World Food Forum Global Youth Forum

Title of the Thesis:

An Empirical Inquiry of the Key Influencing Factors for the Development of Smart Agriculture:

A Principal Component Analysis Approach

Student Name: Ren Changan

Advisor: Cao Lizhe

Shijiazhuang Foreign Language School

June 2024

Abstract

The convergence of state-of-the-art technologies, including the Internet of Things (IoT), artificial intelligence (AI), and big data analytics, is transforming the agricultural industry, establishing smart agriculture as an essential component for the global advancement of sustainable and efficient agricultural practices. This study conducts a comprehensive examination of the determinants of smart agriculture development by scrutinizing contemporary literature and employing principal component analysis (PCA) to synthesize key influential factors from a multi-dimensional dataset. Through a meticulous selection process, the study identifies four primary indicators and thirteen secondary indicators, which are then distilled into three pivotal principal components: digitalization, internet penetration, and infrastructure development. Utilizing SPSS software, the analysis elucidates the interplay between these components and their collective influence on the trajectory of smart agriculture. The findings underscore the imperative to bolster digitalization efforts, enhance internet accessibility, and expedite the establishment of robust digital infrastructure within the agricultural domain. The study's contributions are manifold, offering a robust analytical framework that can inform future research and strategic planning in the realm of agricultural innovation.

Keywords: Smart Agriculture; Principal Component Analysis; Influential Factors.

I. Research Background and Current Status

The agricultural sector is on the cusp of a technological renaissance, driven by the advent of artificial intelligence (AI), machine learning, and blockchain technologies. These innovations are not only redefining conventional farming practices but also steering the industry towards a future characterized by intelligence, automation, and sustainability. At the heart of this transformation is smart agriculture, an advanced model predicated on information and knowledge. It leverages the internet, the Internet of Things (IoT), big data, and AI to foster an eco-system that is green, standardized, interconnected, and intelligent.

The burgeoning prominence of smart agriculture has garnered international interest, prompting nations to craft policies tailored to their agricultural contexts. These policies are instrumental in nurturing smart agriculture and, by extension, in bolstering global food security. Leading agricultural nations, such as the United States and Germany, have spearheaded policy initiatives focused on technological innovation and talent cultivation in the realm of smart agriculture. China serves as a notable example of this proactive approach, having implemented policies that have significantly propelled the integration of modern information technology with agriculture. These measures have not only amplified the informatization of rural areas but have also catalyzed the adoption of smart agriculture practices. The impact of these policies is palpable, as evidenced by the remarkable statistics: the contribution of scientific and technological advancements to agriculture has surpassed 63%, and the mechanization rate for crop cultivation and harvesting has exceeded 73%. The deployment of 2.2 million agricultural machinery terminals equipped with satellite navigation systems stands as a testament to the heightened operational efficiency and precision in the sector. Moreover, the proliferation of nearly 200,000 plant protection drones, which cover an annual operation area of over 2.1 billion mu[1], underscores the sweeping digital transformation in China's agricultural landscape. The nation's pursuit of strengthening agricultural infrastructure and enhancing the comprehensive production capacity, coupled with the strategic application of big data in agriculture and rural development, marks a significant stride towards digitalization. This commitment to innovation and technological integration has yielded tangible outcomes, as demonstrated by China's grain production in 2023, which reached an impressive 1390.82 billion jin[2]. This achievement not only reflects the efficacy of the implemented measures but also signals the potential for smart agriculture to reshape the future of global food production and security.

II. Research Progress

The academic discourse on smart agriculture has been burgeoning, underscoring its burgeoning importance in the contemporary agricultural landscape. This study offers a synthesis of the extensive research conducted by scholars across various dimensions, with a particular emphasis on the pivotal factors that influence the adoption and proliferation of smart agriculture.

A. Government Policies

Scholars such as Guo Yongtian and Qian Jingfei have examined the strategies of agriculturally advanced nations like the United States and Australia. Their work underscores the importance of strategic planning, investment in information technology, talent development, and active government engagement in fostering smart agriculture. Wen Tiejun emphasizes the need for robust institutional frameworks, enhanced rural infrastructure, and policies aimed at elevating farmer incomes. Additionally, Zhang Yucheng highlights the critical role of strategic land use planning in agricultural development, while Hou Xiufang and Zhao Ruixue advocate for the

---

1 A traditional Chinese unit of area, equivalent to 666.7 square meters.
2 Equivalent to 500g.

integration of technology within agricultural production processes and the promotion of supply-side structural reforms.

B. Socio-economic and Cultural Foundations

Zhao Chunjiang identifies significant impediments to the advancement of smart agriculture in China, including low levels of agricultural mechanization, inadequate rural infrastructure, and the intricacies of subsidy mechanisms. Mou Xiaoyan and Zhao Minjuan approach the topic from an economic perspective, stressing the importance of capital allocation, talent, and information flow. Starting from the investigation of rural digital literacy, The Information Research Center of the Chinese Academy of Social Sciences posits that the education level, age, and income of modern agricultural practitioners will affect the promotion of the digital application ability, content creation ability, and collaboration ability of smart agriculture..

C. Talent Development and Integration

Lu Zhen suggests that the development of smart agriculture from the perspective of agricultural refinement should fully consider the talent demand and strengthen the cultivation of new agricultural talents. Peng Jianmin believes that the construction of smart agriculture infrastructure, talent cultivation, and the improvement of the smart agriculture standard system are important factors, and scholars such as Du Pu and Ye Xingyi hold the same view. Luo Xiwen and others believe that the development of smart agriculture should be based on the improvement of agricultural labor productivity, land yield, and resource utilization efficiency. Li Fangmin contends that governmental leadership is indispensable in the development of smart agriculture, strengthen the construction of legal and regulatory systems, accelerate technological innovation, improve smart agriculture infrastructure, and increase the reserve of talent cultivation. Li Jin supports the augmentation of that the modern agricultural technology basic research level should be increased, funding investment should be increased and talent cultivation should be strengthened.

Drawing from these scholarly perspectives, this paper combines the selection criteria of scholars and the research content to finally determine the scope of factor selection and further carry out research.

III. Research Methods and Significance

This study endeavors to isolate the principal determinants through rigorous data analysis, employing both correlation and principal component analyses to achieve a nuanced understanding of the subject. The significance of this study is multifaceted and is discussed below:

Policy Formulation: The findings are instrumental for the development of astute governmental agricultural policies. By deconstructing the key elements that propel smart agriculture, policymakers can devise targeted initiatives that catalyze its advancement, thereby fostering an environment conducive to innovation and efficiency within the sector.

Acceleration of Adoption: Identifying the drivers of smart agriculture development is essential for prioritizing investments in pivotal areas and for addressing and overcoming the constraints that may hinder progress. This proactive approach can enhance the agrarian community's receptivity to smart agricultural practices, thereby facilitating the widespread adoption of innovative farming techniques.

Resource Allocation: By making the market aware of the significant advantages of smart agriculture and clearly understanding the factors that affect its development, it is beneficial to increase the market's level of understanding and attention to smart agriculture, and to invest more funds in the development of smart agriculture, fully exerting the decisive role of the market in resource allocation. At the same time, by combining the effective market with capable government, we can strive to achieve rational allocation of resources related to smart agriculture.

This study comprehensively reviews the criteria for selecting factors impacting smart agriculture development, as identified by scholars. Through correlation analysis, it discerns the influential factors and applies principal component analysis for dimensionality reduction. The approach streamlines the complex determinants into a concise framework of key determinants, offering a coherent framework for academic and strategic discussions in the field of smart agriculture.

IV. Indicator Selection and Methodological Rigor

A. Selection of Influencing Factors and Data Source

This study chooses a set of indicators to capture the multifaceted dimensions of smart agriculture development. The framework established is grounded in scholarly consensus, the research constructs a framework that encompasses four primary and thirteen secondary indicators. This selection is informed by empirical assessments of digital rural development across 2,481 counties and districts, as reported by the New Rural Development Research Institute of Peking University, aligning with the analytical constructs of the Digital Rural Project Team.

The data underpinning this study is sourced from a constellation of authoritative entities, including the Ministry of Agriculture and Rural Affairs, the National Bureau of Statistics, Winds Information, and the New Rural Development Research Institute of Peking University.

Tab 1 Key Influencing Factors Indicators

Primary Indicators	Secondary Indicators
Human Resources Quality	Agricultural Machinery Operators in Rural Areas
	Per Capita Disposable Income of Rural Residents
	Number of College and Above per 100,000 People
	Number of Undergraduate Students
Digitalization Level	Number of Computers per 100 People
	Employees in High-tech Enterprises
	Number of High-tech Enterprises
	Number of Technology Patent Applications
Policy and Institutional Construction	Public Budget Expenditure on Agriculture, Forestry, and Water Affairs
	Length of Optical Cable Usage
Infrastructure Construction	Amount Spent on Agricultural Machinery Purchases
	Number of Agricultural Machinery Organizations
	Urbanization Rate

B. Utilization of Correlation Analysis Methodology

To explore the interlinkages between the selected indicators and the evolution of smart agriculture, this study meticulously performs a correlation analysis. This methodological choice is pivotal in discerning the interdependencies among the indicators and the overall digital rural development spectrum. The analysis delves into the central tendencies, distributions, and outlier characteristics of the data, utilizing the Pearson correlation coefficient to quantify the linear relationships with precision.

Employing IBM SPSS Statistics 27 software, the study quantitatively examines the key factors underpinning the proliferation of smart agriculture, as elucidated in Tables 2 and Tables 3.

In conjunction with the Pearson simple correlation coefficient, the findings reveal significant positive correlations: between the "number of computers used per hundred people" and the prevalence of smart agriculture, and between "financial public budget expenditure on agriculture, forestry, and water" and the total agricultural output value. Moreover, the "urbanization rate" is positively correlated with the spread of smart agricultural practices.

Tab 2 Correlation of Key Factors

		Total Agricultural Output Value	Digital Rural Index 20	Digital Infrastructure Index 20	Rural Economic Digitalization Index 20	Digital Production Index 20
Number of Computers per 100 People	Pearson correlation coefficient	-0.083	**0.997**	**0.995**	**0.997**	**0.998**
	Significance (Two-tailed)	0.650	0	0	0	0
Public Budget Expenditure on Agriculture, Forestry, and Water Affairs	Pearson correlation coefficient	**0.656**	**0.615**	**0.619**	**0.614**	**0.604**
	Significance (Two-tailed)	0.000	0	0	0	0
Urbanization Rate	Pearson correlation coefficient	-0.076	**0.997**	**0.995**	**0.998**	**0.999**
	Significance (Two-tailed)	0.679	0	0	0	0
Number of High-tech Enterprises	Pearson correlation coefficient	0.249	-0.109	-0.114	-0.103	-0.094
	Significance (Two-tailed)	0.170	0.552	0.553	0.575	0.611
Employees in High-tech Enterprises	Pearson correlation coefficient	0.261	-0.127	-0.131	-0.122	-0.113
	Significance (Two-tailed)	0.149	0.488	0.474	0.507	0.536
Number of Technology Patent Applications	Pearson correlation coefficient	0.254	-0.056	-0.057	-0.053	-0.049
	Significance (Two-tailed)	0.160	0.759	0.756	0.775	0.789

Further observations on digitalization indicators correlation, the study observes a robust positive correlation among digitalization indicators, specifically the Number of High-tech Enterprises, Employees in High-tech Enterprises, and Number of Technology Patent Applications.

These correlations extend to other indicators in relation to the Digital Rural Development Index, this interconnectivity points to a multidimensional interplay that can be effectively condensed through principal component analysis(PCA).

Tab 3 Correlation of Digitalization Level Indicators

		Number of High-tech Enterprises	Employees in High-tech Enterprises	Number of Technology Patent Applications
Number of High-tech Enterprises	Pearson correlation coefficient	1.000	**0.991**	**0.091**
	Significance (Two-tailed)		0.000	0.000
Employees in High-tech Enterprises	Pearson correlation coefficient	**0.991**	1.000	**0.930**
	Significance (Two-tailed)	0.000		0.000
Number of Technology Patent Applications	Pearson correlation coefficient	**0.913**	**0.930**	1.000
	Significance (Two-tailed)	0.000	0.000	

V. Principal Component Analysis: Methodological Elaboration and Interpretation

A. Introduction to Principal Component Analysis (PCA)

Principal Component Analysis (PCA) serves as a pivotal analytical tool in the multivariate statistical arsenal, adept at transmuting a constellation of interrelated variables into a more succinct array of uncorrelated, composite indicators. This transformative process, predicated on the tenets of "dimension reduction," amalgamates the original variables into principal components, capturing the essence of the data while being sequentially ranking them by their variance contribution. Without altering the total variance of the variables, PCA employs a linear transformation to convert the interrelated variables into a sequence of uncorrelated groups, ordered by decreasing variance. The group with the largest variance is the first variable group, designated as the first principal component; the second variable group, which has the second largest variance and is uncorrelated with the first, is known as the second principal component. This pattern continues, with the I-th variable group also being referred to as the I-th principal component.

B. Suitability Assessment for PCA

The application of PCA is contingent upon fulfilling two principal criteria for dimensionality reduction: Firstly, a Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy that exceeds the threshold of 0.5, indicative of the suitability of the data for PCA. Secondly, a significance level (Sig value) from Bartlett's test of sphericity that is typically less than 0.05, further substantiating the appropriateness of PCA.

In this study, the data underwent standardization to mitigate dimensional constraints, subsequently subjected to the KMO and Bartlett's test of sphericity. The KMO value of 0.607, surpassing the 0.5 benchmark, alongside a significance level well below 0.05 (0.00), collectively endorse the aptness of the data for PCA.

Tab 4 KMO and Bartlett's Test of Sphericity

KMO Measure of Sampling Adequacy		0.607
Bartlett's test of sphericity	Approximate Chi-Square	585.435
	Degrees of Freedom	78
	Significance	0.000

C. Identification of Principal Components

Tab 5 Total Variance Explained by Indicators

Components	Initial Eigenvalues			Sum of Squared Loadings of Extracted Factors			Sum of Squared Loadings of Rotated Factors		
	Total	Percentage of Variance	Cumulative %	Total	Percentage of Variance	Cumulative %	Total	Percentage of Variance	Cumulative %
1	**4.981**	38.312	38.312	4.981	38.312	38.312	4.724	36.342	36.342
2	**4.171**	32.083	70.394	4.171	32.083	70.394	3.911	30.081	66.423
3	**1.715**	13.19	**83.584**	1.715	13.19	83.584	2.231	17.161	83.584
4	0.949	7.299	90.882						
5	0.431	3.317	94.2						
6	0.387	2.975	97.175						

7	0.175	1.342	98.517
8	0.075	0.58	99.097
9	0.06	0.464	99.561
10	0.031	0.238	99.798
11	0.014	0.106	99.904
12	0.009	0.07	99.974
13	0.003	0.026	100

The Principal Component Analysis (PCA) conducted in this study yielded a total of 13 components, each contributing to the analytical framework. The extraction of components with eigenvalues exceeding 1.0 culminated in the identification of three principal components (F1, F2, F3), characterized by eigenvalues of 4.981, 4.171, and 1.715, respectively. These components contribute to the variance in the data at rates of 38.312%, 32.083%, and 13.190%, respectively, amassing to a cumulative contribution of 83.584%. This substantial representation justifies the focus on these three principal components as the cornerstone for understanding the dynamics of smart agriculture development.

D. Determination of Factor Loadings for Principal Components

Tab 6 Factor Loading Matrix

Influence Factors	**Components**		
	1	**2**	**3**
Public Budget Expenditure on Agriculture, Forestry, and Water Affairs	0.546	-0.58	0.43
Number of High-tech Enterprises	0.962	0.221	-0.026
Employees in High-tech Enterprises	0.964	0.161	-0.044
Number of Computers per 100 People	0.075	0.864	0.015
Number of College and Above per 100,000 People	0.148	0.881	0.002
Number of Technology Patent Applications	0.89	0.073	-0.131
Amount Spent on Agricultural Machinery Purchases	0.019	-0.247	0.916
Number of Agricultural Machinery	0.226	-0.689	0.526

Organizations			
Per Capita Disposable Income of Rural Residents	-0.165	0.352	0.719
Agricultural Machinery Operators in Rural Areas	0.216	-0.663	0.582
Urbanization Rate	0.401	0.768	0.058
Length of Optical Cable Usage	0.74	-0.53	0.118
Number of Undergraduate Students	0.958	-0.031	0.193

The interpretation of the principal components is grounded in the magnitude of their respective factor loadings, influential factors with higher loadings are extracted to interpret the principal component and to name it accordingly.

In the first principal component, the variables of Employees in High-tech Enterprises, Number of High-tech Enterprises, Number of Technology Patent Applications, and Number of Undergraduate Students exhibit higher loadings and are generally related to the level of local technological development. Integrating the correlation analysis, the first principal component can be summarized as the Digitalization Level.

In the second principal component, the variables of Number of Computers per 100 People, Number of College and Above per 100,000 People, and Urbanization Rate show higher loadings and are related to the opportunities for individuals in the region to access and proficiently use internet technology. Combining the correlation analysis, the second principal component is concluded as the Internet Penetration Level.

In the third principal component, the variables of Amount Spent on Agricultural Machinery Purchases, Per Capita Disposable Income of Rural Residents, and Public Budget Expenditure on Agriculture, Forestry, and Water Affairs have higher loadings and are associated with government support and investment in infrastructure construction. With the correlation analysis, the third principal component is encapsulated as the Level of Infrastructure Construction.

E. Expression of Principal Components

The resultant principal components, as delineated in Table 6, were quantified through their respective weight coefficients presented in Table 7. The mathematical expressions for the three principal components, encapsulating the weighted contributions of the original variables, are as follows:

$$F1=0.099x1+0.209x2+0.211x3+0.013x4+0.029x5+0.199x6-0.040x7+0.026x8-0.073x9+0.021x10+0.081x11+0.156x12+0.198x13$$

$$F2=-0.114x1+0.057x2+0.038x3+0.251x4+0.255x5+0.001x6+0.048x7-0.131x8+0.196x9-0.116x10+0.228x11-0.140x12+0.013x13$$

$$F3=0.120x1-0.032x2-0.049x3+0.117x4+0.109x5-0.102x6+0.441x7+0.171x8+0.426x9+0.204x10+0.111x11-0.045x12+0.049x13$$

Tab 7 Indicator Component Score Weight Matrix

	Components		
	1	2	3
Public Budget Expenditure on Agriculture, Forestry, and Water Affairs	0.099	-0.114	0.12
Number of High-tech Enterprises	0.209	0.057	-0.032
Employees in High-tech Enterprises	0.211	0.038	-0.049
Number of Computers per 100 People	0.013	0.251	0.117
Number of College and Above per 100,000 People	0.029	0.255	0.109
Number of Technology Patent Applications	0.199	0.001	-0.102
Amount Spent on Agricultural Machinery Purchases	-0.04	0.048	0.441
Number of Agricultural Machinery Organizations	0.026	-0.131	0.171
Per Capita Disposable Income of Rural Residents	-0.073	0.196	0.426
Agricultural Machinery Operators in Rural Areas	0.021	-0.116	0.204
Urbanization Rate	0.081	0.228	0.111
Length of Optical Cable Usage	0.156	-0.14	-0.045
Number of Undergraduate Students	0.198	0.013	0.049

VI. Summary of Findings and Strategic Recommendations

This study presents a comprehensive analysis of prevailing academic discourses and empirical data, underpinned by the methodological rigor of correlation and principal component analyses, has yielded pivotal insights into the determinants of smart agriculture proliferation. The study's construction of key factors from a hierarchy of indicators—four primary and thirteen secondary—has culminated in the identification of three principal components: digitalization, internet penetration, and infrastructure development. These components are deemed critical for the advancement of smart agricultural practices. The following recommendations are proposed to foster the continued evolution of smart agriculture:

Enhancement of Technological Innovation: It is imperative to bolster science and technology within the smart agriculture sector. This can be achieved by launching flagship projects that focus on pivotal areas and industries, facilitating breakthroughs in core technologies, and tailoring initiatives to regional strengths. The leveraging of modern IT to augment agricultural efficiency is a cornerstone for the robust development of smart agriculture.

Talent Cultivation and Development: Recognizing the pivotal role of digital agriculture talents, it is essential to establish and refine training mechanisms for such professionals. A multi-tiered and diverse educational framework, encompassing vocational institutions and online platforms, should be developed to nurture a cohort of talents well-versed in both agricultural practices and digital technologies.

Infrastructure Investment and Expansion: Amidst the backdrop of digital transformation, there is a pressing need to reinforce rural digital infrastructure. This entails scaling up investment in rural informatization, broadening network accessibility, and enhancing internet speed and quality in rural regions. Such measures are fundamental in establishing a solid foundation for data management within smart agriculture ecosystems.

References

[1] Guo Yongtian. The Characteristics and Enlightenment of Modern Agricultural Development in Australia [J]. World Agriculture, 2016(1):5. DOI: CNKI: SUN: SJNY.0.2016-01-035.

[2] Qian Jingfei, Chen Yangfen. The Enlightenment of Agricultural Information Construction in Typical Developed Countries to China's Agricultural "New Infrastructure" [J]. Research on Science and Management, 2021, 41(23):7.

[3] Wen Tiejun. The Key to Promoting the Modernization of Agriculture and Rural Areas [J]. Chinese Ecological Civilization, 2021, 000(002): P.23-26.

[4] Zhang Yucheng, Wu Wei, Dai Feng, et al. Reflections on the Development Model of Smart Agriculture in the Countryside Revitalization with the Help of Science and Technology. Bulletin of the Chinese Academy of Sciences, 2024, 39(3): 563-571, doi: 10.16418/j.issn.1000-3045.20240117002.

[5] Hou Xiufang, Wang Dong. The Path Choice of China's "Smart Agriculture" Development in the New Era [J]. Macroeconomic Management, 2017(12):5. DOI: CNKI: SUN: HGJG.0.2017-12-015.

[6] Zhao Ruixue, Yang Chenxue, Zheng Jianhua, et al. The Current Status and Prospects of Agricultural Intelligent Knowledge Service Research [J]. Smart Agriculture (Chinese and English), 2022, 4(4):105-125. DOI: 10.12133/j.smartag.SA202207009.

[7] Zhao Chunjiang. The Current Status and Future Prospects of Smart Agriculture [J]. Journal of South China Agricultural University, 2021, 42(6):7.

[8] Mou Xiaoyan, Wu Zitao. Basic Goals, Key Issues, and Implementation Path of Smart Agriculture Development from the Perspective of Green Development Concept [J]. Research on Science Management, 2021, 39(4):131-136.

[9] Chang Qian, Li Xiaoping, Xie Xiong, et al. The Impact of Non-agricultural Employment on Farmers' Ecological Production Behavior: Based on the Mediating Effect of Agricultural Production and Management Characteristics and the Moderating Effect of Family Life Cycle [J]. Chinese Rural Observation, 2020(1):18. DOI: CNKI: SUN: ZNCG.0.2020-01-006.

[10] Research Center for Informationization of the Chinese Academy of Social Sciences. Analysis Report on Rural Digital Literacy Survey under the Background of Rural Revitalization Strategy. 2021.

[11] Lu Zhen. Problems and Solutions in the Promotion of Smart Agriculture [J]. Hebei Agricultural Machinery, 2023(5):46-48.

[12] Peng Jianmin, Song Xuezhao, Han Keli. Analysis of the Development Trend of Smart Agriculture—Based on "Current Status and Prospects of Smart Agriculture" [J]. Hebei Agricultural Machinery, 2023(11):154-156.

[13] Du Pu, Qiao Guanri, Wang Di, et al. Research on the Development of Scientific and Technological Power Support for New Smart Agriculture Applications [J]. Journal of Shandong University of Agriculture and Engineering, 2023, 40(4):26-30.

[14] Ye Xingyi, Xie Chuang. Analysis of the Problems and Strategies of Smart Agriculture Development in China under the Rural Revitalization Strategy [J]. Agricultural Economy, 2023(12):3-6.

[15] Luo Xiwen, Liao Juan, Zang Ying, et al. The Development Direction of China's Agricultural Production: From Mechanization to Smartization [J]. Chinese Engineering Science, 2022, 24(1):9. DOI: 10.15302/J-SSCAE-2022.01.005.

[16] Li Fangmin. Research on the Current Problems of Smart Agriculture Development in China from the Perspective of Rural Revitalization [J]. China's Science and Technology Industry, 2024(03):68-70. DOI: 10.16277/j.cnki.cn11-2502/n.2024.03.032.1

[17] Li Jin, Ma Chen, Zhao Chunjiang, et al. Strategic Path and Policy Suggestions for "Internet Plus" Modern Agriculture [J]. Chinese Engineering Science, 2020, 22(4):8. DOI: 10.15302/J-SSCAE-2020.04.006.