination of linear regression with nonlinear models such as decision tree and support vector machine. However, the research of combining multiple regression model with BP neural network is still in its infancy, and the related theory and application need further exploration. Multiple linear regression is the cornerstone of predictive modeling, and Wang Wei and others use multiple linear regression to model the compressor performance, which shows the effectiveness of traditional methods in scenarios with clear mechanism and approximate linear assumption [7]. However, when the independent variables are multicollinearity or the dimensions are high, the stability of the model will decrease. Therefore, regularization methods such as Lasso and Ridge are widely used. For example, Liu Lina et al. and Feng Jianwen et al. used Lasso-Logistic regression to construct medical prediction models respectively, and made variable selection while estimating parameters through L1 regularization, which improved the generalization ability and simplicity of the model [8,9].

The Logistic regression prediction model constructed by Zhang Yali et al. also follows a similar clear and interpretable modeling path [10]. These efforts represent a research paradigm that addresses complexity by improving traditional models themselves. Liu Xuexiang et al. used an improved RBF neural network for short-term power load forecasting [11]; the RBF network, due to its local approximation characteristics, converges quickly and has advantages in time-series forecasting. Qin Yanjiao's comparative study clearly shows that when dealing with problems affected by complex, multi-factor nonlinearity such as housing prices, the forecast accuracy of ensemble learning algorithms such as random forests is significantly better than that of traditional multiple linear regression [12]. These studies confirm the excellent ability of machine learning models to capture complex patterns in data. A comprehensive analysis of existing studies shows that although hybrid models have become a trend, there is still ample room for combining BP neural networks with multiple regression in an "embedded", low-cost manner. Against this background, this paper proposes a multiple regression model optimization algorithm based on a BP neural network: by embedding the BP neural network into the multiple regression model, it improves forecast accuracy, retains the interpretability of the multiple regression model, and reduces the training cost of the neural network.

Based on the above, this paper proposes a combined forecasting model based on multiple linear regression and a BP neural network. Using China's per capita GDP and related macroeconomic indicators from 2005 to 2023 as samples, this study aims to synergistically optimize the linear analytical capability of the regression model and the nonlinear fitting capability of the neural network by constructing a weighted combination framework. This paper not only elaborates the model construction process in detail, but also conducts comparative empirical studies with several models including ARIMA, decision tree, and ensemble learning, to comprehensively evaluate the performance of the proposed combined model in terms of forecast accuracy and stability, thereby providing a new method with both high accuracy and strong robustness for macroeconomic forecasting research.

## 2. Related Theories and Methods

### 2.1 multiple Regression Model

The main principle of the multiple regression model is to construct a regression equation for a system containing multiple explanatory variables and one dependent variable. Relationships among some events involve multiple aspects, leading to various factors that drive event development. Therefore, using an optimal equation composed of multiple explanatory variables to predict the dependent variable yields more accurate results that are more consistent with actual parameter estimation than using a single explanatory variable. The core of the multiple regression model is to construct a function that minimizes the sum of squared differences between true values and predicted values. Typically, this method constructs a random linear relationship between X and Y, estimates parameters using the ordinary least squares method, conducts significance tests on the parameters, and finally uses the regression equation to forecast data, generally assessing the accuracy of the regression equation by forecasting existing historical data. The multiple regression model is commonly used to describe the random relationship between variable Y and variables X, where the dependent variable is Y and the independent variables are X1 、 X2、…、Xn. The equation is:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n \quad (1)$$

The parameter estimation formula is $\beta = (X'X)^{(-1)X'Y}$, where β is the parameter vector, X is the design matrix, and Y is the dependent variable vector.

## 2.2 Introduction to BP Neural Network

The BP neural network is a multilayer feedforward network with hidden layers. It is the core part of feedforward networks and embodies the essence of artificial neural networks. The learning process of a BP neural network consists of two stages: forward propagation and backpropagation. During forward propagation, input information is transmitted downward layer by layer according to the order of hidden layers, finally obtaining the forecast result at the output layer. Input data are transmitted layer by layer starting from the input layer, and neurons in each layer perform a weighted summation of the received inputs and apply a nonlinear transformation through an activation function. The mathematical expressions are:

$$z^{(l)} = W^{(l)} a^{(l-1)} + b^{(l)}$$
$$a^{(l)} = f(z^{(l)}) \quad (2)$$

Where $W^{(l)}$ and $b^{(l)}$ are the weight matrix and bias vector of the l-th layer, respectively, and $f(\cdot)$ is the activation function. Information passing to one layer affects the next layer but does not affect the transmission of the previous layer.

During training, the BP neural network can perform error backpropagation. When the obtained output does not meet the error requirement, the error signal is propagated back along the original path, and the network is retrained until the output satisfies the error requirement. In the backpropagation stage, the system first calculates the error between the predicted value and the true value using a loss function (such as mean squared error $E = \frac{1}{2}\|\hat{y} - y\|^2$ ), and then uses the chain rule to propagate the error backward from the output layer to the input layer. The error term for the output layer L is:

$$\delta^{(L)} = (a^{(L)} - y) \odot f'(z^{(L)}) \quad (3)$$

The error term for hidden layers is propagated backward layer by layer using:

$$\delta^{(l)} = ((W^{(l+1)})^T \delta^{(l+1)}) \odot f'(z^{(L)}) \quad (4)$$

After obtaining the error terms, the gradients of the loss function with respect to the weights and biases of each layer are computed as:

$$\frac{\partial E}{\partial W^{(l)}} = \delta^{(l)} \left(a^{(l-1)}\right)^T$$
$$\frac{\partial E}{\partial b^{(l)}} = \delta^{(l)} \quad (5)$$

Finally, based on the gradient descent algorithm, the network parameters are iteratively updated with a set learning rate:

$$W^{(l)} \leftarrow W^{(l)} - \eta \frac{\partial E}{\partial W^{(l)}}$$
$$b^{(l)} \leftarrow b^{(l)} - \eta \frac{\partial E}{\partial b^{(l)}} \quad (6)$$

Thereby continuously reducing the forecast error and enabling the model to gradually approach the true mapping relationship. The specific steps of the process are illustrated in Figure 1 below.

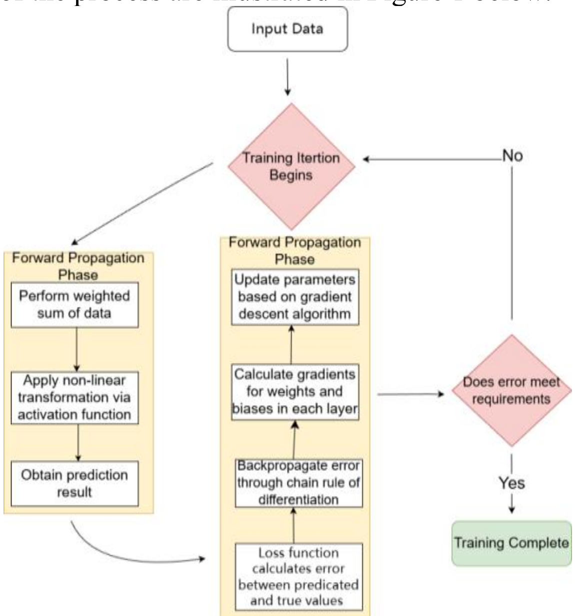


**Figure 1. Flowchart of the Working Principle of BP Neural Network**

## 3. Construction of the MR- BP Combination Model

The core idea of the MR-BP combination model proposed in this study is to capture the trend of linear subjects through a parallel fusion framework, to learn the nonlinear patterns in the residuals by using BP neural network, to train the multivariate linear regression module and BP neural network module cooperatively, and to dynamically weight their outputs by using a learnable weight parameter, so as to realize the deep fusion of linear laws and nonlinear features, and finally to generate an integrated model with stronger prediction ability and robustness. The realization of this model is a process of phased training and overall fine-tuning, and the specific construction ideas are as follows.

### 3.1 Model Structure

This parallel combination model consists of two main parts:

(1) Linear module (multiple regression part): This module aims to capture the stable global linear trend between input features and the target variable. Its output is expressed as:

$$y\_linear = W * X + b \tag{7}$$

where $W=[w1 \ ,w2 \ ,...,wp \ ]^T \in R$ is the weight coefficient vector, b∈R is the bias term, and X is the input feature vector.

(2) Nonlinear module (BP neural network part): This module consists of a multilayer feedforward neural network, used to mine complex local nonlinear relationships and interaction effects in the data. Its structure is defined as:

$$h1 = ReLU(W1 * X + b1)$$
$$h2 = ReLU(W2 * h1 + b2)$$
$$y\_nn = W3 * h2 + b3 \tag{8}$$

where W1,W2,W3 are the weight matrices from the input layer to the first hidden layer, from the first hidden layer to the second hidden layer, and from the second hidden layer to the output layer, respectively; b1,b2,b3 are the corresponding bias terms. The activation function uses ReLU to enhance the nonlinear expressive ability of the model and alleviate the vanishing gradient problem.

(3) Combined output layer: The final forecast of the model is the weighted sum of the outputs of the linear module and the nonlinear module:

$$y\_final = \alpha * y\_linear + (1 - \alpha) * y\_nn \tag{9}$$

where y_fina is the final forecast value of the model, and α is a learnable combination weight parameter in the range [0,1]. The value of this parameter is that it can adaptively adjust the proportion of linear and nonlinear contributions in the final forecast.

### 3.2 Training Strategy

To optimize the model, this study adopts a three-stage training strategy:

Stage 1: Module pre-training. First, the linear regression module and the BP neural network module are trained separately. In this stage, the parameters of the other module are frozen respectively, and each module is initially converged using the training data to obtain its basic predictive ability. This step aims to provide a good parameter starting point for subsequent fusion.

Stage 2: Weighted fusion. After pre-training, a learnable combination weight α is introduced. The outputs of the linear module and the nonlinear module are combined using the weighted formula, and the impact of different α values on overall prediction performance is preliminarily evaluated on a validation set to set a reasonable initial value.

Stage 3: Overall fine-tuning. This is the key step of model optimization. All previously frozen parameters (including W,b of the linear module, W1,b1,W2,b2,W3,b3 of the neural network module, and the combination weight α) are thawed, and the entire combined model is treated as a unified computational graph. Using the backpropagation algorithm with the objective of minimizing the overall prediction loss (such as mean squared error), all parameters are collaboratively fine-tuned in an end-to-end manner. This process enables the linear and nonlinear parts to cooperate and adjust each other, thus obtaining a set of globally better model parameters.

Through the above construction process, the finally obtained equation for y_final is the optimized MR-BP combined forecasting model, which possesses both the robustness of a linear model and the strong fitting ability of a neural network model.

## 4. Empirical Study and Result Analysis

### 4.1 Data Sources and Variable Selection

To explore the growth drivers of China’s per capita GDP and make accurate forecasts, this study selects the period 2005-2023 as the research period. This period covers the critical transition of China’s economy from high-speed growth to high-quality development, and the data are well representative. All research data are sourced from the China Statistical Yearbook (2006-2024) published by the National Bureau of Statistics, ensuring the authority and continuity of the data (Figure 2).

In variable selection, referring to macroeconomic theory and relevant econometric research, the dependent variables and independent variables are finally determined as follows: the dependent variable is per capita GDP (yuan), which is the core index to measure economic development and residents 'well-being; the independent variables are per capita consumption expenditure of residents (yuan), investment in fixed assets of the whole society (hundred million yuan), government budget expenditure (hundred million yuan), total export trade (hundred million yuan) and total employment population (ten thousand people)

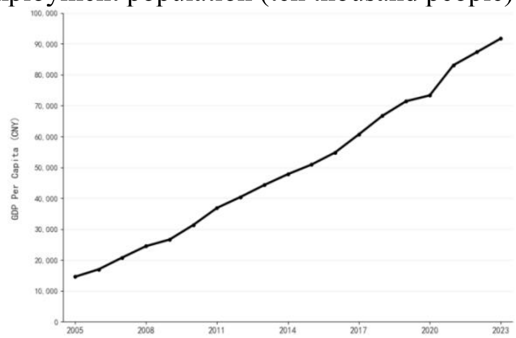


**Figure 2. Trend Chart of China's Annual Per Capita GDP**

**4.2 Data Pre-processing**

To ensure the stability and validity of the model training, the following preprocessing steps were performed on the raw data:

(1) Stationarity test and treatment: Because macroeconomic time series often have nonstationarity, direct regression may lead to "pseudo-regression" problem.

$$\Delta Y_t = \alpha + \beta t + \rho Y_{t-1} + \sum_{i=1}^{p} \gamma_i \Delta Y_{t-i} + u_t \quad (10)$$

Among them: ρ is the unit root test parameter, ρ=0 indicates the existence of unit root, α is the drift term, reflecting the average change of the sequence, β is the time trend term, reflecting the deterministic trend, $\gamma_i$ is the lagged difference term coefficient, reflecting short-term dynamic adjustment, and p is the lag order selected by the BIC criterion.

The test results (all p values are greater than 0.1) show that all the original series are non-stationary series. To solve this problem, all variables are natural logarithm. ADF test is performed on the logarithm series again, and the results show that all the original series are stationary series (all p values are less than 0.05), which meets the basic requirements for stationarity of modeling. Subsequent modeling analysis is based on the processed stationary series.

(2) Data standardization: Due to the huge difference in variable dimensions (for example, fixed asset investment is trillions of yuan, while employment population is ten thousand people), and BP neural network is very sensitive to the scale of input data.

Let the original feature matrix be $X \in \mathbb{R}^{N\times p}$, where the statistic for the j-th column (corresponding to the j-th feature) is:

$$u_j = \frac{1}{N} \quad \sum_{i=1}^{N} x_{ij}, \sigma = \sqrt{\frac{1}{N-1}\sum_{1}^{N}\left(x_{ij} - u_j\right)^2} \quad (11)$$

When $\sigma_j \gg \sigma_k$, during backpropagation in the neural network, the weight update is:

$$\Delta \quad \omega_{jk} \propto \frac{\partial \mathcal{L}}{\partial \omega_{jk}} \quad (12)$$

where $h_j$ is the activation value of the previous layer. If the input feature scales are not uniform, the gradient magnitudes corresponding to different features will vary significantly. In this study, Z-score standardization was performed on the stationary log-transformed data. This method transforms the data into a distribution with a mean of 0 and a standard deviation of 1, with the calculation formula as follows:

$$X_{scaled} = \frac{X_{ij} - \mu_j}{\theta_j} \quad (13)$$

where μj is the sample mean of the j-th feature, and θj is the sample standard deviation of the j-th feature.

After standardization, the condition number of the Hessian matrix of the loss function was significantly improved, approaching the identity matrix, thereby accelerating gradient descent convergence. The convergence speed is related to the condition number κ(H):

$$\text{Convergence speed} \propto \frac{\kappa(H)-1}{\kappa(H)+1} \quad (14)$$

The optimal convergence speed is achieved when κ(H)=1. This standardization processing can accelerate model convergence and prevent certain features from dominating model training due to their large numerical values. Through the above preprocessing steps, a stationary and dimensionally unified standardized dataset was obtained, laying a reliable data foundation for the subsequent construction of a high-precision MR-BP combined forecasting model.

**4.3 Model Training and Prediction**

In order to ensure the objectivity and generalization ability of model evaluation, this study adopts a standardized process for model training and prediction, and compares and analyzes the proposed MR-BP combined model with many mainstream prediction models. The data of this study spans from 2005 to 2023, with a total of 19 sample years. In order to simulate the real time series prediction scene, the data set is divided into: training set (2005-2018): a total of 14 years of data for learning and training model parameters; Verification set (2019-2020): 2 years' data, which is used to adjust the model superparameters (such as the learning rate of neural network, the number of layers, the initialization of combination weight α, etc.) and prevent the model from over-fitting; Test set (2021-2023): 3 years' data, as "unknown" data that did not participate in model training at all, is used to finally and objectively evaluate the generalization performance and prediction accuracy of all models.

To intuitively evaluate the fitting and forecasting effects of each model, this study plotted a comparative distribution diagram of the actual observed values, the forecasted values of the traditional multiple regression model, and the forecasted values of the MR-BP combined model on the test set (2021-2023). The comparison between forecasted and actual values is shown in Figure 3.

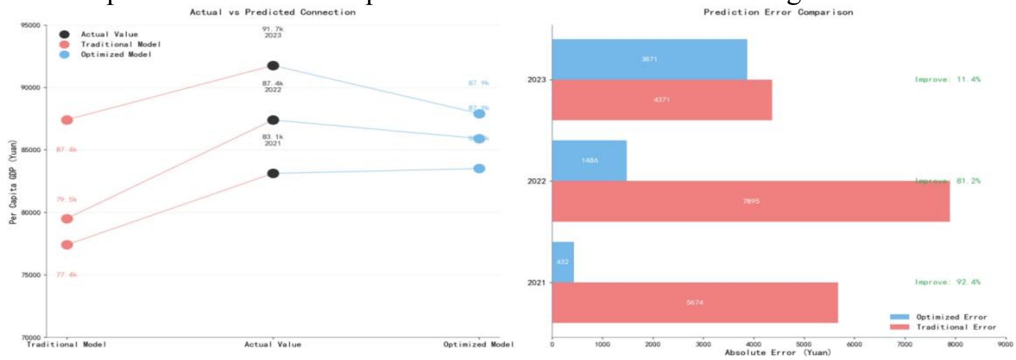


**Figure 3. Comparison of Predictions from Traditional Regression/Optimized Models with Actual Data**

It can be clearly observed from the overall distribution pattern that the forecast curve of the MR-BP combined model fits most closely with the actual curve of real per capita GDP. Whether in tracking the trend or matching specific values, the combined model demonstrates significantly superior performance to the traditional multiple regression model (whose forecast curve usually deviates considerably from the actual values). This visualization result preliminarily proves that by introducing the BP neural network to capture nonlinear patterns, the combined model effectively enhances the ability to characterize the dynamics of complex economic systems, making its forecasts closer to the actual operating trajectory of the economy.

### 4.4 Model Forecast Error Analysis

To quantitatively evaluate the forecasting accuracy of the models, this study calculated the forecast error rates of the traditional multiple regression model and the MR-BP combined model for each year in the test set (2021-2023). The results are shown in Table 1.

**Table 1. Comparison of Model Forecast Error Rates (2021-2023)**

| Modle | 2021 | 2022 | 2023 | Average Error Rate |
|---|---|---|---|---|
| Traditional Model | 6.83% | 9.00% | 4.76% | 6.86% |
| MR-BP Model | 0.51% | 1.71% | 4.23% | 2.15% |

Analysis shows that the MR-BP combined model consistently achieved lower error rates in all forecast years. Specifically, its error rates in each year were significantly reduced compared to the traditional multiple regression model, especially in 2021, where the error rate dropped sharply from 6.83% to 0.51%. In terms of overall performance, the combined model achieved an average error rate of 2.15%, far lower than the traditional model's 6.86%, representing a forecast accuracy improvement of approximately 68.7%. This quantitative result strongly confirms that the MR-BP combined model, by integrating the advantages of linear and nonlinear modeling (Figure 4), can significantly reduce forecast bias and possesses superior forecast accuracy and stability.

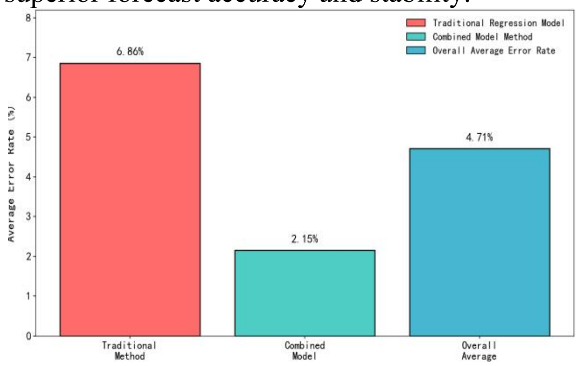


**Figure 4. Comparison of Error Rates between Traditional Regression and Optimized Models**

Based on the above forecast results and error analysis, it can be clearly concluded that the fusion framework combining the BP neural network with the traditional multiple regression model has achieved remarkable results in the task of per capita GDP forecasting. This MR-BP combined model is not a simple superposition of models but achieves deep optimization of the model parameter system through its unique collaborative training and weighted integration mechanism.

Specifically, the optimization effects are mainly reflected in two aspects: First, in terms of forecast accuracy, the combined model successfully absorbs the robustness of multiple regression in capturing macro-level linear trends and the powerful ability of neural networks to mine complex nonlinear relationships, so that the deviation between the final forecasted value and the actual value is systematically reduced, with the average error rate significantly lowered to 2.15%, fully demonstrating the great potential of this fusion strategy in improving forecast accuracy. Second, in terms of model robustness, the combined model performs stably across annual tests, with its forecast curve closely fitting the actual values, effectively avoiding the significant fluctuations in forecast performance exhibited by the traditional regression model in some years (such as 2022), demonstrating superior generalization ability.

In summary, this model optimization, by introducing the BP neural network, effectively compensates for the lack of flexibility in traditional econometric models, constructing an enhanced forecasting framework that possesses both explanatory power and predictive performance, providing an effective new method for the accurate forecasting of macroeconomic indicators.

### 4.5 Model Robustness Test

To evaluate the predictive performance of a single nonlinear model, this study constructed and trained a standard BP neural network model on the same dataset. The model's predictive performance on the test set (2021-2023) was unsatisfactory. Data analysis indicates that the forecast error rates of the single BP neural network model remained consistently high (12.99% - 14.50%), with an average error rate as high as 13.90%, significantly higher than that of the aforementioned MR-BP combined model (2.15%).

This result reveals that although the BP neural network possesses strong nonlinear fitting capabilities, in the context of the limited sample size of this study, its standalone use may struggle to effectively capture and generalize the complex macroeconomic patterns behind per capita GDP, making it prone to overfitting or local optima, thus leading to unsatisfactory out-of-sample forecast accuracy. This conclusion confirms from the opposite perspective the necessity and superiority of the MR-BP framework, which combines a neural network with a linear model having strong theoretical foundations, as the fusion model effectively avoids the potential drawbacks of a single neural network.

## 5. Comparison with Other Model Forecasts

### 5.1 BP Neural Network

Using the BP neural network model to forecast per capita gross domestic product (yuan), and employing the data partitioning method described above, the forecast results for per capita gross domestic product (yuan) from 2021 to 2023 show an error rate of 12.99% in 2021, 14.2% in 2022, and 14.5% in 2023, with an average error rate of 13.9%. The standalone BP neural network model has low forecast accuracy for this type of data, and the data cannot be accurately predicted (Figure 5).

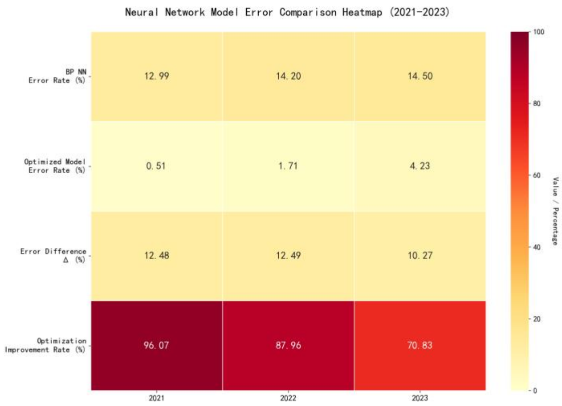


**Figure 5. Comparison of Error Rates for a Single BP Neural Network**

### 5.2 ARIMA Model

In forecasting the data using the ARIMA model, a classic method for time series forecasting, this study constructed an ARIMA model for comparative analysis. Through unit root testing, the series was determined to be first-order integrated. On this basis, grid search was used for iterative training, and the optimal model parameters were ultimately determined to be

ARIMA (1,1,1). As shown in Figure 6, the model's forecast values on the test set (2021-2023) are 77,244.34 Yuan, 81,150.64 yuan, and 85,056.91 yuan, respectively.

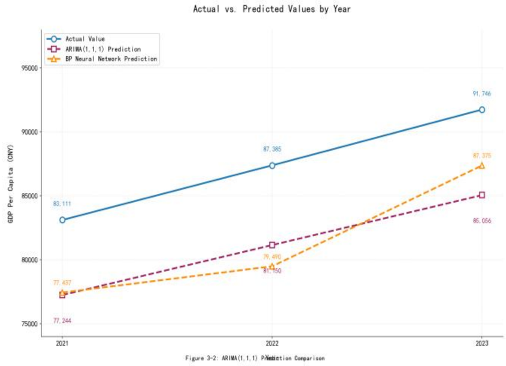


**Figure 6. Comparison of ARIMA (1,1,1) Predictions**

Model evaluation shows that the forecast error rates of the ARIMA(1,1,1) model for each year range between 7.05% and 7.29%, with an average error rate of 7.16%. These results indicate: First, the ARIMA model can capture the time series trend characteristics of per capita GDP relatively well, with its forecast accuracy significantly superior to that of the single BP neural network model (13.90%), reflecting the applicability of traditional time series methods in macroeconomic forecasting. Second, its error rate remains significantly higher than that of the MR-BP combined model (2.15%), primarily because the ARIMA model relies only on the historical information of the series itself and fails to incorporate the explanatory effects of key economic indicators such as household consumption and fixed asset investment. Therefore, when facing a complex economic system influenced by multiple factors, its forecasting capability has inherent limitations.

### 5.3 Ensemble Learning

As an advanced machine learning paradigm, ensemble learning's core idea is to improve the generalization ability and prediction accuracy of the whole model by combining the prediction results of several basic models and using "group wisdom". In this study, Stacking integration strategy is adopted, in which random forest regression and gradient lifting regression are used as basic models to capture diversified pattern features from data; As a meta-model, the linear regression model is used to learn and optimize the optimal combination of the prediction results of the basic model. The absolute prediction errors of the integrated model on the test set (2021-2023) are 588.52 Yuan, 3685.48 yuan and 8046.48 yuan respectively, and the annual error rate and average error rate are shown in Figure 7.

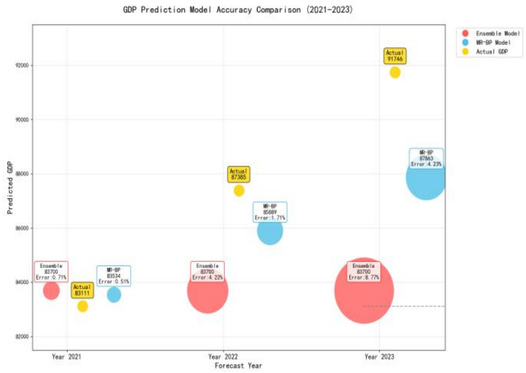


**Figure 7. Comparison of Ensemble Learning Predictions**

Model evaluation results show that the ensemble learning achieves an average error rate of 4.56%, with its forecast accuracy significantly superior to the aforementioned single BP neural network model (13.90%) and ARIMA model (7.16%). On the one hand, this verifies that the ensemble learning method, by integrating the advantages of different algorithms, can effectively enhance the robustness and accuracy of the model. On the other hand, it is observed that the model's errors show an expanding trend in recent forecasts, which may reflect its limitations in capturing the long-term dynamic changes of the economic system. Nevertheless, the overall performance of ensemble learning still proves its application value in the economic forecasting field, yet its accuracy still fails to surpass the MR-BP combined model (2.15%) proposed in this paper, further highlighting the innovativeness and effectiveness of the deep integration of linear models and neural networks.

### 5.4 Comparative Analysis of Errors under Different Algorithm Results

Summarizing the above comparative experimental results, it can be seen that in the task of forecasting per capita gross domestic product, the multiple regression optimization model based on BP neural network (MR-BP) proposed in this paper demonstrates significant performance advantages. By systematically comparing the forecasting effects of multiple methods, including the single BP neural network, the ARIMA time series model, the ensemble learning model, and the traditional multiple regression model in Figure 8, it can be found that:

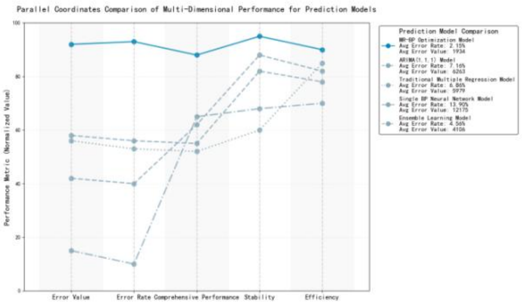


**Figure 8. Comparison between the Optimized Model and Other Models**

First, the MR-BP combined model, with an average error rate of 2.15%, significantly outperforms all comparison models. Its forecast accuracy is improved by approximately 68.7% compared to the traditional multiple regression model (6.86%), by approximately 70% compared to the ARIMA model (7.16%), by approximately 84.5% compared to the single BP neural network (13.90%), and also shows a marked improvement over the ensemble learning model (4.56%).

Second, from the perspective of forecast stability, the fluctuation range of the MR-BP model's forecast error rates over the three years of the test set (0.51% - 4.23%) is significantly smaller than that of the other comparison models, demonstrating more robust forecasting performance.

In summary, although the various comparison models possess certain predictive capabilities to varying degrees, none of them surpass the MR-BP combined model proposed in this paper in the two key dimensions of forecast accuracy and stability. This result fully proves the innovative value of integrating neural networks with classical statistical models: it retains the interpretability advantages of traditional regression models while significantly improving forecasting performance through the nonlinear compensation mechanism of the neural network, providing a more effective solution for the accurate forecasting of macroeconomic indicators.

## 6. Conclusions and Future Work

### 6.1 Main Research Conclusions

This article innovatively integrates the multiple linear regression model with the BP neural network to construct a new combination forecast model, and conducts systematic verification based on China's per capita GDP and related macroeconomic indicators from 2005 to 2023. Empirical research shows that the combined model is significantly better than traditional multiple regression, single BP neural network, ARIMA and ensemble learning comparison models in terms of prediction accuracy and stability. Specifically, its average prediction error rate is controlled at 2.15%, which is a significant improvement of 67% and 84% respectively compared with the traditional regression model (6.86%) and the single BP neural network model (13.9%), which fully reflects the superiority of collaborative modeling of linear and nonlinear features.

In terms of model performance, the multiple regression module effectively captures the linear laws of indicators such as household consumption and fixed asset investment, while the BP neural network accurately depicts complex interactive effects such as export trade and employment population through nonlinear mapping. The complementary advantages of the two significantly enhance the model's adaptability to the dynamic characteristics of the economic system. It is worth mentioning that this combined model effectively controls the training cost of the neural network while maintaining the interpretability of the multiple regression model parameters through hierarchical training and parameter fine-tuning strategies, successfully solving the dual dilemma of poor interpretability and high computational cost faced by pure black box models in practical applications.

### 6.2 Policy Recommendations and Practical Implications

This model has important application value in the field of macroeconomic forecasting. At the policy level, it is recommended that government departments apply it to GDP forecasting, fiscal budget preparation, and industrial policy evaluation, and simulate economic development paths under different policy scenarios by dynamically adjusting the weights of key variables such as fixed asset investment and government consumption. At the enterprise level, the model can provide support for investment decisions and strategic planning, especially during periods of export trade fluctuations, helping enterprises to adjust production capacity allocation in a timely manner based on the nonlinear response characteristics of the model. In order to ensure the timeliness of forecasts, it is necessary to establish a regular update mechanism, focusing on high-frequency changing indicators such as employment population and trade policies.

In terms of technological development, this model still has room for optimization in many aspects. First, attention mechanisms or reinforcement learning strategies can be introduced to achieve dynamic adjustment of combination weights and enhance model adaptability. Secondly, the model performance can be further improved by embedding the LSTM network to improve the ability to capture temporal features, or combining it with Lasso regression to optimize feature selection. In addition, technologies such as knowledge distillation can be used to achieve lightweight models to meet real-time prediction needs. Future research can focus on three directions: first, carry out cross-domain verification and explore the applicability of the model in fields such as carbon emission prediction; second, integrate Bayesian methods to quantify prediction uncertainty; third, develop a real-time prediction system that integrates data visualization and parameter adjustment functions to promote the transformation of research results into practical applications. This model provides an effective solution for the prediction of complex economic systems. Follow-up research should continue to deepen algorithm optimization and expand application boundaries to give full play to its theoretical value and practical significance.

## Acknowledgments

This paper is supported by Changji College 2025 University-Level College Student Innovation Training Project “Multiple Regression Model Optimization Algorithm Based on BT Neural Network” (No. X202510997199) and Changji College 2025 University-Level Teaching Research and Reform Project “Research on Teaching Model Reform of Data Mining and Algorithms Course Based on OBE Concept and BOPPPS Model” (No. 25JYYB041).

## References

[1] Zhang Xiaoguo, Li Shuang, Li Yajie. Grey Relational Analysis of GDP Forecasting and Industrial Structure in Henan Province. China Market, 2025, (28): 12-16.

[2] Liang Peiting, Wang Binhui, Wang Guochang, et al. Stock Price Forecasting Based on Automatic Identification of Linear or Nonlinear Structures. Journal of Systems Engineering, 2025, 40(04): 574-594.

[3] Qiu Dong, Wang Zhongfei, Zhang Suixing, et al. User Electricity Demand Forecasting Method Based on BP Neural Network. Industrial Control Computer, 2025, 38(10): 111-112.

[4] Wu Tingting. Prediction Model of Shoulder-Hand Syndrome in Elderly Stroke Patients Based on Lasso-Logistic Regression Analysis. Chinese Journal of Hospital Statistics, 2025, 32(04): 255-262.

[5] Yao Ting, Li Hua, Wang Yujiang. Multiple Regression Model of Concrete Adiabatic Temperature Rise Based on Statistical Analysis. Concrete, 2025, (07): 51-54+67.

[6] Xu Min, Lin Weixing, Shi Lei, et al. Prediction of Blasting Vibration in Mining Areas Based on I-GWO-BP Neural Network. Mining Research and Development, 2025, 45(10): 121-128.

[7] Wang Wei, Cai Zhongyuan, Chen Jiongzhen. Compressor Performance Prediction Model Based on Multiple Linear Regression. Gas Turbine Technology, 2025, 38(03): 37-42.

[8] Liu Lina, Song Qixiu, Zhang Lun, et al. Lasso-Logistic Regression Analysis of Direct Oral Anticoagulant-Related Issues and Construction of a Nomogram Prediction Model in Real-World Settings. Practical Pharmacy and Clinical Medicine, 2025, 28(06): 406-412.

[9] Feng Jianwen, Wu Yuzhi. Construction of a Risk Prediction Model for Rheumatoid Arthritis-Associated Interstitial Lung Disease Based on Lasso-Logistic Regression Analysis. Journal of Clinical Pulmonary Medicine, 2024, 29(11): 1638-1643.

[10]Zhang Yali, Chen Binha, Jin Ying. Construction of a Ulcer Recurrence Prediction Model for Patients with Lower Extremity Venous Ulcers Based on Logistic Regression Analysis. Chongqing Medicine, 2025, 54(07): 1576-1580.

[11]Liu Xuexiang, Cui Yangyang, Zhu Hailu. Research on Short-Term Power Load Forecasting Based on an Improved RBF Neural Network. Electrical Technology and Economy, 2025, (10): 4-6.

[12]Qin Yanjiao. Comparative Study of House Price Prediction Models Based on Multiple Linear Regression and Random Forest Algorithm. Modern Information Technology, 2024, 8(22): 127-13

 