Journal of Network Intelligence ©2025 ISSN 2414-8105
Taiwan Ubiquitous Information Volume **10**, Number **1**, February **2025**

Consumer Expenditure Forecasting Based on SPSS Mathematical Statistics and Neural Networks

Ling Zhang*

School of Mathematics and Physics Science
Jingchu University of Technology, Jingmen 448000, P. R. China
13032760445@163.com

Elena Zheng

School of Management
Sripatum University, Jatujak, Bangkok 10900, Thailand
un5044@163.com

*Corresponding author: Ling Zhang

Received January 29, 2024, revised May 3, 2024, accepted June 12, 2024.

Abstract. *Focusing on the issues of low convergence speed and the existence of local minima in the error of the existing consumer expenditure prediction methods, a consumer expenditure prediction method based on SPSS mathematical statistics and neural network is designed. Firstly, the Gaussian function is adopted to enhance the gradient descent algorithm of BP neural network, different learning rates are set for different weights and thresholds of the model, and each parameter is optimized in a targeted way to accelerate the training efficiency and convergence speed of the BP neural network model, and secondly, Lasso's mathematical and statistical method is adopted to perform logarithmic downscaling of the variables of the influencing factors to reduce the complexity of the input variables, and then the downscaled influencing factors as inputs and consumption expenditures as outputs, utilizing the enhanced BP neural network to fit the nonlinear relationship among them, suppressing the influence of noise on the fitting accuracy, thus improving the accuracy of prediction. Finally, the statistical data are empirically analyzed with SPSS software, and the study indicates that the MAE, MAPE, RMSE and RMSPE of the model designed in this paper are 0.0598, 0.0427, 0.0746 and 0.0539, respectively, which are lower than that of the comparison model, and the prediction performance has been effectively improved.*
Keywords: Consumer expenditure forecasting; mathematical statistics; BP neural networks; gaussian functions; gradient descent

---

1. **Introduction.** Recently, in the face of the serious ecological environment and resource problems, people's demand for a better life continues to rise, environmental awareness is getting stronger and stronger, and their living habits are gradually tilted in the direction of green environmental protection [1]. For the purpose of meeting the requirements of sustainable growth and reduce the dependence of products on the environment and resources, enterprises upstream and downstream of the supply chain need to actively improve the greening level of the product supply chain. However, when eco-transitioning, enterprises will invest more costs, and these additional costs will cascade down the supply chain and eventually be transferred to consumers [2, 3]. Whether consumers are willing to bear the extra costs is the main source of driving force for companies to develop green strategies [4, 5].

1.1. **Related Work.** Jun and Park [6] analyzed a large amount of data generated from users' historical purchasing behaviors to categorize and segment users, which can significantly reduce the marketing costs of enterprises, and Schaer et al. [7] proposed to use the number of times and the duration of customers' browsing of online platforms in a certain period of time as an indicator to analyze and predict consumers' behaviors. How to use mathematical and statistical modeling to discover useful information among the huge amount of data is getting more and more attention from the industry. Nowadays, there are many mathematical statistical prediction methods, such as Lasso analysis [8], Principal Component Analysis (PCA) [9], Gray Correlation Analysis (GRA) [10], etc. Ogura [11] proposed the "S" curve model, which is used to predict consumer spending through the mathematical relationships derived from the model. Ervural et al. [12] used an Autore-Gressive Moving Average (ARMA) model to predict consumption during peak hours. Kemalbay and Korkmazoglu [13] used a partial least squares regression method to forecast energy consumption of Chinese transportation sector. Li and Zhang [14] used an optimized discrete Gray Model (GM) to forecast the per capita electricity expenditure in a developed region but the prediction performance was poor. Mishra et al. [15] analyzed by SPSS software that the Lasso method is better than the traditional logistic method for predicting consumption expenditures, and Chen et al. [16] used E-Net for the generalized linear model and selected all the group variables when there are group effects in the data under study. Other forecasting models include Gaussian semi-empirical models [17] and probability density models [18].

The behavior of consumers is often influenced by various complex factors, and the relationship between these factors may be nonlinear. Compared with traditional statistical analysis methods, BP neural network has strong nonlinear fitting ability and can effectively model and predict such nonlinear relationship. Consumption expenditure forecasting methods are evolving from traditional models to artificial intelligence models and combinatorial models. Liu et al. [19] used a least squares Support Vector Machine (SVM) model to forecast energy expenditure. Biswas et al. [20] obtained a high accuracy by training and testing the correspondence using BP neural network on the online consumption data of the last year. Li et al. [21] used a combination model of mathematical statistics Lasso method and SVM to forecast the residents' expenditure, but the training time of the model was added long, and the prediction accuracy was not high. Hu [22] proposed a grey model GM with a BP neural network consisting of a combined model to improve the accuracy of consumption prediction. Kheirkhah et al. [23] adopted PCA to deconstruct the impact of consumption expenditure and used BP neural networks to predict consumption. A Learning Vector Quantization-GARCH-BP neural network was created by Yin and Li [24] to analyze the factors affecting consumption preferences, which proved that better prediction results can be obtained when using BP Neural Network for prediction, which provides the basis for this paper to construct a combined model using BP Neural Network and mathematical statistics and predict the consumption expenditures.

1.2. **Contribution.** From the above analysis, it indicates that the method combining linear prediction and nonlinear prediction of neural network is better. In this article, for the issues of high data complexity and slow convergence speed of neural network in the existing consumer expenditure combination prediction model, a consumer expenditure prediction method based on SPSS mathematical statistics and neural network is designed.

(1) We improve the BP neural network, increases the studying rate of the gradient descent method through Gaussian function and effectively accelerates the union speed of the BP neural network model.

(2) Lasso mathematical and statistical method is applied to logarithmic dimensionality reduction of the variables to reduce the complexity of the input variables. Lasso is a method for data analysis and statistical modeling, and it is a regression analysis technique widely used in variable selection and regularization.

Finally, the empirical analysis is carried out with SPSS software, and the outcome indicates that the model offered in this article has a low relative error and can efficiently realize the accurate prediction of consumption expenditure.

2. Theoretical analysis.

2.1. Lasso method.

Lasso algorithm is a mathematical and statistical method for streamlining the set of indicators [25], which provides a better estimation of the unidentified parameters when choosing the factors, and can greatly address the issue of multiple covariance of the model, particularly for addressing high-dimensional data. Lasso method realizes variable selection by using L1 regularization term in the objective function, that is, thinning the regression coefficient, so as to identify the predictive variables that have a significant impact on the response variables.

Assuming modeled data, consider the usual linear model $(X_{j1}, X_{j2}, \ldots, X_{jq}; Y_j)$.

$$Y_j = \sum_{i=1}^{q} \alpha_i X_{ij} + \xi_j \tag{1}$$

where $j = 1, 2, \ldots, n$ is independently allocated and $E[\xi_j] = 0$, $\{X_j; j = 1, 2, \ldots, m\}$ are independent.

Because of $Y \in R$, centering and standardizing the feedback and predictor variables, respectively, i.e., $Y = X\alpha + \xi$, where $Y_{n\times 1}$ is the response vector, $X_{n\times q}$ is the matrix of independent variables, $\alpha_{q\times 1}$ is the vector of coefficients, and noting that $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_q)^T$ $\xi_{n\times 1}$ are normally distributed random errors.

It is probable to produce regression coefficients which are rigorously equal to 0 and to achieve coefficient assessments that lead to models with great ambulatory energy. The estimation of Lasso's method is defined as follows.

$$\hat{\alpha}^{(Lasso)}(\mu) = \arg\min \sum_{i=1}^{n} \left[\left(Y_i - \sum_{j=1}^{q} X_{ij}\alpha_j \right)^2 + \mu \sum_{j=1}^{q} |\alpha_j| \right] \tag{2}$$

where μ denotes the regression parameter, the calculations in this article chiefly adopt the Lars packet in SPSS, for the selection of coefficients utilizing the Mallows criterion. This method, the statistic theory [26], is as follows.

$$C_p = \frac{SSE_p}{s^2} - n + 2p^2, \quad SSE_p = \sum_{i=1}^{n} (Y_i - Y_{pi})^2 \tag{3}$$

where n represents the number of regressions, s represents the penalty parameter, and p represents the number of independent variables involved in regression.

2.2. BP neural network.

BP neural network algorithm is to divide the learning process into the forward propagation of signals and the reverse adjustment of errors, but there are issues such as local minimum and slow algorithm convergence in network training [26], and its network structure is implied in Figure 1.

Assume that the input layer has M input signals, and any input signal stands for m. The obscured level is with k neurons, each neuron stands for j. The output level is with q neurons, and each neuron stands for q. v_{mj} is used to represent the connection

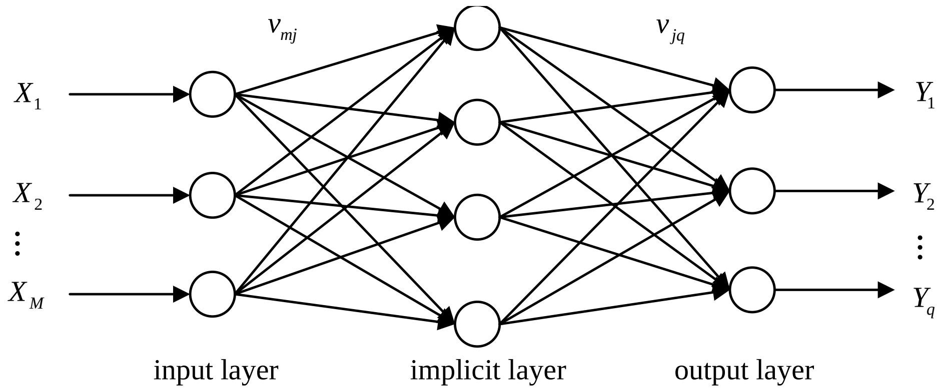

Figure 1. The structure of BP neural network

weight among the obscured level and the input level, and v_{jq} is used to represent the connection weight among the input level and the hidden level. Let the input value be $X = \left[X_1, X_2, \ldots, X_M\right]$, the actual output be $Y = \left[Y_1, Y_2, \ldots, Y_q\right]$, and the expected output be $C = \left[C_1, C_2, \ldots, C_q\right]$.

The overall steps are as follows.

(1) Initialize the weight. Assign a set of relatively small non-zero values to $v_{mj}(0)$ and $v_{jq}(0)$ at random.

(2) Let $X_l = \left[x_{l1}, x_{l2}, \ldots, x_{lM}\right]$ and $l = 1, 2, \ldots, t$ be input vectors and t be the number of training samples. $Y_l = \left[y_{l1}, y_{l2}, \ldots, y_{lq}\right]$ is the actual network output, $C_l = \left[c_{l1}, c_{l2}, \ldots, c_{lq}\right]$ is the expected output.

(3) Input the training sample set $X = \left[X_1, X_2, \ldots, X_l\right]$ sequentially.

(4) Forward propagation. Calculate the network output:
$y_{lq} = \varphi\left(\sum_{i=1}^{I} v_{iq} f\left(\sum_{m=1}^{M} v_{mj} x_{lm}\right)\right)$, and the sample X_l, training error $e_{lq} = c_{lq} - y_{lq}$.

(5) Reverse Propagation. In terms of the error signal, update the weights and thresholds of any level, and determine whether $l > t$, if $l > t$ then go to the next step, otherwise go back to step (3).

BP neural network can effectively deal with complex nonlinear relations, which is very important in the prediction of consumption expenditure, because consumer behavior is often influenced by many factors, and the relationship between these factors is usually nonlinear.

3. **Optimization of BP Neural Networks.** Aiming at the issues of BP neural network algorithm, such as network training is simple to fall under native minima and dense convergence speed, the Gaussian operation is adopted to enhance the gradient descent algorithm of BP neural network, and different learning rates are set for different weights and thresholds of the model, and each parameter is optimized, so that it can effectively accelerate the training effectiveness of BP neural network model and the convergence speed.

Gaussian operation can effectively change the shape of the search space and make the whole optimization process smoother. It can provide a better search strategy for the gradient descent algorithm of BP neural network and help the algorithm converge to the global optimal solution faster.

Firstly, amount of nodes h in the stimulus level, the amount t in the obscured level, the amount of nodes m_t in each obscured level and the amount n in the output level of BP neural network are determined, and the weights and thresholds of any level are initialized to obtain v_{ij}^t and a_i^t. Hidden layer output w_i^t is calculated as follows:

$$w_i^t = f\left(\sum_{j=1}^{m_{t-1}} v_{ij}^t w_j^{t-1} + a_i^t\right) \tag{4}$$

where w_i^t is the i-th yield of the t layer neuron; $f(\cdot)$ is the neuron activation operation *tanh*, and m_{t-1} is the amount of neurons in the $t-1$ level. v_{ij}^t is the junction weight among the j-th neuron in level $t-1$ and the i neuron in level t; w_j^{t-1} is the yield of the j-th neuron in level $t-1$; a_i^t is the bias of the i-th neuron in the t level.

Output layer O_i is calculated as follows:

$$O_i = Q\left(\sum_{j=1}^{m_t} v_{ji}^o w_j^t + a_i^o\right) = \sum_{j=1}^{m_t} v_{ji}^o w_j^t + a_i^o \tag{5}$$

where O_i is the i output of the yield level; $Q(\cdot)$ is the Purelin operation of neuron activation; m_t is the amount of hidden layer neurons in the t level; v_{ji}^o is the junction weight among the t level obscured level neurons; a_i^o is the bias of the i-th neuron in the output layer.

The mean square error among the predicted values is as follows:

$$E(v, a, s) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{l}\sum_{j=1}^{l}(O_{ij} - C_{ij})^2 \tag{6}$$

where s is the amount of training times, l is the amount of neurons in the yield level, n is the samples' amount, $\hat{O}_{ij}$ is the j-th yield value predicted by the i-th sample, and C_{ij} is the j-th expected value of the i-th sample.

The Gaussian function [28] is then introduced into the BP neural network gradient descent method, and the model converges faster when the data size is larger.

$$G\left(\frac{\partial E}{\partial v_{ij}^t}\right) = e^{-\left(\frac{\partial E}{\partial v_{ij}^t}\right)^2/2} \tag{7}$$

$$G\left(\frac{\partial E}{\partial a_i^t}\right) = e^{-\left(\frac{\partial E}{\partial a_i^t}\right)^2/2} \tag{8}$$

where $G\left(\frac{\partial E}{\partial v_{ij}^t}\right)$ and $G\left(\frac{\partial E}{\partial a_i^t}\right)$ are the learning rates of layer t, v_{ij} and a_i calculated by the Gaussian function, $\frac{\partial E}{\partial v_{ij}^t}$ and $\frac{\partial E}{\partial a_i^t}$ are the gradients of layer t v_{ij} and a_i, respectively.

The weights and threshold learning rate matrices for each layer of the improved BP neural network are obtained as follows.

$$kr\left(v^t\right) = \begin{bmatrix} G\left(\frac{\partial E}{\partial v_{11}^t}\right) & \cdots & G\left(\frac{\partial E}{\partial v_{lm_{t-1}}^t}\right) \\ \vdots & \ddots & \vdots \\ G\left(\frac{\partial E}{\partial v_{m_t 1}^t}\right) & \cdots & G\left(\frac{\partial E}{\partial v_{m_t m_{t-1}}^t}\right) \end{bmatrix} \tag{9}$$

$$kr(a^t) = [G(\frac{\partial E}{\partial a_1^t})\cdots G(\frac{\partial E}{\partial a_{m_t}^t})] \tag{10}$$

where $kr(v^t)$, $kr(a^t)$ are the learning rate matrices of the weights and thresholds of the t-th layer, respectively; m_t, m_{t-1} are the neurons' amount in the t-th layer and the $(t-1)$-th layer, individually.

The improved gradient descent algorithm is as follows.

$$v^{s+1} = v^s - kr(v^s) \odot \left(\frac{\partial E}{\partial v^s}\right) \tag{11}$$

$$a^{s+1} = a^s - kr(a^s) \odot \left(\frac{\partial E}{\partial a^s}\right) \tag{12}$$

where v^{s+1}, a^{s+1}, v^s, and a^s are the values of v and a at $s+1$, s iterations respectively; $kr(v^s)$ and $kr(a^s)$ are the learning rate matrices of v and a calculated for the s training respectively; $\frac{\partial E}{\partial v^s}$ and $\frac{\partial E}{\partial a^s}$ are the gradient matrices of v and a calculated at the s iteration respectively.

4. Consumer expenditure forecasting based on SPSS mathematical statistics and neural networks.

4.1. Analysis of consumption influencing factors based on Lasso's method of mathematical statistics. On the ground of the comprehensive consideration of multiple factors affecting consumption, this article adopts Lasso mathematical statistics method to construct the usance expenditure models of Chinese residents respectively, and use SPSS software to empirically analyze the relevant statistical data from 2000-2015 in the experimental session.

Secondly, on the basis of variable selection, the prediction model on the ground of Lasso mathematical statistics method and improved BP neural network is constructed, and the usance expenditure of Chinese residents is predicted, and the specific process is indicated in Figure 2.

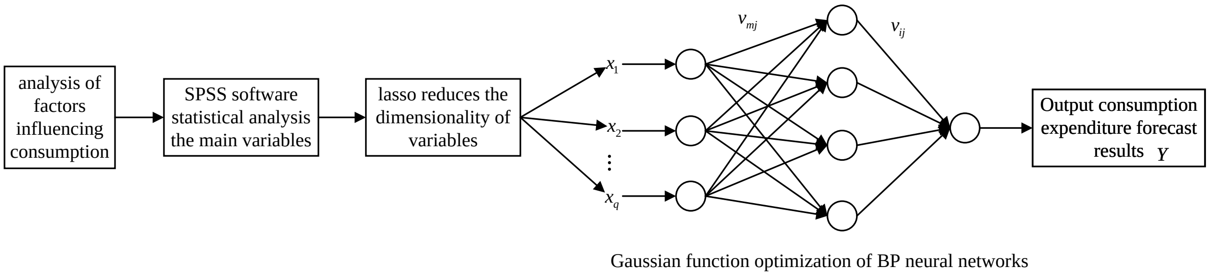

Figure 2. Consumer expenditure forecasting process based on SPSS mathematical statistics and neural networks

In terms of existing studies [29], it can be concluded that the selection of variables is based on three aspects: economic factors, social factors and other influences, and the specific selection of variables and their meanings are indicated in Table 1. Var stands for variable, IV stands for implicit variable, EF stands for economic factor, SF stands for social variable, OI represents other influences, and Y represents Consumption expenditure level (Cel).

X_1 represents Resident Disposable income (Rdi), X_2 represents Economic Growth rate (Egr), X_3 represents Inflation level (Il), X_4 stands for Gross industrial output (Gio), X_5 stands for Investment in fixed assets (Iifa), X_6 stands for Interest rates (Ir), X_7 stands for Size of population (Sop), X_8 stands for Educational situation (Es), X_9 for Employed persons (Ep), X_{10} for Income distribution gap (Idg), X_{11} for Consumer behavior (Cb), X_{12} for Miles of roads (Mor), and X_{13} for Postal and telecommunication traffic (Patt).

Table 1. Selection and definition of var

Basic var	**Var name**	**Var definition**	**Basic var**	**Var name**	**Var definition**
IV	Y	Cel		X_7	Sop
	X_1	Rdi		X_8	Es
	X_2	Egr	*SF*	X_9	Ep
	X_3	Il		X_{10}	Idg
EF	X_4	Gio		X_{11}	Cb
	X_5	Iifa	*OI*	X_{12}	Mor
	X_6	Ir		X_{13}	Patt

Suppose there are q ($q = 13$) independent variables $X_1, X_2, ..., X_q$ and dependent variable Y, and they are linearly modeled as follows.

$$Y = b + \alpha_1 X_1 + \alpha_2 X_2 + \cdots + \alpha_q X_q + \xi \tag{13}$$

where b is a constant term, $\alpha_1, \alpha_2, ..., \alpha_q$ denotes the regression coefficient, and ξ is a random perturbation term.

Let $(X_{i1}, X_{i2}, ..., X_{iq}; Y_i), i = 1, 2, ..., m$ be m sets of observations of the variables, and assume that the data have been centered and standardized, i.e., $\sum_{i=1}^{m} Y_i = 0$, $\sum_{i=1}^{m} X_{ij} = 0$, $\sum_{i=1}^{m} X_{ij}^2 = 1$.

$j = 1, 2, \ldots, q$. Noting $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_q)^T$, the Lasso estimate of the unknown parameters b and α is defined as follows.

$$(\hat{b}, \hat{\alpha}) = \arg\min \sum_{i=1}^{m} (Y_i - b - \sum_{j=1}^{q} \alpha_j X_{ij})^2 \tag{14}$$

where $\sum_{j=1}^{q} |\alpha_j| \leq s$ and s are penalty parameters.

Each value of $s \geq 0$ corresponds to a Lasso solution through Equation (14), and all Lasso solutions under different s values can be obtained after multiple calculations.

Then the parameter s is estimated by AIC criterion. The error of the model follows an independent normal distribution, so that m is the amount of observations and RSS represents the sum of remaining squares. The estimation of s is as follows.

$$AIC_s = 2L + n \ln(RSS/K) \tag{15}$$

where L denotes the amount of parameters and K is the likelihood function.

4.2. **Consumer expenditure forecasting based on SPSS mathematical statistics and neural networks.** To realize the accurate prediction of consumption expenditure, the relevant variables of the influencing factors are first analyzed through mathematical and statistical methods, the linear model is constructed through the Lasso method, and the relevant variables are downgraded, and then the downgraded influencing factors are taken as the inputs, and the consumption expenditure is taken as the outputs, and the improved BP neural network is utilized to fit the non-linear relationship between them, and to inhibit the effect of the noise on the fitting accuracy, so as to optimize the precision of the forecasting.

After the above mathematical and statistical analysis, a linear model of consumption expenditure is established, and the relevant variables are downscaled, and to estimate the impact of the scale, some variables need to be logarithmically treated. Variables X_2,

X_6, and X_{10} are not treated, and the rest of the variables are logarithmically treated and denoted as $\ln Y$, $\ln X_1$, $\ln X_3$, ..., $\ln X_{13}$. The linear model is indicated below.

$$\ln Y = \alpha_0+\alpha_1 \ln X_1+\alpha_2 X_2+\alpha_3 \ln X_3+\cdots+\alpha_6 X_6+\alpha_7 \ln X_7+\cdots+\alpha_{10} X_{10}+\cdots+\alpha_{13} \ln X_{13}+\xi \tag{16}$$

where α_0 is a constant term, $\alpha_1, ..., \alpha_{13}$ is the coefficient corresponding to each variable, and ξ is a random perturbation term.

Using the above dimensionality reduced variable $x_1, ..., x_{13}$ as the input to the BP neural network, the stimulus level output and the obscured level input and output are represented as follows.

$$\begin{cases} O_i^{(a)}(l) = x_{(i)}(l) \\ layer_j^{(b)}(l) = \sum_{i=1}^{N} v_{ij}^{(b)} O_i^{(a)}(l) \\ O_j^{(b)}(l) = f(layer_i^{(b)}(l) - \rho^{(b)}) \end{cases} \tag{17}$$

where the superscript (a) and (b) stand for the stimulus level and the obscured level of the neural network individually; $\rho^{(b)}$ stands for the threshold of the obscured level; $v_{ij}^{(b)}$ denotes the junction weight among the j-th neuron and the i-th neuron in the implicit level; $f(x)$ is the activation function $f(x) = \tanh(x) = \frac{e^x-e^{-x}}{e^x+e^{-x}}$ of the obscured level.

Meanwhile, the stimulus and yield of the output level are as follows.

$$\begin{cases} layer_1^{(c)}(l) = \sum_{i=1}^{Q} v_{li}^{(c)} O_i^{(b)}(l) \\ Q_1^{(c)}(k) = g(layer_1^{(c)}(l) - \alpha_1^{(c)}) \end{cases} \tag{18}$$

where superscript (c) represents the output layer; $\alpha_1^{(c)}$ is the threshold of l neurons in the output level; $v_{li}^{(c)}$ is the connection weight among the i-th neuron in the hidden level and the l-th neuron in the output level; $g(x)$ is the activation operation of the output layer.

The above is a forward propagation process, when the error is not 0, the error mean square value of the reverse error adjustment is as follows.

$$E(v, a, l) = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{1} (Ys(l) - Y(l))^2 \tag{19}$$

where l is the variable and connection weights $v_{ij}^{(b)}$ and $v_{li}^{(c)}$ are adjusted according to the negative gradient rule, and n is the samples' amount. $\hat{O}_{ij}$ is the j-th output value predicted by the i-th sample.

The convergence speed is then increased and the probability of falling into a dead loop is reduced by introducing an inertia term in the Gaussian function. The formula for updating the junction weights of neurons in the output and obscured layers is as follows.

$$\begin{cases} v_{li}^{(c)}(l+1) = \Delta v_{li}^{(c)}(l) - \mu_1 \frac{\partial E(l)}{\partial \alpha_{li}^{(c)}} \beta_1 kr(l-1) \\ v_{ij}^{(b)}(l+1) = \Delta v_{ij}^{(b)}(l) - \mu_2 \frac{\partial E(l)}{\partial v_{ij}^{(b)}} \beta_2 kr(l-1) \end{cases} \tag{20}$$

where β is the inertia term coefficient.

The training is carried out in terms of the improved BP neural network, and the weights and thresholds of the obscured level and output level are constantly updated. When the difference between the mean square error $E(v, a, l)$ obtained by the s computation and $E(v, a, l+1)$ obtained by the $l+1$ computation is less than the training accuracy A, that is, $E(v, a, l) - E(v, a, l+1) < A$, the training is ended, and the opposite is continued.

5. **Performance test and analysis.**

5.1. **Empirical analysis of urban residents' consumption expenditure based on SPSS software.** To estimate the performance of the consumption expenditure forecasting method based on SPSS mathematical statistics and neural network, 16 pieces of network data from China Statistical Yearbook from 2000 to 2015 were selected. The dependent variable in the data set was urban consumption expenditure, and the independent variable was 13 indicators selected according to the qualitative analysis above. After sorting out the statistical results, SPSS software was used for mathematical statistical analysis. To facilitate description, the forecasting model designed in this article is denoted as CEFB, the model proposed in literature [22] is denoted as ECPU, and the model proposed in literature [23] is denoted as IEED, and the prediction outcome of these three methods are compared and analyzed. Since the data samples from 2011 to 2015 are used as the forecast samples in this paper, the data from 2000 to 2010 are used as the data for Lasso statistical analysis. The summarized EXCEL table is imported into SPSS software, and the Lars package in SPSS software is called to perform statistical analysis on the linear parts of CEFB, ECPU and IEED models Lasso, GM and PCA respectively. The results of the linear parts of the three models for predicting the consumption expenditure of Chinese residents from 2011 to 2015 (unit: billion yuan) are indicated in Table 2, and the relative errors are implied in Figure 3.

Table 2. Predictions of consumption expenditures from the linear part of the three models

Vintages	Actual value	Forecast of ECPU-GM	Forecast of IEED-PCA	Forecast of CEFB-Lasso
2011	103964.15	103115.09	104559.26	103805.69
2012	119138.42	119526.74	119324.68	119192.18
2013	136982.37	136537.16	136729.62	137025.37
2014	150693.79	151363.58	150364.56	150738.94
2015	168237.34	168967.29	168715.27	168295.65

From Table 2, when using the GM method to forecast, except for 2011 and 2015 when the forecast value of residents' consumption expenditure is lower than the true value, the forecast values of 2012 to 2014 are all lower than the actual value; when using the PCA method to forecast, except for 2011 and 2013 when the forecast value is lower than the actual value, the forecast values of 2012, 2014 and 2015 are all higher than the actual value; when using the Lasso method to forecast, except for 2011 when the forecast value of consumption expenditure is lower than the actual value, the forecast values of 2012 to 2015 are all higher than the actual value. Thus, the Lasso linear mathematical-statistical prediction method is the best performing of the three methods.

As can be seen from Figure 3, the relative errors in these five years are lower than those predicted by GM and PCA, which indicates that the prediction effect of Lasso is relatively better. In conclusion, through the modeling of the linear part of GM, PCA and Lasso and the forecasting results of urban residents' consumption expenditure from 2011 to 2015, the forecasting effect of Lasso method on urban residents' consumption expenditure is better, closest to the real value, and the relative error is also the smallest.

Because in the actual prediction, there may be a variety of complex factors will have an impact on the prediction object, the use of a single mathematical and statistical methods for its prediction may not be satisfactory results, so the prediction results of the three combinations of ECPU, IEED and CEFB models were analyzed, and the summary of the relative error results are shown in Table 3.

As can be seen from Table 3, the relative error of the combined model CEFB based on Lasso and improved BP neural network is small, with relative errors of 0.0583, 0.0641, 0.0536, 0.0467 and 0.0516 from 2011 to 2015, respectively, and that of the ECPU model,

Table 3. Relative errors of different combination models

Vintages	**2011**	**2012**	**2013**	**2014**	**2015**
ECPU	0.2073	0.3068	0.3489	0.2862	0.3514
IEED	0.1459	0.2172	0.2751	0.2035	0.2159
CEFB	0.0583	0.0641	0.0536	0.0467	0.0516

a combined model based on GM and BP neural network, is larger. This is due to the fact that there is no dimensionality reduction of the statistical data and the output of the prediction results using only the traditional BP neural network, which leads to a large error. The relative error of IEED model, a combination model based on PCA and BP neural network, is larger than that of CEFB model. Although the variables affecting the consumption expenditure are downscaled before prediction, the traditional BP neural network is not optimized, which leads to a higher relative error than that of CEFB, thus the combined forecasting approach of the CEFB model produces better forecasts of consumption expenditures than the ECPU and IEED models.

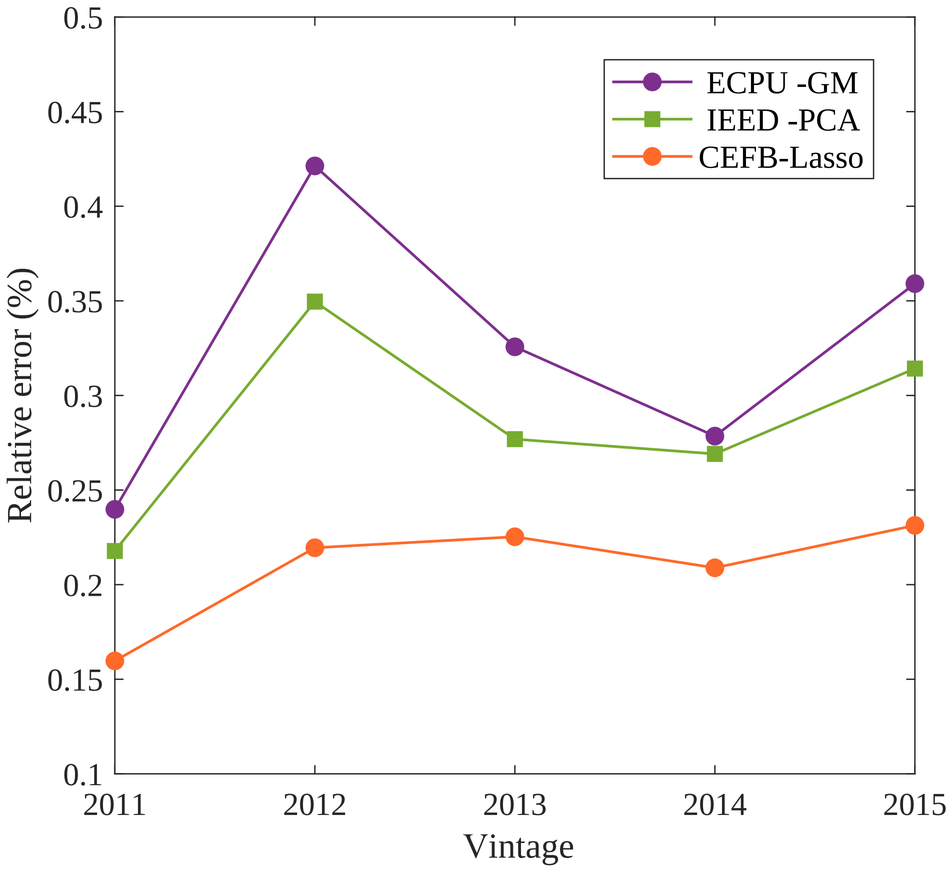

Figure 3. Comparison of relative errors for linear models with different methods

5.2. **Comparison and analysis of prediction accuracy of different models.** To evaluate the prediction accuracy of the models, ECPU, IEED and CEFB models were adopted to train and predict the consumption expenditure data respectively, and the experimental outcome was measured by five indexes: correlation coefficient (R), mean absolute error (MAE), mean absolute percentage error (MAPE), root mean squared error (RMSE) and root mean square percentage error (RMSPE). The outcome of the three model evaluation metrics is given in Table 4 and the results are plotted on a visual bar comparison chart as shown in Figure 4.

Table 4. Comparison of model prediction performance

Model	R	MAE	MAPE	RMSE	RMSPE
ECPU	0.8069	0.2489	0.2793	0.3205	0.2615
IEED	0.8713	0.1329	0.1269	0.1824	0.1226
CEFB	0.9539	0.0598	0.0427	0.0746	0.0539

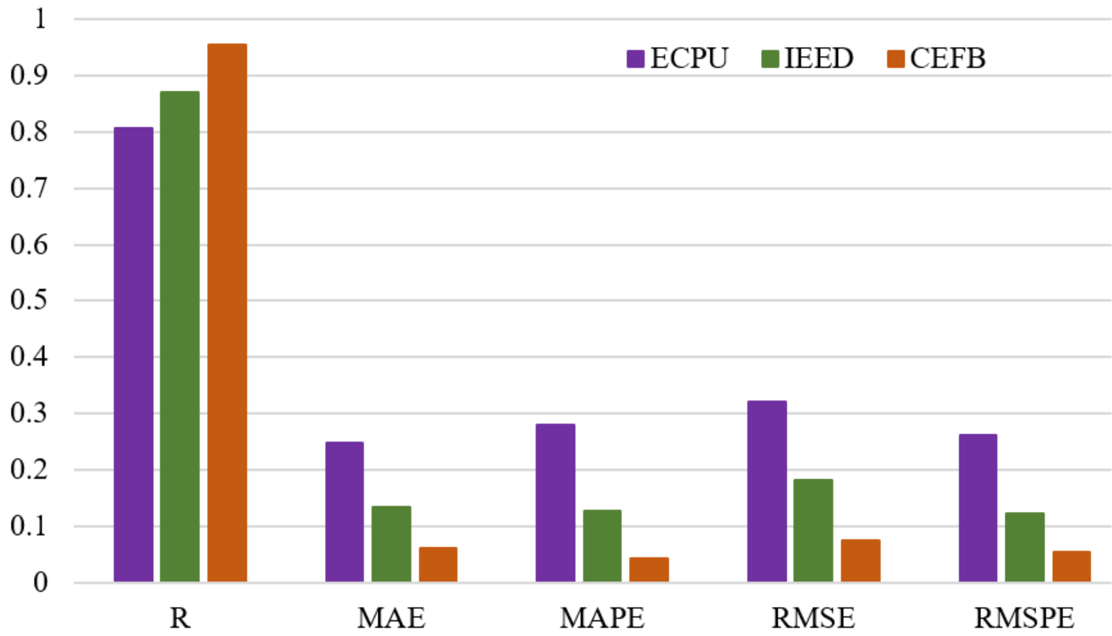

Figure 4. Prediction accuracy comparison

As can be seen from Table 4, various accuracy evaluation indexes of CEFB model are significantly superior to ECPU model and IEED model. The RMSPE of CEFB is 0.0539, which is lower than that of ECPU and IEED by 0.0687 and 0.2076 respectively. MAE is 0.0598, which is 0.0731 and 0.1891 lower than ECPU and IEED, respectively. Compared with ECPU and IEED, MAPE was 0.0427, which decreased by 0.0842 and 0.2366, respectively. RMSE was 0.0746, which decreased by 0.0731 and 0.1891 compared with ECPU and IEED, respectively. In addition, the MAE and RMSE of the CEFB are both lower than 0.15 in the prediction results, which indicates that the CEFB is more suitable for solving the prediction issue of consumption expenditure.

Comparing the correlation coefficient R, it can be seen from Figure 4 that the R value of the CEFB model is 0.9539, which is improved by 9.48% and 18.21% compared with the ECPU model and the IEED model, respectively, which indicates that the use of the Gaussian function to improve the convergence speed of the BP neural network is effective, and that the CEFB model, which introduces the inertia terms of forward and reverse transmission, is superior to the comparative models, and has a better fitting effect on the consumption and expenditure prediction. Therefore, the CEFB model has better fitting effect and prediction accuracy because it reduces the complexity of the input data by dimensionality reduction of the influencing factor variables and optimizes the weight updating algorithm of the BP neural network.

6. **Conclusion.** Aiming at the issues of low convergence speed and the existence of local minima in the error of the existing consumer expenditure prediction methods, a consumer expenditure prediction method on the ground of SPSS mathematical statistics and neural network is designed. Firstly, increase the studying rate of gradient descent algorithm through Gaussian function, control the lstudying rate of gradient descent algorithm when the error is large, and accelerate the speed of convergence of the BP neural network model, and secondly, the Lasso mathematical and statistical method is used to carry out logarithmic downgrading of the variables to reduce the complexity of the input variables, and then the improved BP neural network is adopted to fit the variables between the SPSS mathematical statistics and the neural network, so as to improve the accuracy of prediction. Finally, SPSS software is used to empirically analyze the relevant statistical data, and the method suggested in this article can effectively enhance the correlation coefficient R of the consumption expenditure prediction method, as well as reduce the MAE, MAPE, RMSE, and RMPSE, which are better applied to the forecasting of consumer expenditures.

REFERENCES

[1] W. Zhang, and X. Zhang, "A forecast analysis on fertilizers consumption worldwide," Environmental Monitoring and Assessment, vol. 133, pp. 427-434, 2007.

[2] M. Q. Dao, "Exports, imports, government consumption and economic growth in upper-middle income countries," Progress in Development Studies, vol. 14, no. 2, pp. 197-204, 2014.

[3] Y. Truong, "Personal aspirations and the consumption of luxury goods," International Journal of Market Research, vol. 52, no. 5, pp. 655-673, 2010.

[4] C. Siepmann, L. C. Holthoff, and P. Kowalczuk, "Conspicuous consumption of luxury experiences: an experimental investigation of status perceptions on social media," Journal of Product & Brand Management, vol. 31, no. 3, pp. 454-468, 2022.

[5] L. Yu, Y. Zhao, L. Tang, and Z. Yang, "Online big data-driven oil consumption forecasting with Google trends," International Journal of Forecasting, vol. 35, no. 1, pp. 213-223, 2019.

[6] S.-P. Jun, and D.-H. Park, "Consumer information search behavior and purchasing decisions: Empirical evidence from Korea," Technological Forecasting and Social Change, vol. 107, pp. 97-111, 2016.

[7] O. Schaer, N. Kourentzes, and R. Fildes, "Demand forecasting with user-generated online information," International Journal of Forecasting, vol. 35, no. 1, pp. 197-212, 2019.

[8] L. Melkumova, and S. Y. Shatskikh, "Comparing Ridge and LASSO estimators for data analysis," Procedia Engineering, vol. 201, pp. 746-755, 2017.

[9] M. Greenacre, P. J. Groenen, T. Hastie, A. I. d'Enza, A. Markos, and E. Tuzhilina, "Principal component analysis," Nature Reviews Methods Primers, vol. 2, no. 1, 100, 2022.

[10] G. Liu, and J. Yu, "Gray correlation analysis and prediction models of living refuse generation in Shanghai city," Waste Management, vol. 27, no. 3, pp. 345-351, 2007.

[11] M. Ogura, "Forecasting consumption expenditure using a dynamic panel model with cross-sectional dependence: the case of Japan," SN Business & Economics, vol. 2, no. 9, 140, 2022.

[12] B. C. Ervural, O. F. Beyca, and S. Zaim, "Model estimation of ARMA using genetic algorithms: A case study of forecasting natural gas consumption," Procedia-Social and Behavioral Sciences, vol. 235, pp. 537-545, 2016.

[13] G. Kemalbay, and O. B. Korkmazoglu, "Effects of multicollinearity on electricity consumption forecasting using partial least squares regression," Procedia-Social and Behavioral Sciences, vol. 62, pp. 1150-1154, 2012.

[14] K. Li, and T. Zhang, "Forecasting electricity consumption using an improved grey prediction model," Information, vol. 9, no. 8, 204, 2018.

[15] S. Mishra, D. Mishra, and G. H. Santra, "Adaptive boosting of weak regressors for forecasting of crop production considering climatic variability: An empirical assessment," Journal of King Saud University-Computer and Information Sciences, vol. 32, no. 8, pp. 949-964, 2020.

[16] Y. Chen, C. Deng, W. Yao, N. Liang, P. Xia, P. Cao, Y. Dong, Y.-a. Zhang, Z. Liu, and D. Li, "Impacts of stochastic forecast errors of renewable energy generation and load demands on microgrid operation," Renewable Energy, vol. 133, pp. 442-461, 2019.

[17] A. Javanainen, V. Ferlet-Cavrois, J. Jaatinen, H. Kettunen, M. Muschitiello, F. Pintacuda, M. Rossi, J. R. Schwank, M. R. Shaneyfelt, and A. Virtanen, "Semi-empirical model for SEGR prediction," IEEE Transactions on Nuclear Science, vol. 60, no. 4, pp. 2660-2665, 2013.
[18] J. Fischer, C. E. Mayer, and J. Söding, "Prediction of protein functional residues from sequence by probability density estimation," Bioinformatics, vol. 24, no. 5, pp. 613-620, 2008.
[19] Y. Liu, H. Chen, L. Zhang, X. Wu, and X.-j. Wang, "Energy consumption prediction and diagnosis of public buildings based on support vector machine learning: A case study in China," Journal of Cleaner Production, vol. 272, 122542, 2020.
[20] M. R. Biswas, M. D. Robinson, and N. Fumo, "Prediction of residential building energy consumption: A neural network approach," Energy, vol. 117, pp. 84-92, 2016.
[21] P. Li, B. Zhang, Y. Weng, and R. Rajagopal, "A sparse linear model and significance test for individual consumption prediction," IEEE Transactions on Power Systems, vol. 32, no. 6, pp. 4489-4500, 2017.
[22] Y.-C. Hu, "Electricity consumption prediction using a neural-network-based grey forecasting approach," Journal of the Operational Research Society, vol. 68, pp. 1259-1264, 2017.
[23] A. Kheirkhah, A. Azadeh, M. Saberi, A. Azaron, and H. Shakouri, "Improved estimation of electricity demand function by using artificial neural network, principal component analysis and data envelopment analysis," Computers & Industrial Engineering, vol. 64, no. 1, pp. 425-441, 2013.
[24] X. Yin, and J. Li, "RETRACTED: Financial market trend analysis based on autoregressive conditional heteroscedasticity model and BP neural network prediction," International Journal of Electrical Engineering & Education, vol. 60, no. 1, pp. 3716-3729, 2023.
[25] T.-Y. Wu, H. Li, S. Kumari, and C.-M. Chen, "A Spectral Convolutional Neural Network Model Based on Adaptive Fick's Law for Hyperspectral Image Classification," Computers, Materials & Continua, vol. 79, no. 1, pp. 19-46, 2024.
[26] F. Zhang, T.-Y. Wu, Y. Wang, R. Xiong, G. Ding, P. Mei, and L. Liu, "Application of Quantum Genetic Optimization of LVQ Neural Network in Smart City Traffic Network Prediction," IEEE Access, vol. 8, pp. 104555-104564, 2020.
[27] T.-Y. Wu, L. Wang, and C.-M. Chen, "Enhancing the Security: A Lightweight Authentication and Key Agreement Protocol for Smart Medical Services in the IoHT," Mathematics, vol. 11, no. 17, 3701, 2023.
[28] C.-M. Chen, Q. Miao, F. Khan, G. Srivastava, and S. Kumari, "Sustainable Secure Communication in Consumer-Centric Electric Vehicle Charging in Industry 5.0 Environments," IEEE Transactions on Consumer Electronics, vol. 70, no. 1, pp. 1544-1555, 2024.
[29] Y. Luo, "Resident consumption expenditure forecast based on embedded system and machine learning," Microprocessors and Microsystems, vol. 83, 103983, 2021.