Measuring Agricultural Modernization and Diagnosing Development Bottlenecks in China's Provinces: Evidence from high-quality development and new quality productivity

Shuokai Wang[1], Bo Zeng[1,*]

[1] School of Economics and Management, Beijing Forestry University, Beijing, China 100091

* Email: zengbo@bjfu.edu.cn (Corresponding Author)

Received 20 November 2024 · Revised 4 January 2025 · Accepted 18 February 2025 · Published 30 June 2025

Abstract

This study presents a provincial framework for assessing agricultural modernization through high-quality development and new quality productivity. The reported 2020–2022 analysis combines indicator scoring, clustering, obstacle diagnosis, and a stacking-based prediction design. It describes spatial variation and potential province-specific constraints. Table 1 lists 28 indicators; the complete indicator-to-code mapping, X-category scores, and model-comparison outputs were unavailable for independent verification. The reported rankings, score thresholds, and claims of superior predictive performance therefore require verification against the underlying data and code. The framework and policy implications should be interpreted subject to this verification.

Keywords: High-quality development; new quality productivity; agricultural modernization; stacking ensemble learning; K-means clustering; obstacle-degree analysis

Measuring Agricultural Modernization and Diagnosing Development Bottlenecks in China's Provinces: Evidence from high-quality development and new quality productivity

1. Introduction

China's agricultural modernization goals have continuously evolved in step with national development strategies and the dominant constraints of each era. From the early "four modernizations" agenda and management-oriented reforms, through internationalization and large-scale mechanized production, agricultural modernization has gradually shifted from an output-expansion pathway toward a broader concern with quality, efficiency, sustainability, and innovation (Huang, 2007; Ye et al., 2021; Zheng, 2023). The country has now entered a new stage that places high-quality development at the center and treats new quality productive forces as a foundational driver.

Over the past decade, agricultural modernization in China has been propelled largely by scale expansion and technology diffusion, including mechanization and advances in agricultural biotechnology and input technologies. These efforts have supported stable grain production and a broader, more diversified supply of agricultural and livestock products. National policy documents also report substantial mechanization progress, including a comprehensive mechanization rate for crop cultivation and harvesting of approximately 71% (Central Committee of the Communist Party of China & State Council of the People's Republic of China, 2021). As a result, the long-standing challenge of ensuring basic food security for a large population has been substantially alleviated.

Yet a speed- and output-centered narrative is increasingly insufficient for evaluating modernization today. Modernization quality has often been under-discussed and under-measured, despite persistent constraints involving production efficiency, comparative returns, resource use, and ecological sustainability (Du et al., 2021; He et al., 2017; Xiao & Zhang, 2023). Agricultural labor productivity also remains substantially below that of non-agricultural sectors, underscoring the need to improve productivity rather than merely expand output (Xi, 2023). The core challenge is therefore no longer only "producing enough," but "producing well" under resource and environmental constraints. Accordingly, China requires a multidimensional evaluation system that moves beyond traditional single-dimensional approaches and aligns with contemporary strategic priorities.

Since 2017, "high-quality development" has been established as a major national objective grounded in the new development philosophy. More recently, national policy has emphasized the development of a modern industrial system and new quality productive forces as key foundations for high-quality growth (Li, 2024; National Development and Reform Commission, 2024). Together, these agendas signal a dual shift: from growth that prioritizes quantity to development that prioritizes quality and balance, and from factor-driven expansion to innovation-driven productivity upgrading.

Against this backdrop, evaluating provincial agricultural modernization through the combined lenses of high-quality development and new productive quality forces is both timely and policy relevant. This study therefore reframes agricultural modernization measurement around these two strategic orientations, aiming to (i) produce a comparable provincial benchmarking tool, (ii) identify development types among lower-performing regions, and (iii) pinpoint binding constraints that can guide targeted interventions. By doing so, we seek to provide an evidence-based diagnostic foundation for promoting more comprehensive, higher-quality, and more innovation-driven agricultural modernization in China.

Essential innovations relative to existing literature are threefold. First, unlike studies that evaluate agricultural modernization through single-dimensional output expansion or rely on a single strategic lens, we integrate high-quality development and new quality productivity into a unified provincial assessment framework.

Second, rather than applying only conventional entropy/TOPSIS-style weighting or linear screening, we employ correlation diagnostics and multi-learner, tree-based importance estimation to accommodate nonlinear relationships and high-dimensional indicators; the model family draws on established random-forest and gradient-boosting approaches (Breiman, 2001; Ke et al., 2017; Prokhorenkova et al., 2018). Third, we connect measurement to actionability by coupling K-means typology with Pareto-style obstacle-degree diagnostics, thereby producing cluster-specific "critical few" bottlenecks and policy priorities, and we introduce a stacking ensemble to strengthen predictive monitoring (Wolpert, 1992).

Research problems. This manuscript addresses four questions: RQ1: How can provincial agricultural modernization be measured in a way that simultaneously reflects the policy priorities of high-quality development and the emerging concept of new quality productive forces? RQ2: Given a high-dimensional indicator system, how can redundant indicators be screened and indicator importance be learned under potential non-linearity? RQ3: How can low-performing provinces be classified into interpretable development types and diagnosed in terms of binding bottlenecks rather than broad, non-actionable deficits? RQ4: Can ensemble learning improve the robustness and accuracy of monitoring/predicting key benchmark indicators relative to single learners?

2. Literature Review

To align with this study's objectives, we organize the related scholarship into three streams: (i) high-quality agricultural development, (ii) the concept and measurement of new quality productivity, and (iii) agricultural modernization and its evaluation frameworks.

2.1 Research on High-Quality Agricultural Development

High-quality agricultural development is widely regarded as a core component of national high-quality development, and its conceptual scope has expanded as research deepened after the policy discourse emerged in 2017. Early studies proposed multi-dimensional understandings, for example, emphasizing agricultural economic growth, development outcomes, rural social development, and sustainability (He et al., 2017). Later, scholars further incorporated institutional construction and efficiency-oriented mechanisms, thereby broadening the analytical framework for high-quality agricultural development (Du et al., 2021).

In recent years, research has increasingly moved beyond the generic "new development philosophy" lens and examined high-quality agricultural development through context-specific drivers that reflect contemporary transformation pressures, such as the digital economy and rural revitalization. Evidence suggests that the digital economy can promote high-quality agricultural development through multiple mechanisms (Wang et al., 2023). Related discussions also position digital transformation as a pathway for overcoming structural constraints and enhancing the quality of agricultural development (Wang & Yang, 2023). From a regional perspective, studies have emphasized the role of new-quality productive forces in supporting the high-quality development of agriculture in Northeast China (Hou et al., 2024).

2.2 Definition and Research on New Quality Productivity

New quality productive forces have emerged as a policy-oriented construct that highlights innovation-led upgrading under high-quality development. Policy narratives emphasize that their development should be driven by technological innovation, industrial upgrading, and the modernization of productive factors (National Development and Reform Commission, 2024).

Although the concept is relatively new, domestic discussions have begun to converge on several core elements. A representative definition conceptualizes new quality productive forces as innovation-led, oriented toward strategic emerging and future industries, and capable of balancing digital and green transformation so that potential productivity can be converted into realized productivity (Xu et al., 2023). In this framing, "newness" lies not in abandoning traditional productivity components, but in digitally reconfiguring labor, objects of

labor, and means of production to fit the characteristics of the current technological era, thereby improving the materialization and effectiveness of productive forces (Xu et al., 2023).

Extant applications of the new quality productive forces lens have expanded across multiple domains. Some studies discuss its importance for building an "agricultural powerhouse" and restructuring agricultural development pathways (Lin & Dong, 2024). Others examine enabling conditions in state-owned enterprises and sectoral restructuring (Song & Lyu, 2024), as well as industry-specific transitions such as the high-quality development of the cultural and tourism industry (Qi & Li, 2024). Empirically, emerging work has attempted to measure new quality productive forces at city or regional levels using methods such as TOPSIS, entropy weighting, Dagum Gini decomposition, kernel density estimation, Moran's I, and regression-based analyses (Cao et al., 2024; Fu & Cai, 2024; Han et al., 2024).

2.3 Research on Agricultural Modernization Development

Since the reform and opening-up period, China's agricultural modernization has been characterized by path dependence on national conditions, with the definition and priorities of modernization adjusting to changing macro environments. Under the current development stage, agricultural modernization is increasingly described as a system-level transformation that requires demand-driven orientation, supply-side upgrading, improved benefit distribution, and further progress in marketization, standardization, specialization, integration, precision, and intelligence (Research Group of the Rural Economy Research Department, Development Research Center of the State Council, et al., 2021).

A growing body of work treats high-quality agricultural development as an essential stage within agricultural modernization and applies diverse quantitative methods to measure regional progress and its determinants. For example, studies have examined the relationship between the digital economy and high-quality agricultural development using regression frameworks (Wang et al., 2023), assessed the coupling coordination between water-resources carrying capacity and high-quality agricultural development using weighting and spatial analytical methods (Peng & Wu, 2023), and explored spatiotemporal patterns and influencing factors using entropy weighting, indicator evaluation, and spatial analytical tools (Xiao & Zhang, 2023; Xu et al., 2024).

Despite these advances, research that explicitly integrates high-quality development with new quality productive forces in a unified agricultural modernization assessment remains limited. Moreover, traditional indicator-screening and weighting approaches, such as regression-based selection and TOPSIS-style procedures, may be less flexible when faced with high-dimensional indicators and nonlinear relationships. This motivates the present study to jointly adopt the two strategic lenses and introduce machine-learning-based indicator screening and ensemble prediction, thereby constructing a more adaptive and reproducible evaluation-and-diagnosis pipeline for provincial agricultural modernization.

2.4 Summary of research gaps and this study's contributions.

Existing studies provide rich discussions on (i) multi-dimensional high-quality agricultural development, (ii) early measurement attempts for new quality productivity at city/provincial levels, and (iii) agricultural modernization evaluation methods (entropy weighting, TOPSIS, coupling coordination, grey relational analysis, and spatial statistics). However, three shortcomings remain salient. First, integration is limited: few studies jointly operationalize high-quality development and new quality productivity within the same agricultural modernization benchmark, weakening cross-province comparability under the new policy discourse. Second, many indicator-screening/weighting routines are sensitive to linearity assumptions and may struggle with high-dimensional indicators and non-linear interactions, reducing robustness when indicators are strongly coupled (especially in the digital domain). Third, policy translation is often incomplete: even when composite indices are constructed, the linkage from "score differences" to cluster-specific binding constraints and implementable policy sequencing is not fully established. To fill these gaps, this manuscript proposes a replicable pipeline—indicator construction → redundancy screening → multi-learner scoring →

province benchmarking → clustering typology → obstacle-degree diagnosis → stacking-based prediction—thereby strengthening both measurement validity and policy actionability.

3. Construction of an Indicator System for High-Quality Agricultural Modernization

Building on prior work on agricultural modernization evaluation and high-quality agricultural development (Xin & An, 2019; Hao & Shen, 2022; Zhang & Wang, 2023; Zheng, 2023; Xu et al., 2024), this study constructs an integrated indicator system that links high-quality development with new quality productive forces to assess provincial agricultural modernization.

In contrast to single-dimensional, speed-oriented assessments, the proposed system emphasizes both quality and efficiency and is designed to be consistent with China's contemporary policy priorities concerning rural revitalization and agricultural modernization (Central Committee of the Communist Party of China & State Council of the People's Republic of China, 2021).

3.1 Indicator-selection logic and structure

Table 1 lists 28 indicators: 20 for High-Quality Development and 8 for New Quality Productivity. The discrepancy with the 31 indicators reported for the analysis remains unresolved.

The first-dimension captures innovation, green production, coordination, openness, and shared benefits; the second dimension captures digital rural development and agricultural science-and-technology capability. Table 1 presents the agricultural modernization evaluation index system.

3.2 Evaluation Index System

Table 1. Agricultural Modernization Evaluation Index System

Objective layer	Tier-1 indicator	Tier-2 indicator	Specific indicator	Unit	Attribute
High-Quality Development	Innovation	Innovation input	Total power of agricultural machinery	10,000 kW	+
High-Quality Development	Innovation	Resource-allocation efficiency	Completed fixed-asset investment by rural households	billion yuan	+
High-Quality Development	Green development	Green living	Solar water-heater area	10,000 m²	+
High-Quality Development	Green development	Green production	Pesticide usage	tons	–
High-Quality Development	Green development	Green production	Fertilizer usage	tons	–
High-Quality Development	Green development	Green production	Water-treatment area	1,000 hectares	+
High-Quality Development	Green development	Green production	Agricultural plastic-film usage	tons	–

Objective layer	Tier-1 indicator	Tier-2 indicator	Specific indicator	Unit	Attribute
High-Quality Development	Coordination	Urban–rural dual economic structure	Per-capita consumption expenditure of rural residents	yuan	+
High-Quality Development	Openness	Rural labor force	Urban–rural income ratio (rural = 1)	unitless ratio	–
High-Quality Development	Openness	Rural labor force	Producer Price Index of agricultural products	**Unitless index**	+
High-Quality Development	Openness	Rural labor force	Share of rural employment in primary industry	%	+
High-Quality Development	Openness	Rural labor force	Number of rural laborers	persons	+
High-Quality Development	Openness	Cross-border e-commerce	Number of comprehensive pilot zones for cross-border e-commerce	count	+
High-Quality Development	Sharing	Living standard	Completed housing area	10,000 m²	+
High-Quality Development	Sharing	Infrastructure	Rural water-supply coverage	%	+
High-Quality Development	Sharing	Infrastructure	Rural gas penetration rate	%	+
High-Quality Development	Sharing	Infrastructure	Length of paved (hardened) roads	km	+
High-Quality Development	Sharing	Output level	Agricultural expenditure budget	billion yuan	+
High-Quality Development	Sharing	Output level	Agricultural output value	% (as reported)	+
High-Quality Development	Sharing	Social security	Number of rural doctors	persons	+
New Quality Productivity	Digital rural development	Digital environment	Disposable-income-to-consumption ratio of rural residents	%	+
New Quality Productivity	Digital rural development	Digital environment	Number of Taobao villages	count	+
New Quality Productivity	Digital rural development	Digital environment	Peking University Digital Inclusive Finance Index	**Unitless index**	+

Objective layer	Tier-1 indicator	Tier-2 indicator	Specific indicator	Unit	Attribute
New Quality Productivity	Digital rural development	Digital infrastructure	Rural electricity consumption	100 million kWh	+
New Quality Productivity	Digital rural development	Digital infrastructure	Rural broadband access users	10,000 households	+
New Quality Productivity	Agricultural S&T development	Network infrastructure	Number of domain names	10,000	+
New Quality Productivity	Agricultural S&T development	Network infrastructure	Capacity of mobile telephone exchanges	10,000 households	+
New Quality Productivity	Agricultural S&T development	Network infrastructure	Optical-cable line length	km	+

3.3 Data sources

The indicator data are drawn from China Rural Statistical Yearbook, China Science and Technology Statistical Yearbook, China Environmental Statistical Yearbook, and China Statistical Yearbook (2020–2022), as well as third-party sources including the Alibaba Research Institute, the Spatial Planning Research Center of Nanjing University, and Cube Data Society.

4. Data Preprocessing and Indicator Scoring

4.1 Data Cleaning

The manuscript reports mean imputation for missing indicator values.

Second, given substantial and policy-relevant heterogeneity in agricultural modernization across provinces, outlier detection and winsorization were not performed, so as to avoid mechanically truncating meaningful interregional differences.

Z-score standardization is reported as implemented in Python; the data and code were unavailable for independent verification.

4.2 Feature Engineering

To improve readability and facilitate model training, we re-labeled the two indicator groups via label encoding. Importantly, we unify the notation as follows: high-quality development indicators are denoted by G1…Gm, and new quality productivity indicators by X1…Xn.

4.3 Correlation Analysis and Selection of Predictive Indicators

Following Pearson correlation analysis (Pearson, 1895), this study generated correlation heatmaps for both indicator groups over 2020–2022, corresponding to Figures 1 and 2.

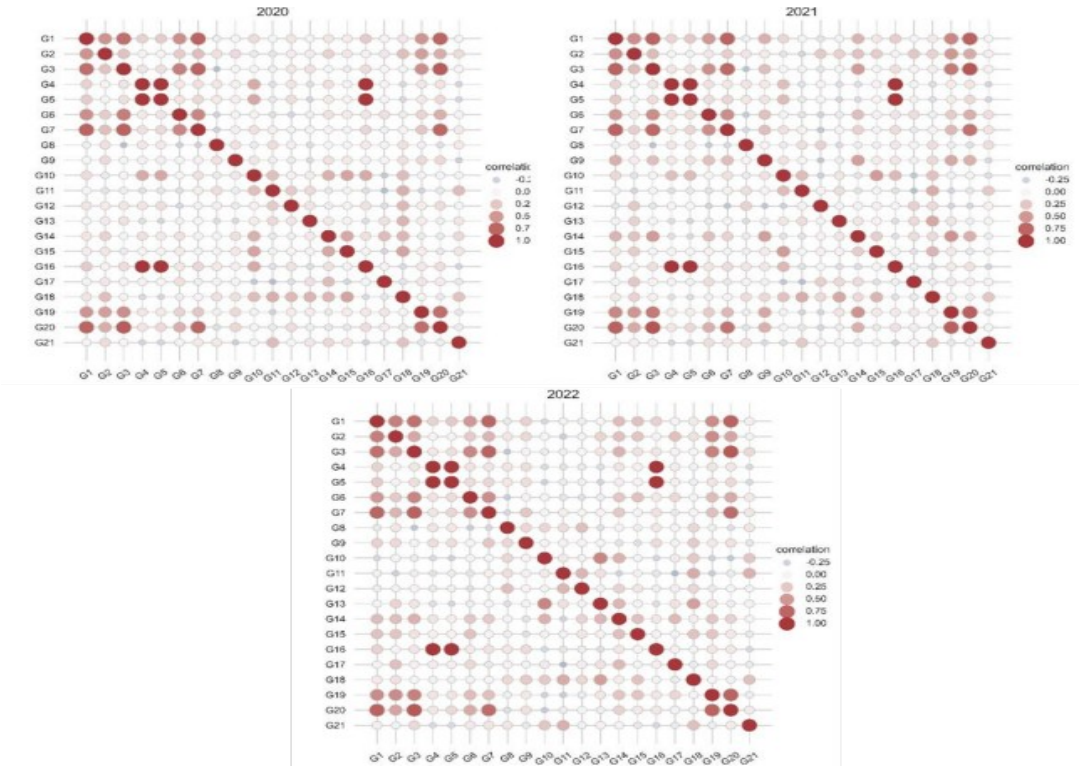


Figure 1. High-Quality Development Indicators (2020)

Across 2020–2022, most high-quality development indicators exhibit weak pairwise correlations, suggesting limited redundancy overall.

However, indicators G4, G5, and G16 remain strongly correlated across years, implying near-duplicate informational content. Accordingly, we retained G4 and excluded G5 and G16 to reduce multicollinearity while preserving interpretability.

Compared with high-quality development, the new quality productivity indicators show materially stronger correlations, consistent with the tighter coupling among digital-infrastructure and e-commerce variables.

Only two highly correlated pairs are salient in 2022 (X4 with X6; X5 with X7). For consistency across years and for model parsimony, we excluded X5 and X6 in all three years and retained X4 and X7.

For predictive benchmarking, we select G1 (agricultural output value) as the core outcome proxy for high-quality development, given its integrative reflection of mechanization, ecological conditions, and production efficiency/benefits.

Similarly, we select X1 (number of Taobao villages) as the key proxy for new quality productivity, because it directly captures rural e-commerce development and aligns with the “technology-driven” emphasis of the new quality productivity concept.

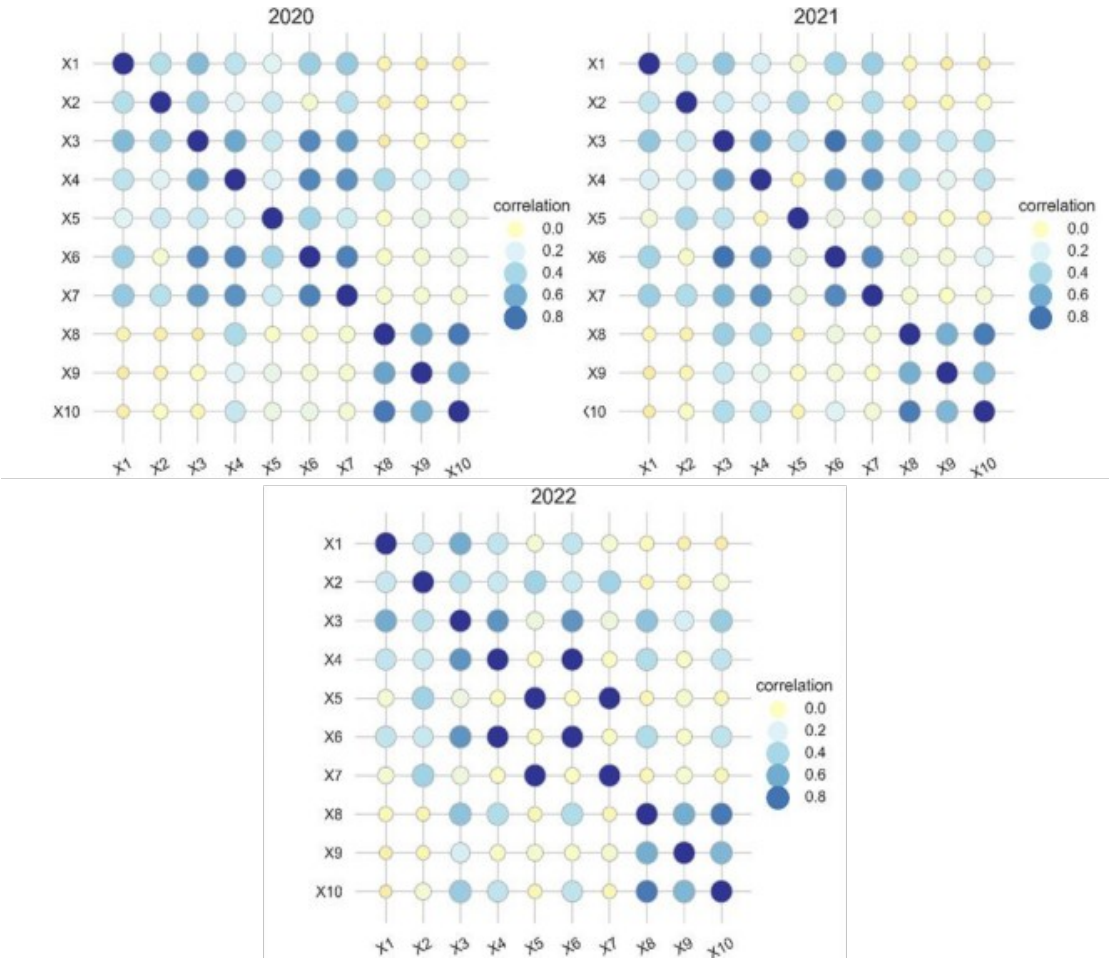


Figure 2. New Quality Productivity Indicators (2022)

4.4 Indicator Scoring Results Based on Multiple Models

Using the selected base learners, this study scores the remaining indicators in each dimension by treating G1 and X1 as benchmark targets. Final indicator scores are computed as the average importance/score across learners, and the results are reported in the text.

The article reports score cutoffs of 0.04 and 0.05, based on averaged model importance, a proposed elbow rule, and a predictive-parsimony check. The underlying indicator-code mapping, feature scores, and cross-validation outputs were unavailable for independent verification. The reported cutoffs and their claimed predictive accuracy therefore require verification.

(1) Scores of High-Quality Development Indicators

A pronounced “fault-line” pattern appears at 0.04, with G2, G7, G9, G19, and G20 lying above this cutoff and exhibiting substantially higher importance for predicting G1.

To enhance predictive efficiency without sacrificing key information, only indicators with scores > 0.04 were retained for subsequent prediction within the high-quality development dimension.

(2) Scores of New Quality Productivity Indicators

The stated X3 score and the 0.05 cutoff require verification against the underlying X-category data.

(3) Visual Analysis

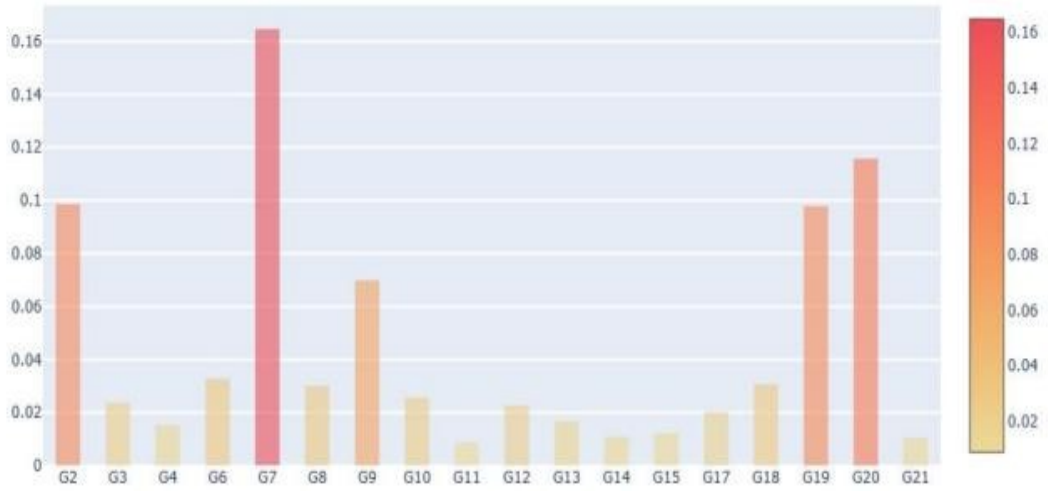


Figure 3. Score Situation of High-Quality Development Indicators

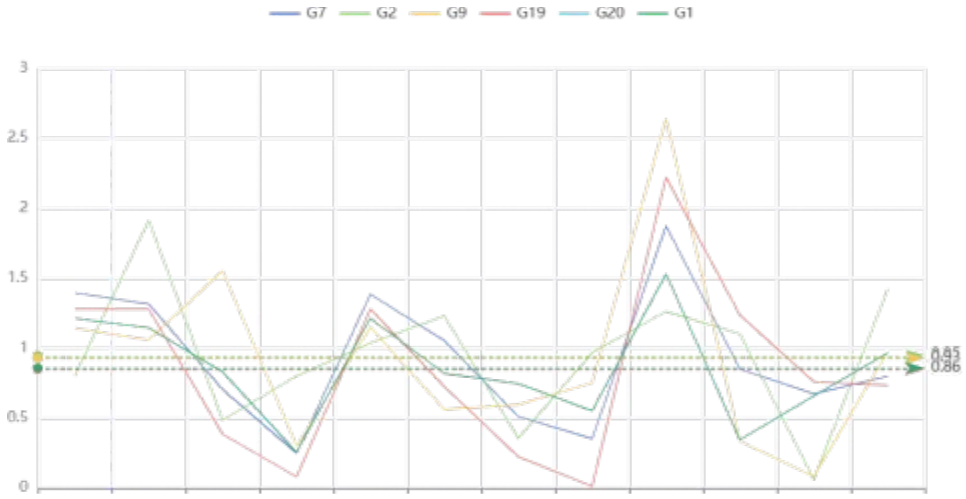


Figure 4. Distribution of Key G-Category Indicators

Using 2022 data for eastern coastal provinces (excluding Taiwan due to data unavailability), we visualize the benchmark indicator G1 together with the retained high-importance G-indicators. The plot suggests clear cross-province dispersion and clustering, with Shandong exhibiting the strongest overall performance and Fujian the weakest. The distribution of the key G-category indicators is shown in Figure 4.

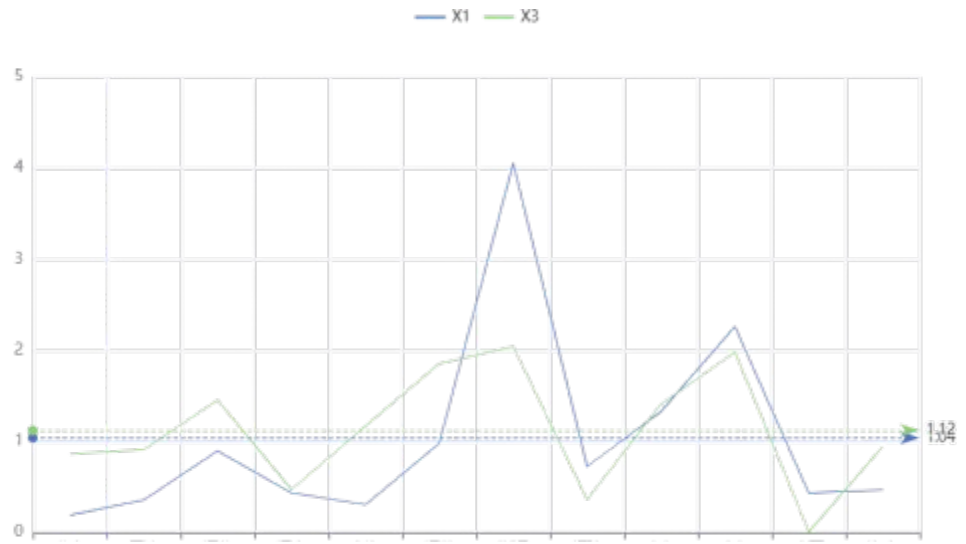


Figure 5. Distribution of Key X-Category Indicators

Analogously, the visualization based on X1 and the key X-indicator(s) indicates that Zhejiang performs best and Guangxi performs worst, while cross-indicator dispersion is relatively modest. The distribution of the key X-category indicators is shown in Figure 5.

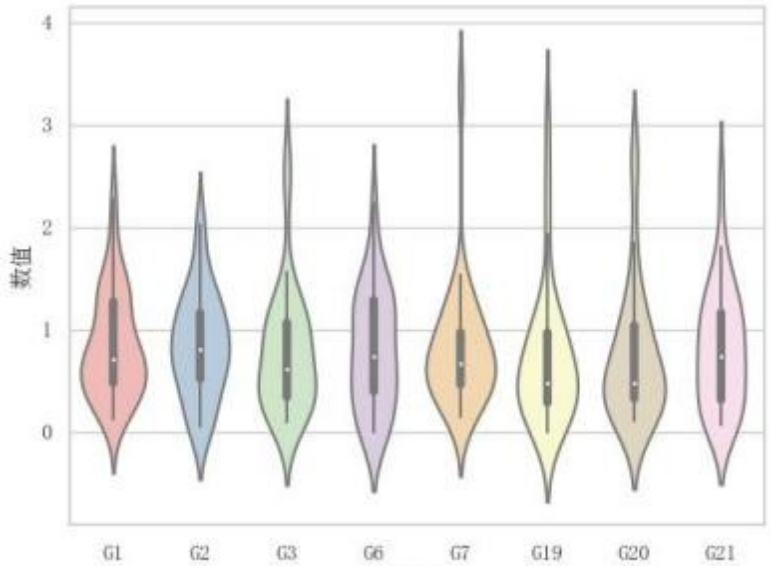


Figure 6. Distribution of Various G-Class Indicators

Most G-indicators lie within a relatively narrow distribution band (roughly 0 to 1.5), while G7, G19, and G20 show longer tails, indicating greater dispersion and potential outliers—consistent with uneven provincial development in these sub-dimensions. These G-class indicator distributions are shown in Figure 6.

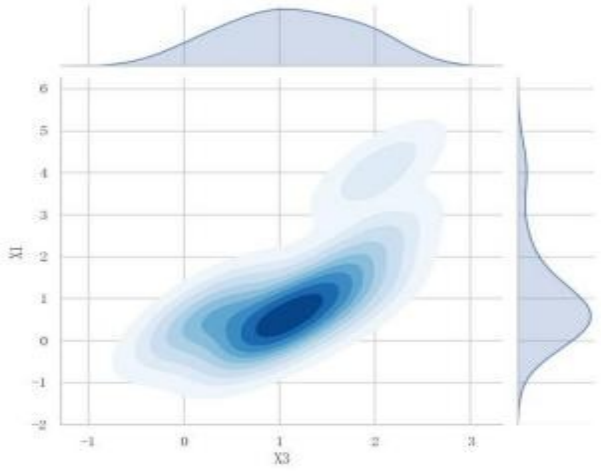


Figure 7. Distribution of the X3 Indicator and the X1 Indicator

X3 appears relatively balanced across provinces, whereas X1 is concentrated at lower levels for most provinces, with only a small subset exhibiting high performance—suggesting that "Taobao village" development remains spatially uneven. Figure 7 compares the distributions of X3 and X1.

(4) Provincial Scores (2020–2022)

Provincial scores for high-quality development, new quality productivity, and the comprehensive level were computed by combining indicator values with their learned scores/weights, producing the score profiles shown in Figures 8–11.

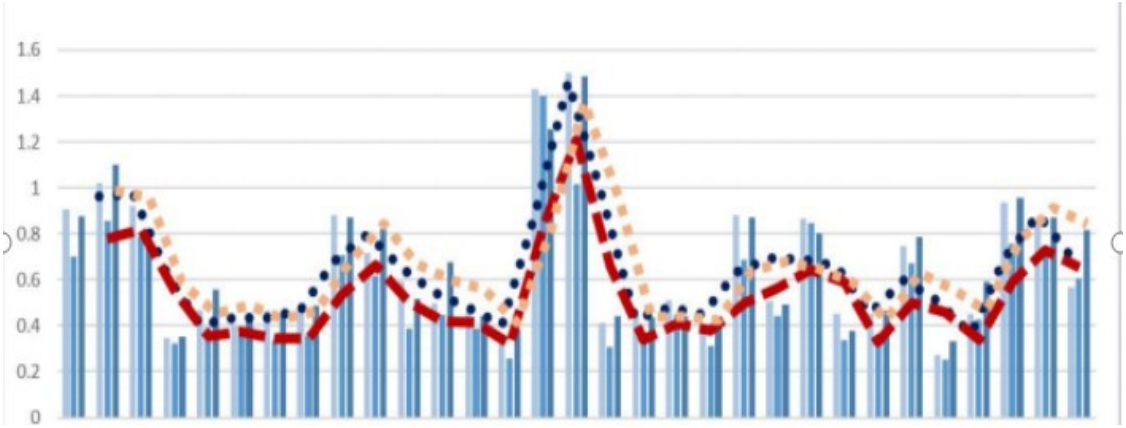


Figure 8. High-Quality Development Scores by Province and Year

High-quality development displays noticeable spatial patterning. The moving-average trend suggests the highest overall level in 2020 and the weakest in 2021, plausibly reflecting widespread disruption to rural production activities during the COVID-19 period.

New quality productivity is highest in major southeast coastal provinces/municipalities (e.g., Jiangsu, Guangdong, Zhejiang, Shanghai) and Beijing, aligning with their strong e-commerce ecosystems and reform-oriented trade intensity.

Temporal variation is also evident: some provinces improve steadily, while others experience sharp declines following the epidemic shock.

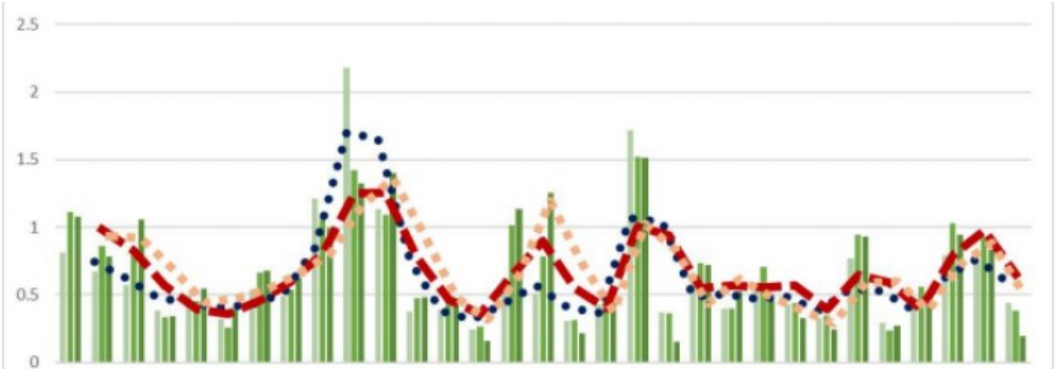


Figure 9. Annual New Quality Productivity Scores of Each Province

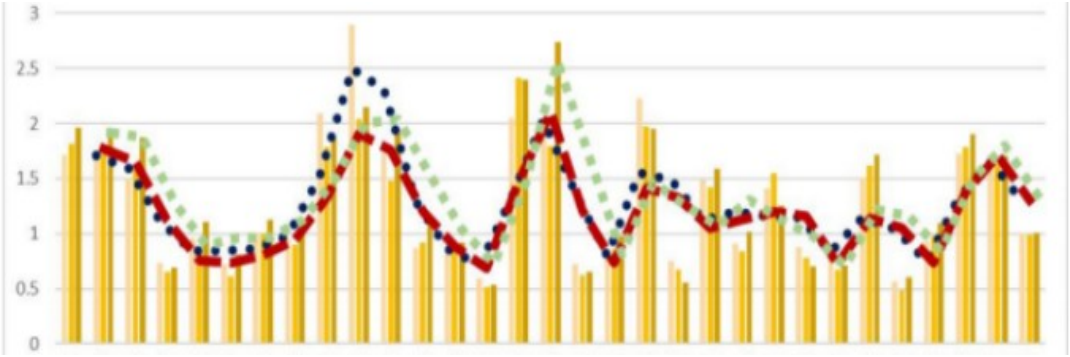


Figure 10. Comprehensive Scores of Each Province in Each Year

The comprehensive trend line indicates heterogeneous epidemic impacts across provinces; for example, Jiangsu shows a persistent decline, while Henan exhibits a sustained rise in the comprehensive score. Table 2 summarizes the reported provincial rankings.

Table 2. Comparison of Comprehensive Scores by Province and Year

Year	**Top 5 provinces/municipalities**	**Bottom 5 provinces/municipalities**
2020	Jiangsu; Guangdong; Shanghai; Shandong; Henan	Shaanxi; Jiangxi; Hubei; Liaoning; Shanxi
2021	Shandong; Jiangsu; Guangdong; Beijing; Henan	Shaanxi; Jiangxi; Liaoning; Hubei; Shanxi
2022	Henan; Shandong; Jiangsu; Beijing; Guangdong	Jiangxi; Guangxi; Shaanxi; Hubei; Shanxi

Across 2020–2022, Jiangsu, Guangdong, Shandong, and Henan appear repeatedly among the top performers, while Shaanxi, Jiangxi, Hubei, and Shanxi repeatedly appear among the bottom performers.

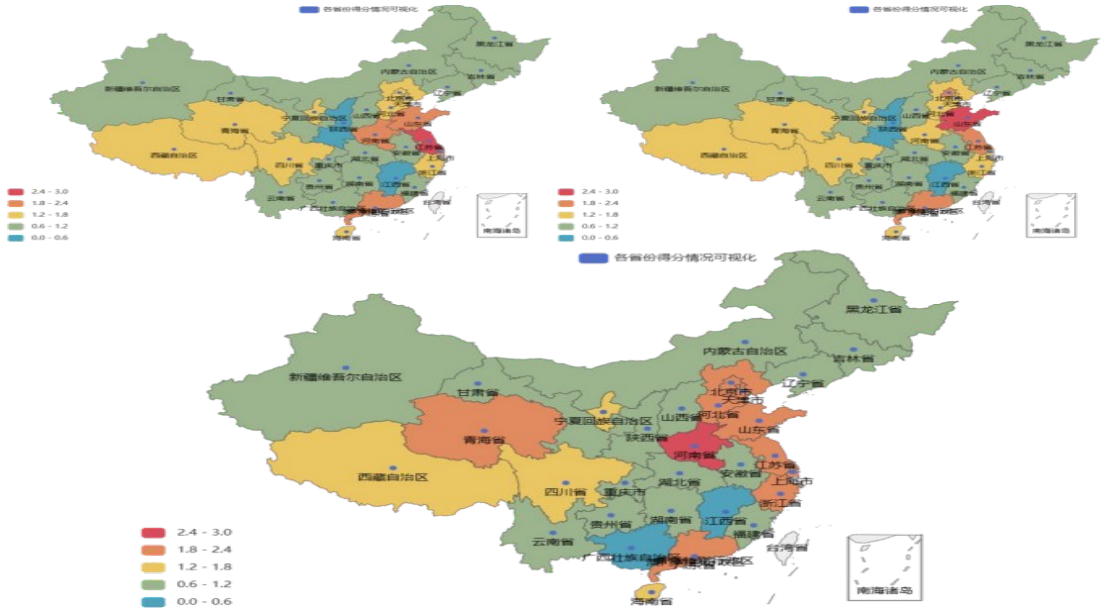


Figure 11. Provincial Distribution of Comprehensive Scores

The spatial visualization further suggests that provinces in the more economically developed southeastern coastal belt tend to exhibit higher comprehensive scores.

5. Analysis and Suggestions on the Obstacles to the Development of Agricultural Modernization in Provincial Regions

5.1 Province Clustering Based on K-means

Building on the composite evaluation results reported in the previous section, this study further investigates the binding constraints that impede agricultural modernization in relatively low-performing provincial regions. To improve interpretability and policy relevance, we focus on provinces whose composite score in 2022 is below 1.2, because provinces above this threshold are more likely to have entered a comparatively stable stage (or have already formed development models suited to local conditions). Given space constraints, the analysis is conducted for 2022 only.

To identify provinces with similar developmental bottlenecks, K-means clustering was applied to the selected provinces, and the silhouette coefficient was used to determine the optimal number of clusters. The clustering outcomes are summarized in Table 3.

Table 3. Cluster Results (2022, Composite Score < 1.2)

Cluster 1	Cluster 2	Cluster 3
Shanxi	Fujian	Guizhou
Inner Mongolia	Jiangxi	Yunnan
Liaoning	Hubei	Shaanxi
Jilin	Hunan	Gansu
Heilongjiang	Guangxi	
Anhui	Chongqing	

Next, Pareto distribution maps were produced for each cluster to examine indicator importance from the perspectives of frequency and cumulative contribution, as shown in Figures 12–14. In the Pareto chart, the left y-axis denotes indicator frequency, while the right y-axis represents cumulative contribution. Following the "critical few" logic, indicators with cumulative contribution below 40% are treated as hindering indicators (i.e., the most salient obstacles), whereas those at 40% or above are treated as development indicators.

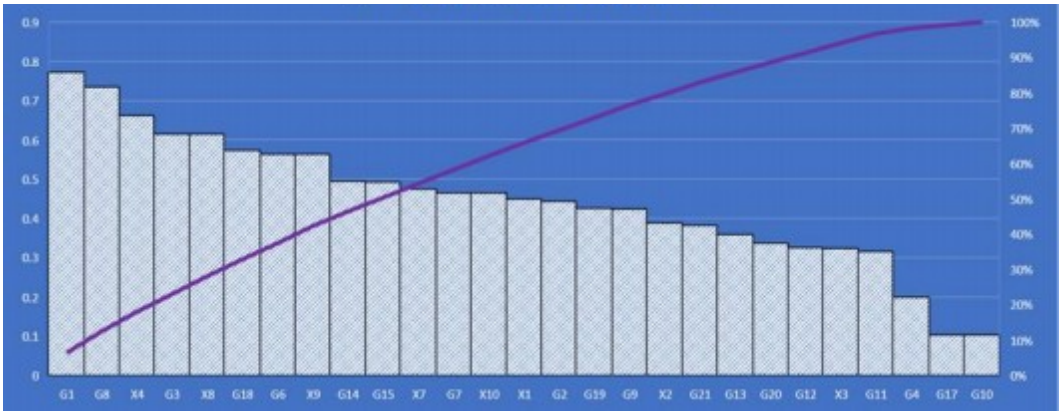


Figure 12. Pareto Distribution of Indicators (Category 1)

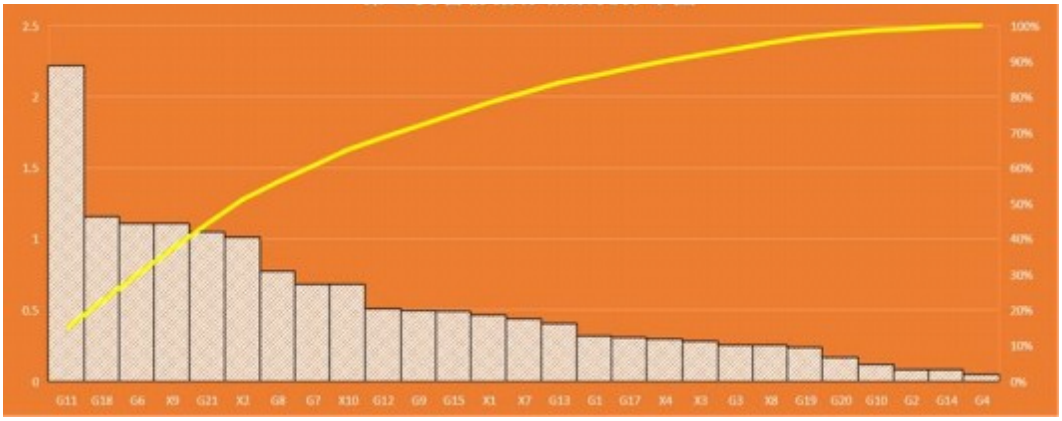


Figure 13. Pareto Distribution of Indicators (Category 2)

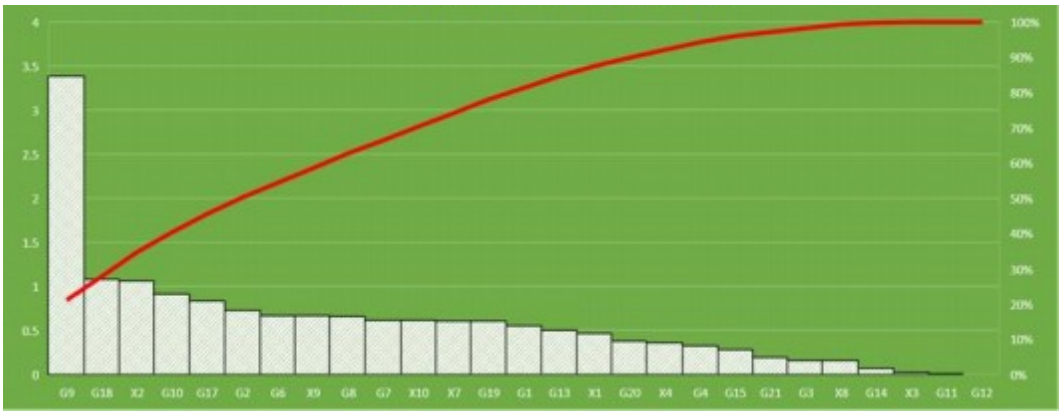


Figure 14. Pareto Distribution of Indicators (Category 3)

5.2 Analysis and Suggestions on the Degree of Obstruction for the First Type of Provinces

This section focuses on Cluster 1 provinces (Shanxi, Inner Mongolia, Liaoning, Jilin, Heilongjiang, and Anhui). We first map each hindering indicator to its corresponding objective layer, Tier-1 indicator, Tier-2 indicator, and specific indicator, thereby forming a structured basis for targeted policy design. The mapping is reported in Table 4.

Table 4. Indicator Correspondence Table (Cluster 1)

Objective Layer	Tier-1 Indicator	Tier-2 Indicator	Specific Indicators	Hindrance indicator
High-Quality Development	Green Development	Green Production	Pesticide Usage Agricultural Producer Price Index	G4 G10
	Openness	Rural Labor Force Cross-border E-commerce Infrastructure	Rural Employment Share in Primary Industry Number of Agricultural Workers Number of Cross-border E-commerce Pilot Zones Length of Paved Roads	G11 G12 G13 G17
	Shared Benefits	Social Security	Number of Rural Doctors Disposable Income-to-Consumption Ratio of Rural Residents	G20 G21
New Quality Productive Forces	Digital Rural Development	Digital Rural Development Environment	Peking University Digital Inclusive Finance Index Rural Electricity Consumption	X1 X3

From Table 4, the dominant constraints for Cluster 1 provinces concentrate on four dimensions: green development, openness, shared benefits, and digital rural development. Based on the indicator meanings and the broader policy context, the following targeted recommendations are proposed.

(1) Green development

Improve fertilizer and pesticide use efficiency by upgrading agricultural machinery and promoting precision application aligned with crop growth stages; precision irrigation can further reduce nutrient loss.

Increase the share of organic fertilizers and promote integrated organic–chemical fertilization, supported by extension services that communicate the risks of excessive chemical use and the agronomic benefits of organic inputs.

Develop crop-specific fertilization formulas based on soil testing, and scale soil-test-based fertilization to guide field practice.

(2) Openness

Cluster 1 provinces face openness-related constraints associated with rural labor supply and structural employment dependence. Recommended actions include:

Expand farmers' income channels by improving the rural operational system and scaling diversified new agricultural business entities, while strengthening the dissemination of pro-farmer policies.

Accelerate land transfer and moderate-scale operation, using mechanization and improved farming techniques to offset labor shortages.

Strengthen local talent pipelines, including incentives for college graduates to work in (or return to) their hometowns, accompanied by improvements in supporting infrastructure and rural service capacity.

Deepen the coverage and usage of digital inclusive finance to improve household financial access and help narrow income disparities among rural laborers (Liu & Han, 2024).

(3) Sharing level

Constraints at the sharing level are linked to infrastructure, rural public services, and social security, which jointly shape residents' welfare and the inclusiveness of modernization gains. Thus:

In line with local specialty industries, scale livestreaming and e-commerce–based agricultural assistance, using digital marketing to strengthen rural industrial resilience.

Increase investment in rural public services (healthcare, education, and basic health), and use policy guidance and fiscal support to strengthen living for rural residents and laborers (Liu, Meng, & Niu, 2024).

Improve bottom-line social protection capacity and strengthen the governance of poverty-prevention mechanisms to reduce the risk of relapse into poverty.

(4) Digital rural development

Digital rural development is hindered by the digital infrastructure and usage environment. Accordingly:

Expand rural broadband and optical-fiber coverage and improve network access capacity, accelerating progress toward near-universal rural connectivity.

Promote agricultural digitalization and adoption of basic digital tools, enabling rural e-commerce and improving the efficiency of rural production–circulation–marketing chains.

In summary, the obstacle diagnosis suggests that lagging provinces are constrained by a small set of "critical" factors rather than by uniformly weak performance across all dimensions. Therefore, policy design should shift from generic capacity building to cluster-specific prioritization, focusing resources on the most binding constraints identified by the Pareto-based obstacle analysis. This "measurement–classification–obstacle identification–targeted recommendation" pipeline not only strengthens the interpretability of the provincial benchmarking results but also improves the actionability of policy interventions. The next chapter concludes by summarizing the study's methodological, empirical, and policy contributions, and by outlining limitations and future research directions.

6. Conclusion and Policy Implications

Contributions. This study makes three contributions. First, methodologically, it constructs an integrated provincial evaluation-and-diagnosis pipeline that combines indicator-system design, correlation-based screening, and machine-learning-based scoring and prediction. Second, empirically, it benchmarks interprovincial heterogeneity in agricultural modernization under the dual lenses of high-quality development and new quality productivity. Third, from a policy perspective, it translates composite scores into cluster-specific obstacle profiles that support targeted intervention prioritization.

This paper develops a provincial assessment framework integrating high-quality development and new quality productivity. The reported 2020–2022 analysis includes indicator scoring, spatial comparison, clustering, obstacle diagnosis, and a stacking ensemble. The data and predictive comparisons require validation. Overall, the

proposed framework forms a replicable closed loop—measurement, classification, obstacle identification, and targeted recommendation—that can be used for interprovincial benchmarking and policy sequencing.

6.1 Main findings

First, the composite evaluation results reveal substantial spatial heterogeneity in provincial agricultural modernization. Provinces in the more developed coastal belt generally achieve higher comprehensive scores, whereas several inland provinces remain persistently lagging. The inter-year comparison across 2020–2022 also suggests that the epidemic period coincided with observable performance fluctuations, indicating that modernization quality and digital productivity can be sensitive to shocks in rural production and circulation systems.

Second, the correlation diagnostics show that redundancy differs across dimensions: high-quality development indicators are mostly weakly correlated, while new quality productivity indicators exhibit stronger coupling among digital infrastructure and e-commerce variables. This justifies screening highly correlated indicators before model-based scoring, improving parsimony and mitigating multicollinearity risks in subsequent modeling.

Third, the clustering and obstacle-degree diagnosis provide a policy-actionable view of lagging provinces. For provinces with lower comprehensive scores, constraints do not appear as a uniform "overall weakness" but as development types with distinct bottleneck structures. The "critical few" obstacles concentrate in areas such as green production efficiency, openness-related labor and infrastructure conditions, shared public services, and the digital rural environment. This implies that one-size-fits-all interventions are unlikely to be efficient; targeted policies should be matched to cluster-specific binding constraints.

Fourth, the article presents stacking as a potential prediction approach. The RMSLE, MSE, and MAE comparisons needed to establish its relative performance are not reported. A claim of superior prediction requires the original results.

6.2 Policy implications

The following policy implications are based on reported patterns and require validation. They consider provincial differences, cluster-specific obstacles, and potential data-driven monitoring.

The findings suggest that policy design should prioritize precision governance rather than average upgrading. For green development constraints, reducing pesticide and fertilizer intensity requires integrated measures that combine soil testing, precision application, and extension services. For openness constraints, policies should jointly address labor structure, land transfer, mechanization substitution, and inclusive finance to expand income channels and improve factor mobility (Liu & Han, 2024). For sharing constraints, strengthening rural public services—especially healthcare capacity, including rural doctors—remains essential to ensure inclusive modernization gains (Liu, Meng, & Niu, 2024). For digital rural development constraints, infrastructure expansion should be coupled with adoption-oriented interventions (training, platform access, and digital business models) to translate connectivity into productivity. Implementation can adopt a "sequence-and-monitor" approach: (i) prioritize interventions targeting each cluster's top obstacle indicators (the "critical few" below 40% cumulative contribution); (ii) align resources with these obstacles (precision input reduction for green production; labor mobility/land transfer and mechanization for openness; rural healthcare staffing and service provision for sharing; broadband/optical fiber plus adoption training for digital rural development); (iii) consider the stacking ensemble, subject to validation, to periodically update G1 and X1 and evaluate whether bottlenecks are relaxing.

6.3 Limitations and future research

This study has several limitations. First, the analysis focuses on 2020–2022 and may not fully capture longer-run structural changes. Future research could extend the time span and evaluate dynamic transitions before and after major policy nodes. Second, although multi-source data were combined, indicators may still be affected by reporting differences across provinces. More granular micro-data (e.g., county-level or household/firm-level evidence) would help validate mechanisms and improve causal interpretability. Third, the current obstacle diagnosis adopts a Pareto-style threshold to identify key constraints. Future work could compare alternative diagnostic rules and explicitly incorporate spatial dependence to account for regional spillovers and policy diffusion.

In sum, this paper provides an integrated measurement and diagnostic toolkit for provincial agricultural modernization under the dual strategic orientation of high-quality development and new quality productivity. By combining indicator-system construction, machine-learning-based scoring, typology identification, and obstacle diagnosis, the proposed framework supports evidence-based prioritization and can be adapted to other regions or sectors facing similar modernization transitions.

Acknowledgement

The authors are grateful to the editorial team and to three anonymous reviewers, whose detailed comments sharpened the measurement framework and the policy discussion. Thanks are also due to Beijing Forestry University for access to the statistical databases and computing facilities used in this study.

Author Contributions

Conceptualization, S.K.W. and B.Z.; Methodology, S.K.W. and B.Z.; Formal Analysis, S.K.W. and B.Z.; Investigation, S.K.W.; Data Curation, S.K.W.; Software, S.K.W.; Validation, S.K.W. and B.Z.; Resources, S.K.W. and B.Z.; Visualization, S.K.W.; Writing – Original Draft, S.K.W.; Writing – Review & Editing, S.K.W. and B.Z.; Supervision, B.Z.; Project Administration, S.K.W. and B.Z. We have read and approved the final manuscript and take responsibility for the work.

Funding

Not applicable.

Conflicts of Interest

The authors declare no conflict of interest.

Ethical Approval and Consent

Not applicable.

Data Availability Statement

The indicator data analysed in this study are compiled from publicly available statistical yearbooks, namely the China Statistical Yearbook, the China Rural Statistical Yearbook, the China Science and Technology Statistical Yearbook, and the China Environmental Statistical Yearbook (2020–2022), together with openly accessible third-party reports. The compiled provincial dataset and the calculation code are available from the corresponding author on reasonable request.

Declaration of Generative AI Use

The authors did not use generative AI tools in the preparation of this manuscript.

References

Breiman, L. (2001). Random forests. Machine Learning, 45, 5–32. https://doi.org/10.1023/A:1010933404324

Cao, Q., Chen, R., & Ma, N. (2024). Measurement and regional differences of new-quality productive forces in China's urban industry. Lingnan Journal, (3), 98–113.

Central Committee of the Communist Party of China, & State Council of the People's Republic of China. (2021, February 21). Opinions on comprehensively advancing rural revitalization and accelerating the modernization of agriculture and rural areas. Chinese Government Network. https://www.gov.cn/zhengce/2021-02/21/content_5588098.htm

Du, Z., Luo, Q., & Yang, X. (2021). Connotations, characteristics, development constraints, and implementation paths of high-quality agricultural development: A literature review. Journal of the Management Cadre College of the Ministry of Agriculture and Rural Affairs, (4), 14–25.

Fang, K., Wu, J., Zhu, J., & Xie, B. (2011). A review of random forest research. Statistics & Information Forum, 26(3), 32–38.

Fu, L., & Cai, Y. (2024). New quality productive forces in China's prefecture-level cities: Temporal evolution, group characteristics, and development strategies. Review of Industrial Economics, (4), 5–22.https://doi.org/10.19313/j.cnki.cn10-1223/f.20240428.001

Han, W., Zhang, R., & Zhao, F. (2024). Measuring new-quality productive forces and the new drivers of China's economic growth. The Journal of Quantitative & Technical Economics, 41(6), 5–25. https://doi.org/10.13653/j.cnki.jqte.20240418.001

Hao, X., & Shen, Z. (2022). Evaluation of the effectiveness of rural e-commerce in empowering rural revitalization and analysis of obstacle factors. Journal of Xi'an University of Finance and Economics, 35(5), 40–52.

He, H., Song, L., & Li, G. (2017). Spatiotemporal differences in the quality of China's agricultural economic growth. Economist, (7), 87–97.

Hou, G., Zhang, Z., & Dong, S. (2024). New-quality productive forces empowering high-quality agricultural development in Northeast China: Theoretical logic, key issues, and practical pathways. Hunan Social Sciences, (1), 69–76.

Huang, P. (2007). The course and innovative development of agricultural modernization in China. Research of Agricultural Modernization, 28(2), 129–134.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. In Advances in Neural Information Processing Systems 30 (pp. 3146–3154).

Li, Q. (2024, March 5). Report on the work of the government: Delivered at the second session of the 14th National People's Congress. Chinese Government Network. https://www.gov.cn/yaowen/liebiao/202403/content_6939153.htm

Lin, W., & Dong, X. (2024). Reflections on building an agricultural powerhouse led by new-quality productive forces. Journal of Nanjing Agricultural University (Social Sciences Edition), 24(3), 18–27. https://doi.org/10.19714/j.cnki.1671-7465.20240430.001

Liu, G., Meng, H., & Niu, H. (2024). Practical constraints and policy choices: The role of the state in addressing rural social security difficulties in the new era. Rural Economy, (2), 103–113.

Liu, X., & Han, Q. (2024). Digital inclusive finance and rural labour wage income: An empirical analysis from the perspectives of total income and income inequality. Shandong Social Sciences, (1), 183–192.

National Development and Reform Commission. (2024, March 15). Developing new-quality productive forces is an intrinsic requirement for promoting high-quality development. https://www.ndrc.gov.cn/wsdwhfz/202403/t20240315_1364971.html

Pearson, K. (1895). Note on regression and inheritance in the case of two parents. Proceedings of the Royal Society of London, 58, 240–242. https://doi.org/10.1098/rspl.1895.0041

Peng, S., & Wu, X. (2023). Coupling and coordination between water-resources carrying capacity and high-quality agricultural development in China. Yangtze River, 54(10), 52–59. https://doi.org/10.16232/j.cnki.1001-4179.2023.10.008

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: Unbiased boosting with categorical features. In Advances in Neural Information Processing Systems 31 (pp. 6638–6648).

Qi, R., & Li, F. (2024). The study on empowering high-quality development of the cultural and tourism industry with new-quality productive forces. Journal of Yantai University (Philosophy and Social Science Edition), 37(5), 112–124. https://doi.org/10.13951/j.cnki.issn1002-3194.2024.0411.02

Research Group of the Rural Economy Research Department, Development Research Center of the State Council, Ye, X., & Cheng, Y. (2021). Connotations, characteristics, and an evaluation system of agricultural and rural modernization in the new development stage. Reform, (9), 1–15.

Shi, J., & Zhang, J. (2019). Load forecasting based on a multi-model fusion stacking ensemble-learning approach. Proceedings of the CSEE, 39(14), 4032–4042. https://doi.org/10.13334/j.0258-8013.pcsee.181510

Song, D., & Lyu, M. (2024). Dialectical reflections on new-quality productive forces and the high-quality development of state-owned enterprises. Social Sciences Front, (5), 256–261.

Wang, J., Shi, M., & Li, Z. (2023). Digital economy promotes high-quality agricultural development: Internal mechanisms and empirical evidence. Journal of Xi'an University of Finance and Economics, 36(4), 78–89. https://doi.org/10.3969/j.issn.1672-2817.2023.04.007

Wang, Q., & Yang, J. (2023). Digital new-quality productive forces and the high-quality development of agriculture in China. Journal of Shaanxi Normal University (Philosophy and Social Sciences Edition), 52(6), 61–72.

Wolpert, D. H. (1992). Stacked generalization. Neural Networks, 5(2), 241–259. https://doi.org/10.1016/S0893-6080(05)80023-1

Xi, J. (2023, March 16). Accelerating the construction of an agricultural powerhouse and advancing the modernization of agriculture and rural areas. Qiushi, (6). http://www.qstheory.cn/dukan/qs/2023-03/15/c_1129431619.htm

Xiao, L., & Zhang, Y. (2023). Spatiotemporal characteristics and influencing factors of high-quality agricultural development in the Northeast China urban agglomeration. Journal of South China Agricultural University (Social Science Edition), 22(6), 74–87. https://doi.org/10.7671/j.issn.1672-0202.2023.06.007

Xin, L., & An, X. (2019). Construction and measurement of an evaluation system for high-quality agricultural development in China. Economic Review Journal, (5), 109–118. https://doi.org/10.16528/j.cnki.22-1054/f.201905109

Xu, J., & Yang, Y. (2018). Ensemble learning methods: A review. Journal of Yunnan University (Natural Sciences Edition), 40(6), 1082–1092. https://doi.org/10.7540/j.ynu.20180455

Xu, Z., Liu, K., & Lin, Z. (2024). Evaluation of the level of high-quality agricultural development in the new era: A supply-and-demand perspective. Journal of Dalian University of Technology (Social Sciences), 45(1), 52–61. https://doi.org/10.19525/j.issn1008-407x.2024.01.006

Xu, Z., Zheng, L., & Cheng, M. (2023). The internal logic and practical conception of new-quality productive forces empowering high-quality development. Contemporary Economic Research, (11), 51–58.

Zhang, H., & Wang, L. (2023). Measurement and advancement pathways of digital rural development in Western China. East China Economic Management, 37(11), 70–78.

Zheng, J. (2023). Promoting agricultural modernization through new-quality productive forces: Theoretical logic and development pathways. Price: Theory & Practice, (11), 31–35. https://doi.org/10.19851/j.cnki.CN11-1010/F.2023.11.429

Zhou, Z.-H. (2012). Ensemble methods: Foundations and algorithms. Chapman & Hall/CRC.