Research on Sustainable Development of Food Systems Based on Data Mining Technology

Xinwei Yu*, Meixuan Zhu

College of Information Engineering, Shenyang University, Shenyang, China, 110044

*Corresponding author: dreamer4203@163.com

Abstract. With the progress of society, the problem of people's food and clothing has been basically solved, but the unbalanced development makes the level of sustainability of the food system incompatible with the level of modern technological development. Based on SPSS, MATLAB, and PYTHON computing software and data mining techniques such as factor analysis and time series analysis, this paper constructs a food system security evaluation index system and gives the contribution rate of each index to the efficiency, effectiveness, equity, and sustainability of the food system security. A linear programming model favoring the sustainable development of the food system was constructed, and the incentive strategy for sustainable development was given according to the load of sustainable development factors. The validity and feasibility of the model were tested.

Keywords: Data Mining Techniques, the Food System Security, Equity, Sustainability, the Load of Sustainable Development Factors.

1. Introduction

With social progress, people's food and clothing problems have been basically solved, but the development is not balanced [1]. Since the current global food system management model prioritizes efficiency and profitability, although this food system management model is very effective, there is an unfair distribution of food and the sustainable development of the food system. In particular, the current COVID-19 pandemic has increased the insecurity of the global food system. In recent years, experts and scholars have conducted a lot of research on food security, and many countries and regions have also been affected by food security, which will have far-reaching significance and influence on the sustainable development of this era and the future. Current research on food security is usually to build a national or regional framework for food trading systems so that the food system model to maintain a good competitive relationship within the scalability and adaptability [2]. More researchers focus on the efficiency and profitability of food system security, while the fairness and sustainability of food system security are ignored. Based on data mining techniques, this paper discusses how to improve the sustainability of food production, improve the unfairness of food systems, and improve the security of the Earth's food system.

To improve the security of the Earth's food system, this paper discusses how to improve the sustainability of food production and the unfairness of the food system based on data mining technology.

2. Global Food System Security Analysis

2.1. Factor Analysis of the Food System

Based on food security considerations, the current state of the global food system is analyzed in terms of food production efficiency, profitability, fairness, and sustainability [3].

Based on the current status data of the food system queried by the official websites of FAO, the United Nations, the World Trade Organization, etc., combined with the data information characteristics of the food system efficiency, profitability, fairness, and sustainability, select the food system per capita GDP/daily per capita calorie intake Evaluation indicators such as the amount of agricultural employment, the ratio of agricultural employment to total employment, and per capita food supply, and perform factor analysis on the indicator data. Based on food security, the current situation of the global food system is analyzed from four aspects of food production efficiency, profitability, fairness, and sustainability. According to the FAO Food and Agriculture Organization, the WTO World Trade Organization, and other official websites of the food system from 2005 to 2020 data, combined with the food system efficiency, profitability, fairness, and sustainability of data and information characteristics, selection the food system per capita GDP / daily per capita calorie intake, the agricultural employment share of total employment, per capita food supply, gas emissions / total grain production, agricultural land area ratio, per capita arable land, agricultural capital investment and total grain exports / total grain production and other indicators 8 evaluation indicators. And factor analysis of the index data, as shown in Table 1.

Table.1. Factor Analysis Results

Component	Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	Variance%	Cumulative%	Total	Variance%	Cumulative%
1	2.313	28.911	28.911	2.303	28.782	28.782
2	1.632	20.398	49.309	1.600	20.004	48.787
3	1.225	15.308	64.617	1.160	14.503	63.289
4	1.041	13.018	77.635	1.148	14.345	77.635
5	0.938	11.727	89.361			
6	0.577	7.210	96.572			
7	0.257	3.218	99.790			
8	0.017	0.210	100.000			
Extraction Method: Principal Component Analysis.						

We adopt the principal component analysis of the correlation matrix and obtain four common factors based on the selection rule of common factors with characteristic roots greater than 1[4]. The cumulative contribution rate of 4 common factors was 77.635 %.

2.2. Construction of the global food system security evaluation index

To explain and name the physical meaning of the common factor, we rotate the factor load matrix. The rotated factor load matrix is shown in Table 2 [5].

Extract four common factors F_1, F_2, F_3, F_4, and physically name them. Based on the actual physical significance of the four common factors, name F_1 as the efficiency factor of the food system, F_2 as the profit factor, F_3 as the fairness factor, and F_4 as the sustainability factor.

The contribution rate of F_1 is 28.782 %, and the contribution rate of F_2 is 20.004 %, which indicates that it is reasonable to give priority to the safety of the food system with efficiency and profitability, but the fairness and sustainability of the food system are neglected.

This paper studies how to improve the security of the food systems by regulating the sustainability of the food systems.

Table.2. Rotated Component Matrix

	Component			
	F_1	F_2	F_3	F_4
Quantity of food export/Quantity of food output x_1	0.203	0.147	0.346	0.643
Per capita GDP/ Daily per capita calorie intake x_2	0.974	-0.056	0.135	0.011
Per capita food supply x_3	-0.220	0.891	0.144	-0.079
The proportion of land used for agriculture x_4	0.181	0.068	0.893	0.094
Agricultural investment x_5	0.074	-0.018	-0.055	-0.693
The ratio of agricultural employment to total employment x_6	0.961	-0.002	0.045	0.046
Arable land per capita x_7	-0.429	-0.227	0.512	0.445
GHG emissions/Quantity of food output x_8	0.203	0.852	-0.124	0.196

2.3. The global food system security evaluation index factor score

By factor analysis, we also get the factor score matrix, as shown in Table 3 [6].

Table.3. Component Score Coefficient Matrix

	Component			
	F_1	F_2	F_3	F_4
Quantity of food export/Quantity of food outputx_1	0.098	0.044	-0.243	0.538
Per capita GDP / Daily per capita calorie intakex_2	0.427	-0.053	0.123	0.046
Per capita food supplyx_3	-0.117	0.575	0.148	-0.108
The proportion of land used for agriculture x_4	0.080	0.074	0.774	-0.010
Agricultural investmentx_5	0.016	0.020	-0.103	-0.615
The ratio of agricultural employment to total employmentx_6	0.420	-0.023	0.049	0.066
Arable land per capitax_7	-0.205	-0.140	0.399	0.428
GHG emissions/Quantity of food outputx_8	0.073	0.519	-0.063	0.125

From Table 3, we can obtain the mathematical model of factor scores, namely

$$F_1 = -0.098x_1 + 0.427x_2 - 0.117x_3 + 0.080x_4 + 0.016x_5 + 0.420x_6 - 0.205x_7 + 0.073x_8 \quad (1)$$

$$F_2 = 0.044x_1 - 0.053x_2 + 0.575x_3 + 0.074x_4 + 0.020x_5 - 0.023x_6 - 0.140x_7 + 0.519x_8 \quad (2)$$

$$F_3 = -0.243x_1 + 0.123x_2 + 0.148x_3 + 0.774x_4 - 0.103x_5 + 0.049x_6 + 0.339x_7 - 0.063x_8 \quad (3)$$

$$F_4 = 0.538x_1 + 0.046x_2 - 0.108x_3 - 0.010x_4 - 0.615x_5 + 0.066x_6 + 0.428x_7 + 0.125x_8 \quad (4)$$

According to the above factor score matrix and model, the scores of four factors in the global food system from 2010 to 2020 are calculated, and the scores are shown in Figure 1.

Globally, the sustainable development of the food system is poor, although it has improved in recent years, it is still a negative state; the fairness of the food system is relatively good, which is also the result of the use of global average index information [7]. Of course, there should be some food insecurity areas, and it is also very serious, but in general, the average state of food security is relatively good. The efficiency of the food system has a downward trend, while the profitability of the food system is on the rise.

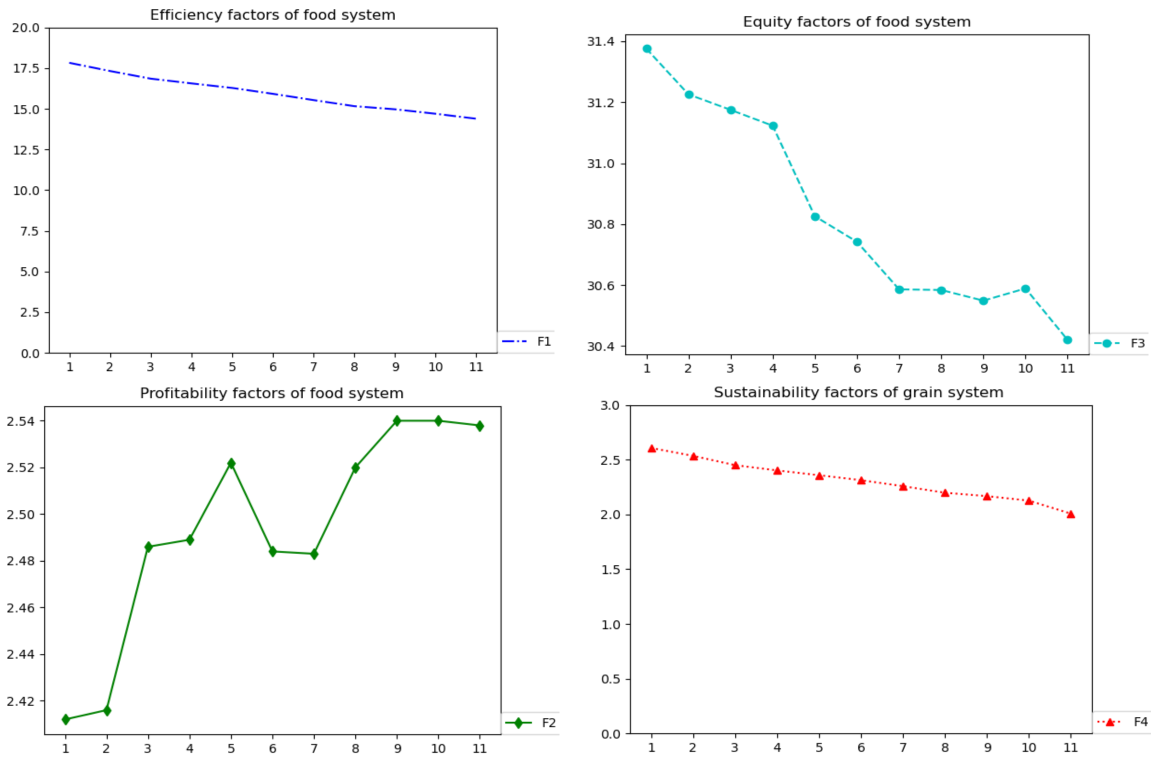

Figure 1. The global food system factors score

3. Construction of Linear Programming Model for Sustainable Development of Grain System

In order to formulate a plan to improve the rationality of the food system, a linear programming model conducive to the sustainable development of the food system is proposed with sustainability as the optimization index [8]. According to the practical significance of the indicators and the load on sustainability factors, several adjustable indicators of the food system are selected and the control strategy is determined. Constructing a linear programming model to improve the rationality of the food system and improving the level of sustainability factors under the condition of ensuring the comprehensive score of each factor is not reduced are conducive to the virtuous cycle strategy of sustainable development of food system.

Considering the actual significance of each index and the influence of the index on the common factor, since the index x_2, x_3, x_4, x_7, x_8 is not adjustable, we only consider the regulation of the index x_1, x_5, x_6. Suppose the annual adjustment ratio parameters of the three indicators are k_1, k_2 and k_3 respectively, that is, the three indicators after adjustment are k_1x_1, k_5x_5 and k_6x_6.

Estimate the control ratio parameters with the data from 2015 to 2020. We use SPSS statistical software to perform regression analysis according to its data characteristics and get the data estimation model after 2020.

To simplify the problem, it may be useful to assume that indicator x_2, x_3, x_4, x_7, x_8 remains unchanged and to consider how indicator x_1, x_5, x_6 changes to allow a higher score of the food system sustainability factor. So we get four common factor scoring models as follows.

$$F_1 = 0.098(k_1 - 1)x_1 + 0.016(k_5 - 1)x_5 + 0.420(k_6 - 1)x_6 + 12.6032 \tag{5}$$

$$F_2 = 0.044(k_1 - 1)x_1 + 0.020(k_5 - 1)x_5 - 0.023(k_6 - 1)x_6 + 2.6243 \tag{6}$$

$$F_3 = -0.243(k_1 - 1)x_1 - 0.103(k_5 - 1)x_5 + 0.049(k_6 - 1)x_6 + 29.8679 \tag{7}$$

$$F_4 = 0.538(k_1 - 1)x_1 - 0.615(k_5 - 1)x_5 + 0.066(k_6 - 1)x_6 + 1.4787 \tag{8}$$

A linear programming model is established, namely

$$\max F_4 = 0.538(k_1 - 1)x_1 - 0.615(k_5 - 1)x_5 + 0.066(k_6 - 1)x_6 + 1.4787 \tag{9}$$

$$\text{s.t.}\begin{cases} F_1 + F_2 + F_3 + F_4 \geq 12.6033 + 2.6243 + 29.8679 + 1.4787 = 46.5743 \\ F_4 \geq 1.4787 \end{cases} \tag{10}$$

Among them, are the global predicted values of the three indicators in 2020.

In order to solve the problem that the sustainability level of the food system does not continue to decline, we construct a linear programming model and find the optimal solution of k_1, k_5, k_6. Using MATLAB software, the optimal solution for linear programming is

$$k_1 = 1.04657, k_5 = 1.04647, k_6 = 1.00998 \tag{11}$$

4. Model validity and feasibility test

We use the data information of indicators $x_1, x_2, \ldots x_8$ from 2005 to 2020 to construct regression analysis and obtain the regression mathematical model as follows.

$$x_1 = \begin{cases} 0132 + 0.0003t(21 \leq t \leq 27) \\ 0.192(k_1 - 1)^{t-20}(t > 27) \end{cases} \tag{12}$$

$$x_2 = 0.251 - 0.004t \tag{13}$$

$$x_3 = 0.7773 \tag{14}$$

$$x_4 = 38.333 - 0.073t \tag{15}$$

$$x_5 = \begin{cases} -0.3688(21 \leq t \leq 27) \\ -0.3688(k_5 - 1)^{t-20}(t > 27) \end{cases} \tag{16}$$

$$x_6 = \begin{cases} 38.701 - 0.781t(21 \leq t \leq 27) \\ 23.3113(k_6 - 1)^2(t > 27) \end{cases} \tag{17}$$

$$x_7 = 0.224 - 0.002t \tag{18}$$

$$x_8 = 0.034 - 0.00044t \tag{19}$$

Data information for the period 2021 to 2027 of the 8 indicators$x_1, x_2, \ldots x_8$is estimated by the regression mathematical model [9]. Substituting the data information from 2005 to 2020 and the estimated values from 2021 to 2027 in the eight indicators into the factor score mathematical model, and obtaining the factor score from 2005 to 2027, as shown in Figure 2.

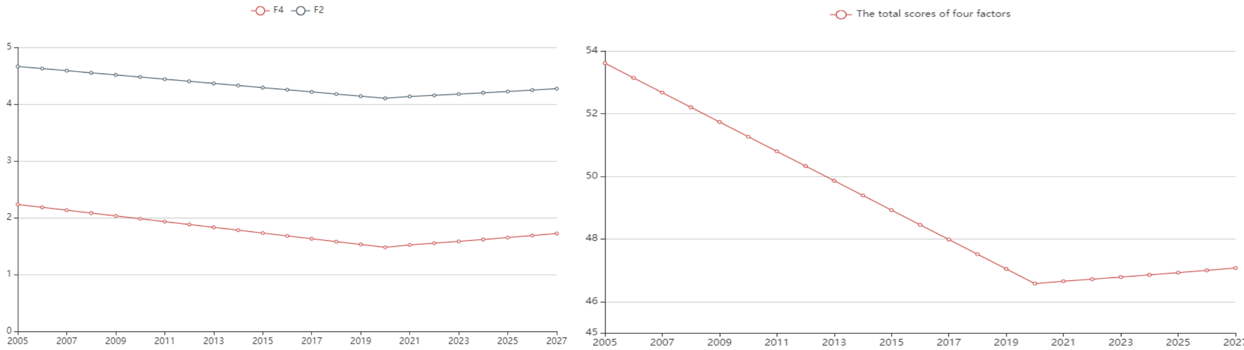

Figure 2. Forecast results from 2005 to 2027

It can be seen from Figure 2 that starting from 2021, under the premise of ensuring that the sum of the scores of the various factors of the food system does not decrease, the construction of a linear programming model to increase the level of sustainability factors is conducive to a virtuous cycle of sustainable development of the food system Effective strategy [10].

5. Conclusions

In this paper, we construct a global food system regulatory strategy based on improving the equity and sustainability of food system security by developing a factor scoring model. The strategy shows a high priority for efficiency and profitability of the current global food system, which fits with the current state of the food system. We focus on optimizing the sustainability of the current food system model based on the factor contribution ratio. The regulation strategy makes the efficiency, profitability and equity of the food system reach the world universal level while greatly improving sustainability and the overall score is improved, which facilitates decision makers to take optimization measures based on the indicators under the four dimensions. This regulation strategy model is of practical significance to change the current inequitable situation of the global food system and is conducive to the sustainable development of the global food system. It has certain theoretical and practical significance for studying the security of the global food system. Due to the limitation of data selection of countries and indicators, it makes the model a low fit for countries with very special situations, which leads to the limit of adaptability of the model cannot reach a high degree. Due to the limitation of space, it is not discussed here and is subject to further study in the future.

References

[1] Xu Tian. Research on the Construction of China's Food Security Early Warning System[D]. Harbin Institute of Technology, 2020.

[2] Zhang Qianyu, Bi Wenbo, Wang Lujin. Construction of World Food System Optimization Scheme Based on Model Analysis [J]. Scientific and Technological Innovation, 2021(13): 37-38.

[3] Irtyshcheva I, Ponomarova M, Dolzhykova I. Conceptual fundamentals of development of the food security system [J]. Baltic journal of economic studies, 2019, 5(2): 57-64.

[4] Wang R, Chen J, Li M. Coupling and Coordinating Relationship between Agricultural Eco-Efficiency and Food Security System in China [J]. International Journal of Environmental Research and Public Health, 2023, 20(1): 431.

[5] Cheng P, Tang H, Lin F, et al. Bibliometrics of the nexus between food security and carbon emissions: Hotspots and trends [J]. Environmental Science and Pollution Research, 2023, 30(10): 25981-25998.

[6] Kaur H. Modelling internet of things driven sustainable food security system [J]. Benchmarking: An International Journal, 2021, 28(5): 1740-1760.

[7] Kogo B K, Kumar L, Koech R. Climate change and variability in Kenya: a review of impacts on agriculture and food security [J]. Environment, Development and Sustainability, 2021, 23: 23-43.

[8] Ben Hassen T, El Bilali H. Impacts of the Russia-Ukraine war on global food security: towards more sustainable and resilient food systems? [J]. Foods, 2022, 11(15): 2301.

[9] Kaur H. Modelling internet of things driven sustainable food security system [J]. Benchmarking: An International Journal, 2021, 28(5): 1740-1760.

[10] Mukhopadhyay R, Sarkar B, Jat H S, et al. Soil salinity under climate change: Challenges for sustainable agriculture and food security [J]. Journal of Environmental Management, 2021, 280: 111736.