International Journal of Social Sciences and Public Administration
ISSN: 3005-9836 (Print), ISSN: 3005-9585 (Online) | Volume 5, Number 2, Year 2024
DOI: https://doi.org/10.62051/ijsspa.v5n2.24
Journal homepage: https://wepub.org/index.php/IJSSPA/index

Evaluation and Predictive Analysis of the Development of the First Industry in Fujian Province under the Background of Rural Revitalization Strategy

Yuxin Zhao*

Future Technology Institute, Fujian Agriculture and Forestry University, Fuzhou, Fujian, 350002, China

*Corresponding author's e-mail: zhaoyuxin20040103@163.com

ABSTRACT

As the number of factors impacting the primary industry grows, interdisciplinary research intensifies, and cross-species studies multiply, the economic implications conveyed by various indicators and their fluctuating patterns over the years become intricate. Randomly examining individual or a subset of these related indicators inevitably lacks systemic rigor, is susceptible to randomness, and fails to provide a comprehensive and visual representation of the overall changes across industries. It is inadequate for accurately forecasting the economic outlook of the primary sector. Presently, Principal Component Analysis (PCA) has seen ample application in agriculture, particularly in fruit and vegetable maturation harvesting and preservation techniques [1-4].Given the current theoretical frameworks, we are unable to differentiate among the 23 influencing factors in agriculture, fisheries, animal husbandry, and forestry – including per capita GDP, workforce numbers, forest nurturing areas, and expenditures on agricultural, forestry, and aquatic affairs – to identify those with statistically significant relevance to estimating the economic status of the primary industry. This research aims to synthesize and investigate recent data on these 23 factors, conducting correlation analyses. Leveraging the scientific methodology of PCA, we will derive representative factors indicative of the economic level of the primary industry. Further, by employing ARMA autoregressive models to predict the future five-year trends of these representative factors, we will forecast the economic prospects of the primary industry in Fujian Province. Consequently, this will facilitate the provision of well-timed and rational recommendations for adjusting the current development strategies of the primary industry.

KEYWORDS

Primary Industry; Correlation Analysis; Principal Component Analysis; ARMA (Auto Regressive Moving Average) Model; Predictive Analysis.

1. INTRODUCTION

In addition, efforts will be made to improve public services and social welfare in rural areas, including the development of education, culture, medical and healthcare, and infrastructure. By optimizing the institutional and mechanistic framework for agricultural and rural development, implementing targeted poverty alleviation measures, the goal is to achieve the target of basic poverty alleviation for the designated population as stipulated by the state. At the same time, it is imperative to accelerate the growth rate of farmers' income and narrow the income gap between urban and rural residents. Strengthening ecological civilization construction in rural areas and consolidating the CPC's leading position in rural areas are also crucial to maintaining social harmony and stability. These achievements and experiences have laid a solid foundation for the "three agricultural" work in the new era. Meanwhile, with the continuous progress of globalization and information technology, accurate forecasting of economic prospects has become a key reference for policymakers to formulate future development strategies. Fujian Province, as an important province on China's southeast coast, has always been the subject of widespread attention for its economic development. The first industry, as the foundation of economic development, is of vital importance to Fujian's economic development.

The predecessors explored the main influencing factors of agriculture, animal husbandry, fishery and forestry by factor analysis, SBM-Tobit model measurement efficiency, generalized impulse response function and other methods. However, there are few studies combining these methods to explore the important influencing factors and predict the development prospect of the primary industry economy in Fujian Province. This study will summarize and explore the data indicators of 23 impact factors in recent years and conduct correlation analysis on them, and obtain the representative impact factors for estimating the economic level of the primary industry through the scientific method of principal component analysis. Then the linear relationship between the added value of the primary industry and the representative influencing factors is predicted by the method of multiple regression, and then the economic development prospect of the primary industry in Fujian province is predicted, and reasonable suggestions are provided for the timely adjustment of the current development strategy of the primary industry.

2. DATA INTRODUCTION

2.1. Indicators

In 2018, the National Bureau of Statistics updated the Third Version of the Three-Industry Classification Rules to align with the newly released National Economic Industry Classification. It clarified that China's primary industry includes agriculture, forestry, animal husbandry, and fishery (excluding related professional and auxiliary activities). Currently, domestic researchers have utilized these classifications to analyze and forecast economic indicators in agriculture, forestry, animal husbandry, and fishery sectors. For instance, Xu Tinglin applied Schultz's traditional agricultural transformation theory along with factor analysis and multiple linear regression models to evaluate how per capita GDP, research personnel numbers, R&D spending as a percentage of GDP, and local state-owned enterprise agricultural technicians influence high-quality agricultural development [5] . Ren Chuantang et al. assessed efficiency and its determinants through a value chain perspective using the SBM-Tobit model; they identified key factors affecting fisheries such as workforce size in fisheries, intermediate inputs used, aquaculture area size, total processed aquatic product value, and processed aquatic product values[6]. Chen Xinglin analyzed how annual average temperature and precipitation levels alongside afforestation efforts impact forestry economy growth by employing accumulation methods along with VAR models for forest carbon sinks and generalized impulse response functions[7]. Hu Shulin investigated factors influencing economic efficiency within livestock industries utilizing matrix network DEA models among others[8]; findings revealed significant influences from output-to-worker ratios in livestock production sectors as well as capital stock proportions relative to labor force sizes in those industries-alongside educational attainment levels among rural workers-and fiscal allocations towards agriculture-related expenditures compared to overall fiscal spending including R&D investments.

Among them, the number of fishery employees = the number of professional fishery employees +0.5× the number of part-time fishery employees +0.25× the number of temporary fishery employees; Fishery intermediate input = total fishery output value - fishery added value; Forestry wood price = wood output value ÷ wood production.

2.2. Data Sources

The added value of the primary industry is derived from the Statistical Bulletin of National Economic and Social Development of Fujian Province. The agricultural data came from Fujian Provincial Statistical Yearbook and Fujian Provincial Financial Yearbook. Fishery data came from China Fishery Statistical Yearbook. The forestry data came from Fujian Climate Bulletin and China Forestry and Grassland Statistical Yearbook. Animal husbandry data came from Fujian Provincial Statistical Yearbook. A small number of missing data are obtained by regression prediction based on 1-2 years' data, and the data are scientific and reliable.

2.3. Data Analysis

Excel 2021 software was used for statistics and drawing of all initial and forecast data, Python 3.8 (-64bit) was used for regression prediction of missing data and AMRA autoregressive time series prediction. IBM SPSS Statistics 26 software was used for standardized processing, correlation analysis and principal component analysis of the data over the past years.

3. INITIAL SCREENING OF INDICATORS -- CORRELATION ANALYSIS

3.1. Correlation Analysis

Before judging whether the original variables are suitable for PCA, we need to first determine whether there is a correlation between these variables [9]. Correlation analysis aims to analyze whether two sets of data influence each other and whether they are independent of each other.

The Pearson correlation coefficient (usually denoted by r) is a key statistical measure for evaluating the degree of linear association between two variables, with a numerical range limited to -1 to 1. When r is positive, it means that there is a positive relationship between the two variables, meaning that an increase in the value of one variable is accompanied by an increase in the value of the other variable. Conversely, if r is negative, it indicates that there is an inverse relationship between the two variables, with an increase in one variable leading to a decrease in the other. Additionally, the closer the absolute value of r is to 1, the stronger the linear association between the two variables; while when r is close to 0, it indicates that there is little or only a weak linear association between the two variables.

Pearson coefficient calculation formula:

$$r = \frac{\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i-\bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i-\bar{Y})^2}} \tag{1}$$

3.2. P-value Test

When discussing the correlation between two variables, it is not enough to only focus on the size of the Pearson correlation coefficient without considering the significance level. Therefore, it is necessary to analyze the Pearson correlation coefficient and p-value together.

The p-value is the probability of obtaining the current observation or a more extreme result under the assumption that the null hypothesis is true. Compared with the F-test and t-test, the p-value method can not only be used to evaluate statistical significance, but also provide a quantitative indicator of the strength of the basis for rejecting the null hypothesis, which helps to objectively select the optimal predictive model. In other words, by using p-values, one can determine statistical significance: if the p-value is less than or equal to the pre-set significance level α, it is considered that the result has statistical significance and can pass the significance test.

In addition, the P-value test can also give the strength of the basis for rejecting the null hypothesis [10] : (1) If $p \leq 0.01$, it means that the basis for rejecting the null hypothesis is strong; (2) If $0.01 < p \leq 0.05$, it means that the basis for rejecting the null hypothesis is strong; (3) If $0.05 < p \leq 0.1$, it means that the basis for rejecting the null hypothesis is weak; (4) If $p > 0.1$, then there is no basis for rejecting the null hypothesis. It can be seen that the P-value test can objectively select the best prediction model [11].

3.3. Correlation Analysis Results of Each Index

Table 1. Correlation analysis results of each index

	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14	R15	R16	R17	R18	R19	R20	R21
R1	1																				
R2	0.959**	1																			
R3	0.927**	0.944**	1																		
R4	-0.975**	-0.987**	-0.964**	1																	
R5	0.421	0.398	0.582	-0.519	1																
R6	0.517	0.445	0.381	-0.378	-0.432	1															
R7	0.778	0.672	0.563	-0.644	-0.155	0.884*	1														
R8	0.908*	0.844*	0.82**	-0.827*	0.101	0.814*	0.909*	1													
R9	0.269	0.147	-0.045	-0.096	-0.58	0.748	0.789	0.469	1												
R10	0.3	0.21	-0.029	-0.146	-0.592	0.699	0.777	0.46	0.982**	1											
R11	0.11	0.197	0.282	-0.204	-0.341	0	0.001	0.034	-0.086	-0.143	1										
R12	0.245	0.112	-0.097	-0.106	-0.324	0.242	0.452	0.25	0.572	0.646	-0.761	1									
R13	-0.843*	-0.951**	-0.825*	0.906*	-0.27	-0.332	-0.548	-0.704	-0.094	-0.207	-0.117	-0.158	1								
R14	-0.840*	-0.835*	-0.953**	0.887*	-0.727	-0.183	-0.374	-0.692	0.266	0.257	-0.147	0.108	0.705	1							
R15	0.902*	0.946**	0.8	-0.922**	0.281	0.397	0.692	0.749	0.309	0.401	0.187	0.269	-0.946**	-0.647	1						
R16	-0.931**	-0.952**	-0.863*	0.960**	-0.477	-0.298	-0.633	-0.729	-0.193	-0.266	-0.256	-0.185	0.908*	0.751	-0.975**	1					
R17	0.725	0.792	0.919**	-0.828*	0.747	0.055	0.196	0.553	-0.434	-0.413	0.309	-0.333	-0.707	-0.959**	0.598	-0.7	1				
R18	0.808	0.762	0.919**	-0.825*	0.704	0.322	0.475	0.719	-0.091	-0.145	0.441	-0.279	-0.55	-0.908*	0.581	-0.706	0.866*	1			
R19	-0.916*	-0.985**	-0.89*	0.957**	-0.334	-0.429	-0.661	-0.791	-0.196	-0.272	-0.278	-0.091	0.967**	0.744	-0.974**	0.961**	-0.727	-0.689	1		
R20	0.888*	0.963**	0.893*	-0.927**	0.284	0.435	0.597	0.808	0.065	0.151	0.024	0.152	-0.965**	-0.807	0.886*	-0.859*	0.778	0.646	-0.938**	1	
R21	-0.6	-0.491	-0.72	0.578	-0.644	-0.317	-0.388	-0.588	0.07	0.186	-0.457	0.365	0.216	0.741	-0.284	0.435	0.679	-0.931**	0.4	-0.349	1

R1: Per capita GDP

R2: Number of research and experimental personnel

R3: The proportion of R&D expenditures in GDP

R4: Number of agricultural professionals in local state-owned enterprises and institutions

R5: Expenditures on agriculture, forestry, and water affairs

R6: Number of fishery workers

R7: Intermediate inputs

R8: Cultivated area

R9: Total volume of aquatic products processed

R10: Production of aquatic products processed products

R11: Annual average temperature

R12: Annual average precipitation

R13: Area of afforestation and forest conservation

R14: Area of forest nursery and reforestation

R15: Timber production

R16: Timber price

R17: Area affected by disasters

R18: Total output of the livestock industry

R19: Proportion of expenditures on agriculture, forestry, and water affairs to the total budgetary expenditures

R20: Investment in R&D (research and experimental development)

R21: Ratio of industrial added value to regional GDP

3.4. Conclusion of Correlation Analysis

The data shown in the table reveals that there is a notably positive relationship between per capita GDP and various factors, including aquaculture area, timber output, research and development (R&D) spending, as well as the number of personnel involved in research and experimentation. Additionally, a significant correlation exists between the proportion of internal R&D expenditure relative to GDP and overall R&D spending ($P < 0.05$). Furthermore, per capita GDP correlates positively with the count of research personnel and their relation to both internal R&D expenditure proportions in GDP and timber output. Other notable correlations include those between research staff numbers and aquaculture area or disaster-affected regions; these relationships are highly significant ($P < 0.01$).

Moreover, certain variables exhibit an extremely strong negative correlation at $P < 0.01$ level: for instance, per capita GDP's connection with timber prices; the link between research personnel counts and agricultural technicians within local state-owned enterprises; as well as connections involving afforestation areas or fiscal expenditures related to agriculture.

In conclusion, from this analysis emerged several key influencing factors deemed significantly correlated. The results indicate that sixteen specific elements-such as per capita GDP itself, counts of research personnel, internal R&D expenditure ratios against GDP figures, agricultural technician numbers in state enterprises along with fishery worker statistics-effectively encapsulate economic progress within primary industries.

4. SELECTION OF REPRESENTATIVE EVALUATION INDEX -- PRINCIPAL COMPONENT ANALYSIS

4.1. Principal Component Analysis

Principal Component Analysis (PCA) is a widely used method in multivariate statistical analysis. It employs linear transformations to reduce multiple variables into a smaller set of significant variables known as principal components, which are essentially linear combinations of the original variables. The main goal of PCA is to decrease the dimensionality of the dataset while maintaining its variance information. This process simplifies data complexity and makes it easier for analysis and interpretation[12].

The mechanism behind PCA involves applying linear transformations to the original dataset's variables to create new principal components. These components are then ordered by their contribution to variance, with the first component accounting for the highest amount of variance, followed by subsequent components in descending order of their contributions. By selecting only a few principal components that contribute most significantly to variance, one can effectively reduce the dataset's dimensionality while preserving essential variance information within it.

However, the principal component analysis data should be standardized before the principal component analysis, and the relevant formulas are as follows:

There are 6 index variables for principal component analysis (data from 2017 to 2022 for each factor), X_1, X_2,…, X_6.There are 16 evaluation objects in total (that is, the impact factor of the first industry that passes the preliminary screening). The JTH index of the ITH evaluation object is X_{ij}, and the value of each index X_{ij} is converted into a standardized index $\widetilde{X_{\iota J}}$.

$$\widetilde{X_{\iota J}} = \frac{X_{ij}-\widetilde{X_J}}{S_j} (i = 1,2,\cdots,16; j = 1,2,\cdots,6) \tag{2}$$

$$\widetilde{X_J} = \frac{1}{6}\sum_{i=1}^{6} X_{ij} \tag{3}$$

$$S_j = \frac{1}{6-1}\sum_{i=1}^{6}\left(X_{ij} - \widetilde{X_J}\right)^2 \tag{4}$$

Formula

$\widetilde{X_J}, S_j$ are respectively the sample mean and sample standard deviation of the JTH index. The corresponding variables are called standardized indicator variables.

$$\widetilde{X_\iota} = \frac{X_{ij}-\widetilde{X_J}}{S_j} (i = 1,2,\cdots,16; j = 1,2,\cdots,6) \tag{5}$$

The original sample matrix is normalized to:

$$X = \begin{bmatrix} X_{11} & \cdots & X_{116} \\ \vdots & \ddots & \vdots \\ X_{61} & \cdots & X_{616} \end{bmatrix} = (X_1, X_2, \cdots, X_n) \tag{6}$$

The covariance matrix of the standardized sample:

$$R = \begin{bmatrix} r_{11} & \cdots & r_{116} \\ \vdots & \ddots & \vdots \\ r_{61} & \cdots & r_{616} \end{bmatrix} \tag{7}$$

Among $r_{ij} = \frac{1}{16-1}\sum_{k=1}^{16}(X_{ki} - \overline{X_\iota})\left(X_{kj} - \overline{X_J}\right) = \frac{1}{16-1}\sum_{k=1}^{16} X_{ki}X_{kj}$:

$$R = \frac{\sum_{k=1}^{16}(X_{ki}-\overline{X_\iota})(X_{kj}-\overline{X_J})}{\sqrt{\sum_{k=1}^{16}(X_{ki}-\overline{X_\iota})^2 \sum_{k=1}^{16}(X_{kj}-\overline{X_J})^2}} \tag{8}$$

Calculate the eigenvalues and eigenvectors of R:

Characteristic values: $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$(R is positive semi-definite matrix, and the $tr(R) = \sum_{k=1}^{n} \lambda_k = p$)

Eigenvector:

$$a_1 = \begin{bmatrix} a_{11} \\ \vdots \\ a_{p1} \end{bmatrix}, a_2 = \begin{bmatrix} a_{12} \\ \vdots \\ a_{p2} \end{bmatrix}, a_3 = \begin{bmatrix} a_{13} \\ \vdots \\ a_{p3} \end{bmatrix} \tag{9}$$

Calculate principal component contribution rate and cumulative contribution rate:

$$principal\ component\ contribution\ rate = \frac{\lambda_i}{\sum_{k=1}^{16} \lambda_k} (i = 1,2,3,\cdots,16) \tag{10}$$

$$cumulative\ contribution\ rate = \frac{\sum_{k=1}^{i} \lambda_k}{\sum_{k=1}^{16} \lambda_k} (i = 1,2,3,\cdots,16) \tag{11}$$

Write principal component:

$$f_i = a_{1i}X_1 + a_{2i}X_2 + \cdots a_{pi}X_p (i = 1,2,3,\cdots,m) \tag{12}$$

m Number of primary components

4.2. Data Processing Results of Principal Component Analysis

Table 2. Component score coefficient matrix

	Component1	Component2	Component3
Per capita GDP	.081	.044	.026
Number of researchers and experimentalists	.081	.017	-.101
Internal expenditure on R&D as a proportion of total GDP	.080	-.110	.054
Number of agricultural professionals and technicians in local state-owned enterprises and institutions	-.081	.035	.060
Number of people employed in fisheries	.040	.376	.303
Intermediate input	.058	.340	.181
Breeding area	.073	.179	.192
Afforestation conservation area	-.073	-.044	.287
Forest tending area	-.072	.225	-.072
Timber yield	.074	.104	-.224
Timber price	-.077	-.008	.168
Affected area	.067	-.290	-.002
Gross output value of animal husbandry	.069	-.190	.260
Expenditure on agriculture, forestry and water resources as a proportion of government expenditure	-.078	-.049	.165
Research and experimental development (R&D) funding	.077	.029	-.162
Ratio of industrial added value to gross regional product	-.050	.199	-.455

Table 3. Principal component analysis total variance interpretation

Total variance interpretation						
Component	Initial Eigenvalue			Extract the sum of squared loads		
	SUM	Percent Variance	Accumulate %	SUM	Percent Variance	Accumulate %
1	12.157	75.982	75.982	12.157	75.982	75.982
2	1.883	11.767	87.749	1.883	11.767	87.749
3	1.487	9.297	97.046	1.487	9.297	97.046
4	.356	2.226	99.272			
5	.116	.728	100.000			
6	2.009E-15	1.256E-14	100.000			
7	5.461E-16	3.413E-15	100.000			
8	3.707E-16	2.317E-15	100.000			
9	2.454E-16	1.534E-15	100.000			
10	1.797E-16	1.123E-15	100.000			
11	1.201E-17	7.509E-17	100.000			
12	-6.573E-17	-4.108E-16	100.000			
13	-2.809E-16	-1.755E-15	100.000			
14	-3.284E-16	-2.053E-15	100.000			
15	-3.840E-16	-2.400E-15	100.000			
16	-5.907E-16	-3.692E-15	100.000			

4.3. Screening of Representative Evaluation Indicators

As can be seen from Table 1, the evaluation indexes with high scores in principal component 1 are, in turn, per capita GDP, number of research and experimental personnel, proportion of internal R&D expenditure to total GDP, investment in research and experimental development (R&D), and wood production. In principal component 2, the evaluation indexes with high scores were the number of fishery employees, intermediate inputs, forest tending area, aquaculture area, and the ratio of industrial added value to gross regional product. In principal component 3, the evaluation indexes with high scores were the number of fishery employees, afforestation area, total output value of animal husbandry, breeding area and intermediate input.

According to the contribution rate formula, the contribution rates of principal component 1, principal component 2 and principal component 3 are 78.29%, 12.13% and 9.58%, respectively. Take it as the weight assigned to these 16 preliminary screening evaluation indicators, and get the top 60% of the comprehensive score (i.e. 8 after rounding) evaluation indicators as the representative impact factors on the economic level of the primary industry, as follows: Number of fishery employees, intermediate inputs, farming area, per capita GDP, total output value of animal husbandry, number of research and experimental personnel, proportion of internal expenditure of R&D funds to total GDP, timber production.

5. TIME SERIES PREDICTION -- ARMA AUTOREGRESSIVE MODEL

5.1. Autoregressive Model (AR)

Autoregressive Model (AR model) is a statistical model that describes how a time series depends on its historical values, using the past values of a time series to predict its current or future values. It assumes that the observations at the current moment are a linear combination of the observations at the past moments plus a white noise error term, reflecting the persistence and dynamics within the sequence.

The formal expression of AR model is as follows:

$$Y_t = c + \emptyset_1 Y_{t-1} + \emptyset_2 Y_{t-2} + \cdots + \emptyset_p Y_{t-p} + \varepsilon_t \tag{13}$$

Where Y_t represents the observed value of a time series at a time point t, c is the constant term, $\emptyset_1, \emptyset_2, \cdots, \emptyset_p$ are autoregressive coefficients, respectively, which determine how much a past observation affects the current value, and ε_t is the error term, which is usually assumed to be an independent, identically distributed zero-mean random variable representing random fluctuations that the model fails to capture.

The activities of the primary industry are often affected by seasonal, cyclical and long-term trends, and the number of fishery employees, intermediate inputs, farming area, per capita GDP, total output value of animal husbandry, the number of research and experimental personnel, the proportion of internal expenditure of R&D funds in the total GDP, and wood production are directly affected by the past state. AR models can effectively capture the long-term dependencies in these time series, helping to identify and analyze the continuous trend of industrial development.

5.2. Moving Average Model (MA)

The Moving Average Model (MA model) is another time series analysis tool that focuses on describing the relationship between random fluctuations in current observations and past error terms. Unlike AR models, MA models do not rely directly on historical observations of the sequence, but rather on random error terms for those observations. Its basic form is:

$$Y_t = \mu + \theta_1 \cdot \varepsilon_{t-1} + \theta_2 \cdot \varepsilon_{t-2} + \cdots + \theta_q \cdot \varepsilon_{t-q} + \varepsilon_t \tag{14}$$

Here, μ is the mean of the sequence, $\theta_1, \theta_2, \cdots, \theta_q$ is the moving average coefficient, representing the weight of the contribution of the past error term to the current value, while $\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \cdots, \varepsilon_{t-q}$ is still the error term, representing the random component in the sequence.

The influencing factors of the first industry, such as intermediate input, breeding area, total output value of animal husbandry and wood production, are susceptible to random factors such as climate and natural disasters, resulting in sudden fluctuations. MA model is good at dealing with such random shocks, and can reflect the short-term volatility well by considering the influence of the recent error term on the current value.

5.3. ARMA

ARMA model, full name of Auto-Regressive Moving Average model, is a common method in time series analysis, which combines Auto-Regressive model (AR) and moving average model (MA).

Moving Average is a characteristic used to describe and predict the behavior of a time series. This model is especially suitable for time series data with autocorrelation and random fluctuation. Reasonable selection of ARMA model order can reduce calculation cost on the premise of ensuring simulation accuracy [13]. Common order criterion functions include Akaike's information theoretic criterion (AIC), Bayesian information theoretic criterion (AIC), and Bayesian Information theoretic criterion (AIC). BIC, final prediction error criterion (FPE), etc.

ARMA(p, q) model is a combination of AR model and MA model, which captures the autoregressive property and moving average property that exist simultaneously in the sequence. The model form is:

$$Y_t = c + \emptyset_1 Y_{t-1} + \emptyset_2 Y_{t-2} + \cdots + \emptyset_p Y_{t-p} + \theta_1 \cdot \varepsilon_{t-1} + \theta_2 \cdot \varepsilon_{t-2} + \cdots + \theta_q \cdot \varepsilon_{t-q} + \varepsilon_t \quad (15)$$

5.4. Prediction and Analysis of Time Series Data

In order to perform predictive analysis on time series data, we use Python programming language and its powerful data analysis library Pandas and statistical modeling library Statsmodels.

First, the data is processed by importing pandas as pd. Introduction of Statsmodels library ARIMA module (from Statsmodels. Tsa. ARIMA. Model import ARIMA), an ARIMA model is established. Based on ARMA(1,1)[14], we call ARIMA class to instantiate the model, and use.fit() method to fit the data, and we get the best estimate of the model parameters. The.forecast(steps=forecast_steps) method is used to forecast the following forecast_steps time points (set to 5 here) based on the fitted model.

Finally, the time series forecast data of the number of fishery employees, intermediate input, farming area, per capita GDP, total output value of animal husbandry, number of research and experimental personnel, the proportion of internal expenditure of R&D funds in total GDP, and wood production in the next 5 years are obtained, and the line chart is drawn as follows:

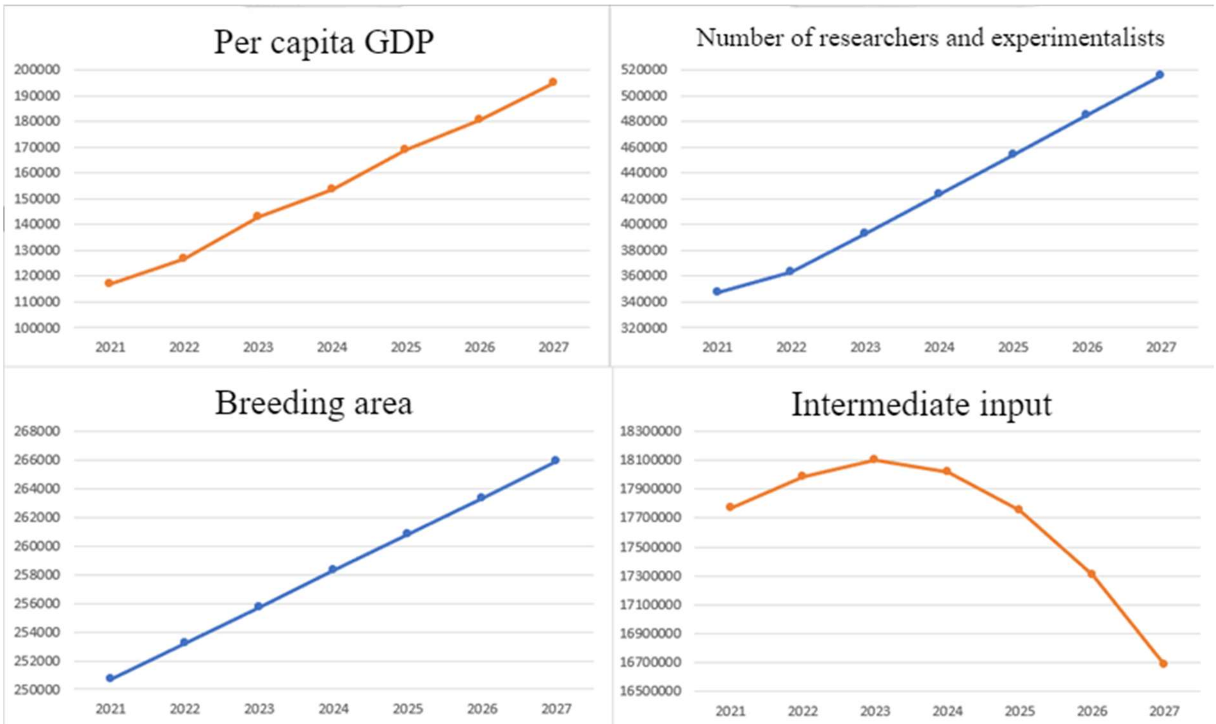

Figure 1. Represents the Development Trend of Factors①

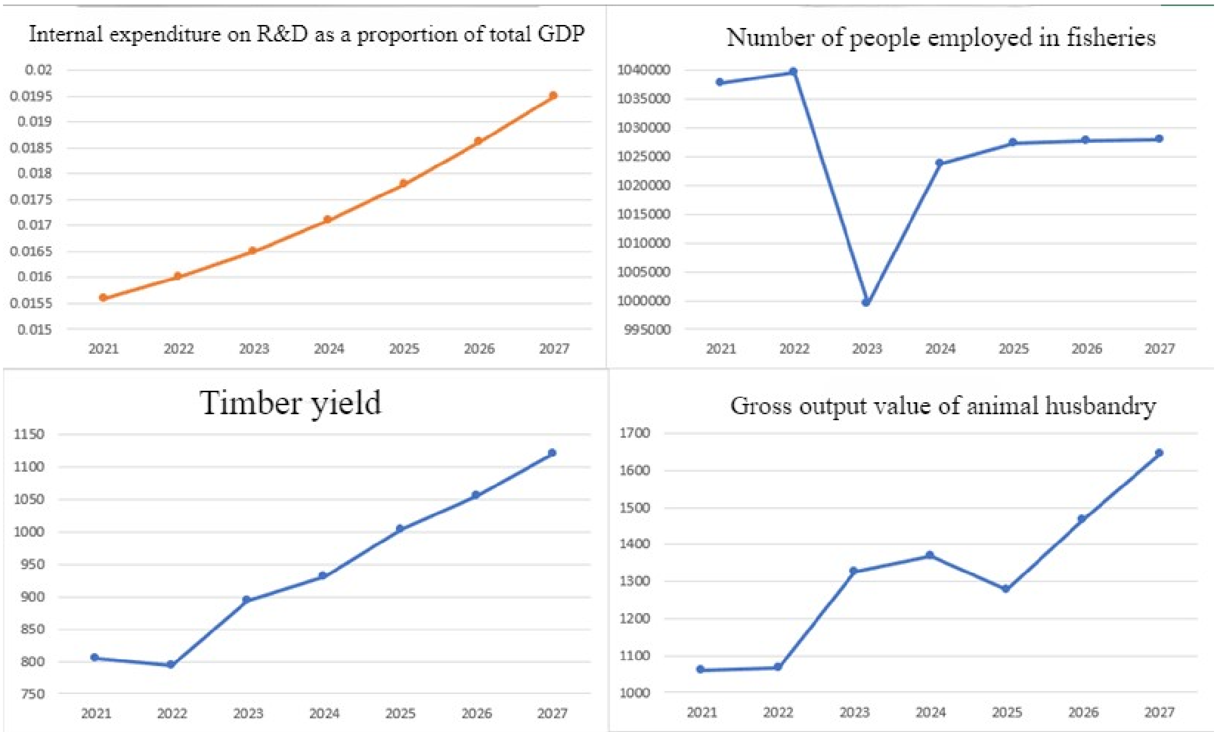

Figure 2. Represents the Development Trend of Factors②

According to the development trend line chart drawn based on the forecast data obtained by ARMA autoregressive forecasting method, the proportion of per capita GDP, the number of research and experimental personnel, aquaculture area, and internal expenditure of R&D funds in the total GDP showed an increasing trend, and the intermediate input of fisheries showed an obvious decreasing trend. The number of fishery employees is expected to decline significantly in 2023 and then rise significantly and level off, timber production will increase year by year after 2022, and the total output value of animal husbandry is expected to rise gradually after a significant inflection point in 2025.

6. CONCLUSION AND SUGGESTIONS

6.1. Conclusion

According to the ranking of the comprehensive score of the principal component analysis, the number of fishery employees, intermediate input, farming area, per capita GDP, total output value of animal husbandry, the number of research and experimental personnel, the proportion of internal expenditure of R&D funds in the total GDP, and the importance of wood production were adjusted, and the forecast development trend line chart was referred to.

As a result, the economic base of Fujian's primary industry will continue to rise steadily, the degree of emphasis on scientific research will continue to increase, and the economic growth and R&D investment will increase. The structure of the primary industry has changed, the scale of the aquaculture industry has expanded, and the fishery sector has undergone structural adjustment and industrial transformation. With the increasing market demand of forestry resources, sustainable development should be paid more attention to. The positive impact of internal changes in the animal husbandry industry and improvement of the market environment.

6.2. Suggestions

(1) Accelerate the transformation of Marine economy and fisheries. Promote the blue economy, strengthen the construction of Marine pastures, and promote the intelligent and ecological transformation of mariculture industry; Pay attention to the protection of fishery resources, and formulate strict measures for the protection of Marine resources.

(2) Technological innovation and industrial upgrading. Give full play to location advantages and strengthen cross-strait scientific and technological exchanges and cooperation; We will support the rapid development of the digital economy, build new heights of the digital economy by relying on the foundation of digital economy development in Fuzhou, Xiamen and other places, and promote the digital transformation of traditional industries.

(3) Attach importance to sustainable agricultural development. Utilizing the rich geographical and climatic conditions of Fujian, we will develop the cultivation of characteristic fruits, flowers and Chinese herbal medicine, etc., and create the development of characteristic agriculture.

(4) Ecological and green development. To strengthen the protection and utilization of forest resources, strengthen the protection of forest resources, rationally develop forest tourism and forest economy; Promote green transportation and energy, reduce carbon emissions, and promote green and low-carbon development.

(5) We will maintain balanced development among regions. Strengthen the connectivity of infrastructure in the two major collaborative development zones of northeast Fujian and southwest Fujian, and optimize the regional industrial layout; We will strengthen the rural revitalization strategy, increase investment in rural infrastructure and public services, and realize rural revitalization and integrated urban and rural development.

REFERENCES

[1] Nowicka, P., Wojdylo, A., Laskowski, P. (2019) Principal component analysis (PCA) of physicochemical compounds content in different cultivars of peach fruits, including qualification and quantification of sugars and organic acids by HPLC. European Food Research and Technology, 245(4): 929-938.

[2] Allegre, A., Gallotta, A., Carimi F., et al. (2018) Metabolic profiling and post-harvest behavior of “Dottato” fig (Ficus carica L.) fruit covered with an edible coating from O. ficus-indica. Frontiers in Plant Science, 9: 1321.

[3] Kim, T.J., Hyeon, H., Park, N.I., et al. (2020) A high-throughput platform for interpretation of metabolite profile data from pepper (Capsicum) fruits of 13 phenotypes associated with different fruit maturity state. Food chemistry, 331: 127286.

[4] Cheng, J.H., Zhou, X.Q., Liu, B.D., Sun, J., Zhang, M., Liu, Y., Cui, K.B., Zhu, X. (2024) Establishment of mango maturity evaluation index system based on principal component analysis and partial least square method. Modern food technology:1-11.

[5] Ren, C.T., Wei, S.Q., You, X.J. (2022) Comparison of fishery economic efficiency and its influencing factors from the perspective of value chain in Fujian Province. China Agricultural Resources and Regional Planning, 43(7):252-261.

[6] Chen, X.L. (2018) Study on economic value assessment and influencing factors of forestry carbon sink: A case study of Fujian Province. Thesis of Fujian Agriculture and Forestry University.

[7] Xu, T.L. (2023) Study on high quality development level and influencing factors of agricultural economy in Fujian Province. Thesis of Fujian Agriculture and Forestry University.

[8] Hu, S.L., Luo, C.W. (2022) Study on the difference and influencing factors of circular economy efficiency in three regions of animal husbandry. Journal of Southwest University of Science and Technology: Philosophy and Social Science Edition, 39(1):60-68.

[9] Fan, X., Liu, B., Cao, J., et al. (2019) Dehydrofreezing of peach: Blanching, D-sodium erythorbate vacuum infiltration, vacuum dehydration, and nitrogen packaging affect the thawed quality of peach. Journal of food biochemistry, 43(7): 12830.

[10] Wang, J., Xiao, H.D., You, Q.F., Shi, X.Z., Zhu, L.J., Meng, W. (2019) Application of P-value test to the displacement prediction of tunnel surrounding rock. Highway Engineering, 48(6):66-73.

[11] Meng, W., He, C., Yan, Q.X., et al. (2021) Application of P-value method to inversion of initial geostress field in rock mass. China Railway Science, 42(1):71-79.

[12] Jin, W.L. (2023) Research on high dimensional time series data prediction based on dynamic principal component analysis. Thesis of Shanxi University.

[13] Li, Y.G., He, Y.P., Yang, Y. (2010) Application of ARMA Model with model order in Wind speed simulation of wind power system. East China Electric Power, 38(3):395-398.

[14] Si, X., Zhu, Q.N., Liu, H.Z., et al. (2023) Prediction of syphilis incidence in Tongzhou District of Nantong City based on ARMA(1, 1) model. Journal of Nantong University (Medical Science Edition), 43(6): 557-560.