CITESCORE 3.5 SCOPUS

06,2026
Volume 4 Issue 2

Agricultural & Rural Studies

EDITOR-IN-CHIEF:JUNBIAO ZHANG

ISSN 2959-9784

SCC PRESS

https://sccpress.com/ars

OPEN ACCESS

ISSN 2959-9784

Cover Story

In Dali, Yunnan, the walnut (*Juglans sigiIlata Dode*) is more than a crop; it is a web of life. With 370,000 farming households and an average planting area of 5.2 mu per farmer, these figures outline the substantial foundation of this pillar ecological industry. Among the mountains of Yangbi, a walnut-based agro-forestry-pastoral system that has endured for a millennium is demonstrating, through the complete logic of a "socio-ecological system," how agriculture can build multiple forms of resilience in its encounter with modernity. It stands as a key case study, significant both in scale and theoretical depth.

The core of this system lies in its inherent structure of resilience, manifested in three mutually supportive dimensions: ecology, livelihoods, and socio-culture. Ecologically, the three-dimensional biological network formed by walnut trees, understory crops, and livestock not only achieves soil and water conservation and carbon sequestration but also creates a "natural buffer pool" against climate fluctuations—an adaptive mechanism based on biodiversity. In terms of livelihoods, diversified outputs provide smallholder families with a risk-dispersed livelihood portfolio, granting them crucial flexibility when facing market volatility and natural risks. Socio-culturally, intergenerationally transmitted local ecological knowledge continuously guides the system's adaptive management, forming a form of "place-based wisdom" with low external dependency that sustains community identity.

The resilience of this system is not static; it has achieved structural innovation through integration with modern elements. Its contemporary evolution is revealed as a stable triangular support framework of "organization-market-technology." First, farmers' cooperatives enhance the collective bargaining power of smallholders through organization and standardization. Second, the "Yangbi Walnut" National Geographical Indication (GI) certification successfully transforms a local ecological product into a high-value commodity carrying a unique narrative of terroir, achieving value capture in specific markets. Particularly crucial is the role of the Science and Technology Commissioner system, which acts as a "technology graft" and "knowledge bridge." The precision cultivation techniques they promote directly address industry bottlenecks, and increased yield data from demonstration bases empirically proves that this system can inject iterative innovative momentum into industrial resilience.

The profundity of the Yangbi case lies in its empirical validation of a rural development paradigm distinct from the industrialized, homogenized path. It demonstrates that socio-ecological systems based on ecological principles, smallholder agency, and cultural continuity possess not only historical legitimacy but also practical vitality in confronting future complex challenges. This offers a significant insight for the global sustainable transformation of agriculture: there is an urgent need to shift from seeking external substitutes towards identifying, empowering, and integrating the endogenous resilience capital embedded within local systems. This is how to establish an organic, mutually supportive connection between ecological security, livelihood assurance, and industrial revitalization. Thus, the walnut forests of Yangbi transcend being merely an economic plantation, becoming a "living laboratory of ideas" concerning the future of agriculture, community resilience, and how humans and nature can coexist sustainably.

(Xueliang Xi, Research Fellow, Yunnan Academy of Forestry and Grassland, China)

Agricultural & Rural Studies

Vol. 4, No. 2 June, 2026

Contents

About the Journal

Agricultural & Rural Studies (**A&R**, **ISSN 2959-9784**) is an exclusively digital, open-access journal dedicated to advancing interdisciplinary scholarship at the critical nexus of agricultural sustainability, rural revitalization, and farmer well-being. Published quarterly, **A&R** features a range of content types—including original research, reviews, perspectives, and commentaries—serving as a professional and innovative platform for rigorous academic dialogue and global knowledge dissemination.

Core Features:

(1) **Interdisciplinary Scope:** A&R centers on interdisciplinary and multidisciplinary inquiry, integrating insights from economics, sociology, human geography, and cognate disciplines to address pressing contemporary issues in agriculture, rural development, and farmer-related policies—bridging theoretical frameworks to advance holistic understandings of these dynamic systems.

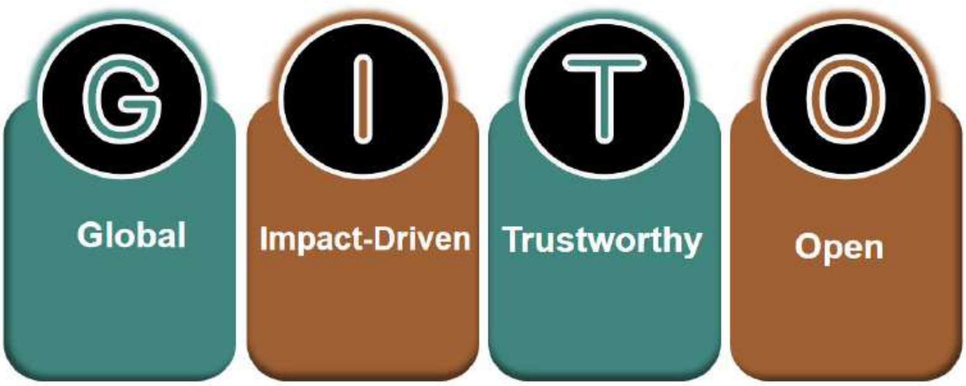
(2) **Global Editorial Excellence:** The journal is guided by a distinguished, highly specialized, and globally diverse editorial board, ensuring rigorous peer review, cultural sensitivity, and alignment with international scholarly standards.

(3) **Culture-Driven Operations:** Rooted in our mission, vision, and values, cultural stewardship lies at the core of A&R's identity. Our unwavering commitment to ethical principles, community engagement, and purposeful impact permeates every facet of the journal's operations.

(4) **Sustainable Non-Profit Model:** No Article Publishing Charges (APC) for the first four years. Beyond this, it will maintain a reasonable level of APC for authors worldwide in the future.

(5) **Fully Independent OA Platform:** As an independent OA publication (https://sccpress.com/ars), A&R prioritizes equitable global access to scholarship while actively promoting your work to a worldwide audience, maximizing its reach and impact.

(6) **Tailored Publication Support:** Committed to fostering a seamless research journey, A&R offers comprehensive, efficient, and professional editorial services. Recognizing the unique demands of each study, our dedicated team provides structured guidance—from refining article frameworks to managing references—simplifying the publication process and empowering authors to focus on their core scholarship.



Agricultural & Rural Studies

Review

Corn Stover: Revisiting of the Opportunities and Barriers of Its Composition, Bioenergy Applications, and Sustainability

Anthony Amotoe-Bondzie [1], Isaac Slaven [1,*], Kelly J. Best [1], Nichole Hugo [1]

1 Eastern Illinois University, Charleston, IL 61920, United States; aamotoebondzie@eiu.edu (A.A.-B.); aamotoebondzie@eiu.edu (K.J.B.); nhugo@eiu.edu (N.H.)
* Correspondence: islaven@eiu.edu

Abstract: Corn Stover (CS), rich in lignocellulosic materials, poses challenges in biodegradability, necessitating pretreatment methods to enhance biofuel yields. Biogasification through anaerobic digestion and bioethanol production via fermentation are promising pathways, though both face technical barriers, particularly in pretreatment efficiency and enzyme limitations. Economic factors, including high feedstock collection and transportation costs, hinder large-scale adoption. The environmental implications of CS removal, such as soil nutrient depletion and erosion risks, highlight the need for balanced residue management practices. Technological advancements, such as improved pretreatment techniques, biomass densification, and co-digestion strategies, have shown potential to enhance process efficiency and reduce costs. However, integrating circular economy principles by valorizing co-products like lignin and digestate further strengthens the sustainability of CS utilization. This review examines the composition of CS, its applications in bioenergy, and the environmental and economic considerations associated with its use. Future research directions emphasize genetic and process innovations to boost biogas and bioethanol yields, scalable industrial applications, and policy frameworks that support large-scale deployment. Ultimately, CS holds significant promise in contributing to global renewable energy goals, provided that technological, economic, and environmental challenges are effectively addressed.

Keywords: biogasification; bioethanol; bioenergy applications; corn stover; lignocellulosic biomass

Check for updates

Citation: Amotoe-Bondzie, A., Slaven, I., Best, K. J., & Hugo, N. (2026). Corn Stover: Revisiting of the Opportunities and Barriers of Its Composition, Bioenergy Applications, and Sustainability. *Agricultural & Rural Studies*, *4*(2), 17. https://doi.org/10.59978/ar04020007

Received: 24 February 2026
Revised: 17 April 2026
Accepted: 27 April 2026
Published: 18 May 2026

1. Introduction

In current societies, the existence of continuous, sustainable, and economic energy is necessary for economic development and growth (Søndergaard et al., 2015). Modern civilization is so dependent on energy that it is impossible to imagine life without it. A disruption or stoppage of its supply would halt the economic machine (Kougias et al., 2014; Labatut et al., 2011). Currently, issues such as hazardous waste materials, the exhaustibility of fossil resources, and increasing energy consumption are subjects of intense research globally (Anal, 2019; Kakihana et al., 2012). These issues make it paramount that it is no longer feasible to rely on existing energy sources (Teng et al., 2014). Moreover, research into new, sustainable sources developed over the past few decades highlights the importance of these concerns and the associated sciences (Lindkvist, 2020).

Global energy policies that promote the inefficient use of fossil fuels have proven environmentally irresponsible because they cause significant environmental damage at local, regional, and global levels. Studies have shown that integrating renewable energy sources into the overall energy mix can mitigate or prevent these negative impacts (Jameel et al., 2024; Tsapekos et al., 2017). While biogas has a long history of recognition, its widespread use has increased primarily over the last century, especially in the last three decades. Biogas, derived from biomass, is particularly useful in rural areas, as it is both inexpensive and locally produced (Fotidis et al., 2014; Y. Li et al., 2013).

The escalating global grain production, which reached 1,030 million tons in 2016, has amplified interest in corn stover (CS) as a renewable energy feedstock (Czajkowski et al., 2019). Despite its promising applications, several challenges hinder the widespread utilization of CS. Residue removal impacts soil health, potentially leading to erosion and reduced soil organic carbon (SOC) sequestration (Mann et al., 2002; Swan et al., 1996). In addition, innovations such as improved techniques for rind and pith components, aim to optimize stover's utility while mitigating environmental concerns (H. Y. Li et al., 2014). Economic considerations, including collection and transportation costs, remain significant barriers to its broader adoption. Given the growing emphasis on renewable energy and sustainable agricultural practices, CS's potential as a bioenergy feedstock warrants a comprehensive assessment. Recent work also indicates that the future competitiveness of lignocellulosic bioenergy will depend not only on improvements in conversion efficiency but also on the integration of digital optimization and broader system-level assessment. Machine-learning approaches are increasingly being used to identify nonlinear interactions across pretreatment, hydrolysis, fermentation, and anaerobic digestion, thereby reducing empirical trial-and-error in process design, while residue-based feedstocks such as corn stover are gaining additional relevance because they can expand bioenergy supply without requiring additional land area. Accordingly, a comprehensive assessment of corn stover bioenergy should consider both artificial intelligence (AI)/ machine learning (ML)-assisted process optimization and the wider socio-economic context, including energy security, carbon-intensity governance, and rural value creation (Farheen et al., 2026; International Energy Agency [IEA], 2023, 2024; Khan et al., 2023; Phromphithak et al., 2021; REN21, 2024; C. Wang et al., 2023). Thus, situating CS within the broader transition toward sustainable energy systems, this study seeks to clarify the key challenges and opportunities shaping its future development.

2. Corn Stover

Corn stover (CS) refers to husks (8%), cobs (15%), leaves (28%), and stalks (48%) left on the farm after harvest. Despite its abundance, only about 6% of corn stovemer is typically collected for use, with the majority left to decompose and enrich soil organic matter (Menardo et al., 2015; Sokhansanj et al., 2002). CS, a typical lignocellulosic biomass, generally consists of 70% cellulose, hemicellulose, and 15–20% lignin, forming a complex and recalcitrant structure (Cui et al., 2012; Guan et al., 2025; Sluiter et al., 2010; Zhong et al., 2021). The heterogeneous composition of corn stover affects its degradability and suitability for various applications (Sokhansanj et al., 2010). About half of the corn plant is corn stover (CS), meaning the mass ratio of corn stover to corn grain is 1:1 (Perlack et al., 2005). CS is the most abundant agricultural residue in the United States (Aghaei et al., 2022).

3. Composition

3.1. Chemical Composition

Lignocellulose is considered an attractive raw material to produce volatile fatty acids (VFAs) through co-digestion with waste-activated sludge (WAS), due to its availability in large quantities and low cost (Guo et al., 2015; A. Zhou et al., 2013). Cellulose is an unsubstituted homopolysaccharide consisting of β-1,4-linked D-glucopyranosyl units. Many cellulose chains aggregate to form crystalline, highly organized microfibrils via extensive inter- and intra-molecular hydrogen bonds, which hinder cellulose degradation. Only amorphous cellulose sections—regions where the ordered structure is lost—are readily accessible to hydrolytic enzymes (Perrot et al., 2022).

Glucuronoarabinoxylan (GAX), the primary matrix polysaccharide in grasses, features a highly O-acetylated linear β-1,4-linked D-xylopyranose backbone, further decorated with pentose, hexose, and (methyl)uronic or hydroxycinnamic acid side chains (Houfani et al., 2020; A. Zhou et al., 2016). This structural heterogeneity impedes enzymatic degradation, necessitating multiple enzymes with specific substrate recognition and cleavage capabilities to efficiently convert GAX.

Lignin, an aromatic heteropolymer composed mainly of guaiacyl, syringyl, and/or p-hydroxyphenyl monolignol subunits, presents the most challenging component for enzymatic depolymerization due to its highly hydrophobic and heterogeneous nature (Gao et al., 2020). Together, these structural and compositional complexities contribute to the recalcitrance of CS lignocellulose to enzymatic degradation, requiring the coordinated action of multiple hydrolytic enzymes.

Recent studies suggest that the presence of cellulose and hemicellulose in pretreated lignocellulosic residues enhances WAS acidification, potentially influencing the composition and metabolic activity of fermentation bacteria (Guo et al., 2015). Moreover, CS hydrolysates contain large amounts of monomeric sugars (e.g., glucose, xylose, and arabinose), making them ideal carbohydrate feedstocks for VFAs production. Furthermore, understanding the contribution of whole CS to VFA recovery from WAS digestion provides a foundation for cost-efficient biomass stabilization and bioenergy recovery in wastewater treatment plants (WWTPs; A. Zhou et al., 2016).

3.2. Nutrient Distribution Characteristics

Studies on CS consistently show strong partitioning of nutrients between grain and vegetative fractions, with grain retaining most nitrogen (N) and phosphorus (P), while stalks, leaves, and husks concentrate substantially higher potassium (K) levels, often nearly threefold greater than grain on a dry matter basis (H. Y. Li et al., 2014; Sawyer & Mallarino, 2014). However, standing-plant measurements frequently overestimate K export because K^+ is highly soluble and subject to intense leaching from vegetative tissues following physiological maturity, particularly after rainfall events between the black layer and grain harvest (Karlen et al., 2015; Sawyer, n.d.; Sawyer & Mallarino, 2012). In contrast, phosphorus concentrations remain comparatively stable over this period (Oltmans & Mallarino, 2011).

Field surveys in Iowa indicate that stover collected at combine harvest typically contains approximately 3 lb P_2O_5 and 19 lb K_2O per dry ton, although observed ranges are wide due to site conditions and weather variability, with potassium occasionally exceeding 40 lb K_2O per ton (Darr et al., 2014; Oltmans & Mallarino, 2011). Nutrient concentrations measured from baled stover often show even greater dispersion, reflecting biological heterogeneity and soil contamination during collection, which can artificially inflate apparent mineral content, particularly for K and trace elements (Darr et al., 2014; Karlen et al., 2015).

4. Bioenergy Applications

Bioenergy encompasses energy production methods that utilize biomass, renewable, plant-derived organic matter. This includes dedicated energy crops, agricultural residues, forestry waste, aquatic vegetation, and municipal waste. Certain characteristics make some biomass more suitable for energy production, such as energy density, moisture content, chemical composition, particle size, production rate, and sustainable availability (Zych, 2008). Among crop residues, CS shows strong potential as a biomass feedstock (Graham et al., 2007).

4.1. Biogasification (Anaerobic Digestion): Process, Efficiency, and Yield

Anaerobic digestion (AD) is a complex biological and physicochemical process that converts organic matter into biogas and digestate in the absence of oxygen (Kangle et al., 2012). Key stages include hydrolysis, acidogenesis, acetogenesis, and methanogenesis. AD can be conducted under two conditions: liquid anaerobic digestion (L-AD, TS < 15%) and solid-state anaerobic digestion (SS-AD, TS > 15%; Ge et al., 2014). SS-AD offers several advantages, such as compost by-products, reduced energy inputs, fewer moving parts, and minimal floating/stratification issues (Brown & Li, 2013). CS has high total solids and contains recalcitrant carbohydrates (J. Zhu et al., 2015). Hence, co-digestion is often required to enhance its degradation. Co-digestion not only improves digestion efficiency and microbial diversity but also sustains microbial growth (El-Mashad & Zhang, 2010). A diverse microbial community is essential for effective anaerobic digestion, and understanding microbial interactions is key to optimizing reactor performance (Liu et al., 2010, 2016; J. Zhou et al., 2013). Additional microbial community exploration can help establish more precise relationships between structure and function.

Co-digestion reduces environmental impact and reliance on fossil fuels (Radwan et al., 1993). But a corn stalk is hard to biodegrade (Myint et al., 2007; Schober & Trösch, 2000; Vedrenne et al., 2008). To enhance lignocellulose hydrolysis, several pretreatment methods have been applied, including size reduction, steam explosion, fungal degradation, ammonification, and alkaline treatments (Bauer et al., 2009; Fernandes et al., 2009). Co-digestion offers promising benefits such as diluting toxic compounds, balancing nutrients, stimulating microbial synergy, and increasing biogas yield (Ağdağ & Sponza, 2007; Parawira et al., 2004, 2008). While research is extensive for co-digestion involving straw and other biosolids, fewer studies focus specifically on corn stalks and other biowaste solids (G. Chen et al., 2010; Parawira et al., 2008; H. Wang et al., 2009).

G. Chen et al. (2010) reported that mono-digestion of corn stalk achieved a maximum methane yield of 217.60 ± 13.87 mL/g TS at an initial TS of 4.8%, with acidification occurring at 6.0% TS and a pH of 5.10 on day 4. Co-digestion improved methane yields by 4.42–58.61% through enhanced VFAs production and better pH regulation. The highest biogas yield (410.30 ± 11.01 mL/g TS) and methane yield (259.35 ± 13.85 mL/g TS) were observed with 40% vermicompost (VC) addition. X-ray diffraction analysis revealed that VC co-digestion reduced corn stalk crystallinity by 29.4%, suggesting improved biodegradability.

Recent studies show that data-driven models (such as the use of AI/ML not as a substitute for process understanding, but as a practical layer for prediction, control, and optimization under highly nonlinear operating conditions) can identify critical operational variables, optimize lignocellulosic biomass-to-manure ratios, and construct soft sensors for instability indicators, suggesting that corn stover digestion can be improved more effectively through multi-parameter, data-assisted control than through isolated one-factor adjustments (Ganeshan et al., 2024; Khan et al., 2023; Sonwai et al., 2023; Zou et al., 2024).

4.2. Bioethanol Production: Pretreatment Strategies and Fermentation Challenges

The introduction of bioethanol into the transport sector was a vital step in reducing fossil fuel dependence. Countries such as the USA, Brazil, and India already blend bioethanol with gasoline

(Morales et al., 2015; Soam et al., 2016; Wojtusik et al., 2016). For sustainable production, lignocellulosic waste, rather than food crops, should serve as the primary feedstock (Agostinho et al., 2015). CS is one of the most abundant lignocellulosic residues, given the global scale of corn production (1.06×10^9 t/year), surpassing wheat and rice (Y. Zhao et al., 2018). However, CS's complex lignocellulosic matrix resists biological conversion (Loow et al., 2016). Pretreatment is thus essential to break down crystalline cellulose and free cellulose and hemicellulose from lignin for enzymatic hydrolysis and fermentation. While downstream steps are well established, pretreatment remains a key research focus (Capolupo & Faraco, 2016; R. Kumar et al., 2016).
Pretreatment methods include:

- **Physical:** Milling, extrusion, microwave.
- **Physicochemical:** Steam explosion, LHW, AFEX, supercritical CO_2.
- **Chemical:** Acid/alkaline treatments, organic solvents, ionic liquids.
- **Biological:** Use of fungi and other microbes (Aditiya et al., 2016).

Enzymatic hydrolysis dominates carbohydrate conversion, though acid and hydrothermal techniques are also used. These technologies vary widely in efficiency, operating conditions, and ethanol yields. Commercial-scale implementations remain limited, with no single method yet preferred (Aditiya et al., 2016; H. Chen & Fu, 2016). Bioethanol is a widely used conventional biofuel and gasoline additive (Balat et al., 2008). It is categorized into four generations, with second-generation ethanol, produced from crop residues like CS, offering a more sustainable path than first-generation sources that compete with food and require high water input. Second-generation ethanol avoids the food-vs-fuel dilemma and utilizes agricultural waste. However, the required pretreatment step adds technical and economic complexity. Innovations are increasingly targeting this challenge (Lászlok et al., 2020; Y. Zhao et al., 2018), helping move second-generation biofuels closer to commercial viability. The EU, for example, aims to source 3.5% of its transport biofuel consumption from advanced biofuels by 2030 (Aghaei et al., 2022). Effective pretreatment must disrupt the crystalline structure of cellulose and separate it from lignin to enable enzymatic hydrolysis. This step is one of the costliest and technically demanding stages of bioethanol production (Shadbahr et al., 2015), accounting for up to 16–19% of a biorefinery's capital costs (Da Costa Sousa et al., 2009). Moreover, enzyme costs can make up 70% of hydrolysis expenses (Baral & Shah, 2017), emphasizing the importance of efficient pretreatment. Pretreatment methods—physical, chemical, biological, or hybrid—must be evaluated based on yield, waste generation, cost, chemical recyclability, and feedstock characteristics (Karimi et al., 2013).

4.2.1. Physical Pretreatment

This aims to reduce particle size and increase the surface area to volume ratio, typically through milling, pyrolysis, or irradiation. Although physical pretreatment alone is usually insufficient, it can be used before, during, or after other treatments. For example, post-chemical pretreatment size reduction offers advantages like lower energy consumption. However, not all processes benefit from this step (Karimi et al., 2013).

4.2.2. Chemical Pretreatment

Chemical pretreatment modifies the structure of biomass using acids, alkalis, solvents, or ionic liquids. Alkaline pretreatment—among the oldest methods—uses agents like NaOH, $Ca(OH)_2$, or Na_2CO_3 to break lignin–carbohydrate bonds. Its effectiveness depends on concentration, temperature, and duration (Karimi et al., 2013; Molaverdi et al., 2019). Mirmohamadsadeghi et al. (2016) achieved a 95% glucose yield with Na_2CO_3 pretreatment. Zheng et al. (2009) developed a wet-state NaOH method that reduced pretreatment time by 86% and used 66.7% less NaOH than the solid-state alternative. Acid pretreatments like steam explosion, LHW, and dilute acid treatments target hemicellulose. Dilute acid pretreatment converts hemicellulose to xylose and then furfural, an inhibitor of fermentation (Mosier et al., 2005). Y. Zhu et al. (2005) showed that preheating to remove moisture before dilute acid treatment improved sugar yields. Um et al. (2003) found sulfuric acid superior to phosphoric acid for this application.

4.2.3. Physicochemical Methods

Steam explosion involves high-pressure, high-temperature treatment followed by rapid depressurization, releasing acetic and uronic acids that autocatalyze hemicellulose breakdown (K. Wang et al., 2015). Catalysts like sulfuric acid or SO_2 can be added, reducing pH to 3–4. Chang et al. (2012) reported substantial reductions in cellulose, hemicellulose, and lignin contents following this treatment. AFEX, a variation of steam explosion using liquid ammonia, is performed under mild temperatures and high pressures. Teymouri et al. (2005) optimized AFEX for CS, achieving nearly 100% glucose and 80% xylose yields.

4.2.4. Emerging Chemical Methods

MgO pretreatment, as shown by Aghaei et al. (2022), neutralizes acetic acid during liquid hot water (LHW) pretreatment, preventing the formation of inhibitors like furfural. Compared to LHW, MgO pretreatment improved hemicellulose recovery, enhanced lignin removal, and increased sugar yields by 6%. Organosolv pretreatment uses organic solvents (e.g., methanol, ethanol) with or without catalysts to fractionate lignocellulose into cellulose, hemicellulose, and lignin. It minimizes carbohydrate degradation and enables easy solvent recovery (X. Zhao et al., 2009). Qing et al. (2017) combined alkaline and organosolv treatments to achieve a 98.6% sugar yield.

4.2.5. Ionic Liquids and DESs

Ionic liquids are highly effective in dissolving biomass, enabling efficient separation of glucan-rich fractions (L. Sun et al., 2013; Uppugundla et al., 2014). However, nutrient supplementation may be required due to losses during pretreatment. Geng and Henderson (2012) achieved a 96% glucose yield by combining mild alkali pretreatment with ionic liquids. Deep eutectic solvents (DESs), a biocompatible class of ionic liquids, consist of hydrogen bond donors and quaternary ammonium salts. Unlike traditional ionic liquids, DESs can be derived from non-ionic components (A. K. Kumar & Sharma, 2017). G.-C. Xu et al.(2016) achieved 99% glucose yield using optimized DESs. Microwave-assisted DES pretreatment, currently under development, reduces processing time significantly (Z. Chen & Wan, 2018).

4.2.6. Biological pretreatment

Biological pretreatment is a safe and environmentally friendly method that utilizes lignocellulose-degrading microorganisms to enhance the digestibility of CS. Various biological strategies, such as fungal treatment, microbial consortia, and enzymatic pretreatment, have been employed as upstream processes in biofuel production from CS (Tabatabaei et al., 2020). Soft- and brown-rot fungi primarily degrade cellulose, while white-rot fungi are more effective at lignin degradation (Singh et al., 2018). These organisms secrete extracellular enzymes, such as lignin peroxidases and laccases, which facilitate the breakdown of lignin and improve the accessibility of cellulose and hemicellulose.

Although biological pretreatment is not energy-intensive (Da Costa Sousa et al., 2009), it is generally time-consuming and requires large-scale infrastructure and equipment (Tabatabaei et al., 2020). Several studies have successfully applied different biological pretreatment strategies to CS (Saha et al., 2016; F.-h. Sun et al., 2011; Wan & Li, 2010). For example, Song et al. (2013) demonstrated biological pretreatment using Irpex lacteus in the presence of manganese ions. Their enhanced method, which involved manganese supplementation, achieved a glucose yield that was 61.39% higher than that of the conventional biological approach.

Combining multiple pretreatment methods can lead to improved yields, though this approach requires a detailed economic evaluation. For instance, combining steam explosion with alkali treatment is more effective than single-method treatments, as each technique targets different structural bonds within lignocellulose (Karimi et al., 2013).

Comparative analyses of pretreatment methods—including steam explosion, AFEX, dilute sulfuric acid, and biological pretreatment—reveal significant differences in cost and efficiency. Biological pretreatment typically requires twice the amount of feedstock compared to diluting sulfuric acid and demands over ten times the capital investment. Operating costs are also roughly double, underscoring the need for further development before this method becomes industrially viable (Baral & Shah, 2017). However, in terms of external energy requirements, biological pretreatment consumes only one-fifth the energy required for ammonia fiber explosion, making it attractive from an energy conservation standpoint.

From an environmental perspective, the use of greener solvents such as deep eutectic solvents (DESs), ionic liquids, and supercritical fluids is gaining attention. Among these, DESs are especially promising due to their environmental friendliness and cost-effectiveness (Roy et al., 2020). Nevertheless, all these emerging methods require additional research to reach industrial scalability.

In a notable life cycle assessment (LCA) study, Smullen et al. (2019) evaluated four pretreatment strategies—NaOH, ammonia, sulfuric acid, and methanol—across multiple environmental categories. Methanol was found to have the lowest global warming potential but the highest eutrophication impact. Overall, ammonia and methanol emerged as the most favorable options in terms of impacts on air, soil, and water. Conversely, sulfuric acid and NaOH had higher impacts on global warming, eutrophication, and photochemical oxidation potential.

4.3. Digital and AI- Assisted Process Optimization in CS Bioconversion

Unlike conventional mechanistic or statistical models, AI-based approaches, especially artificial neural networks (ANNs), show superior capability in modeling complex, multi-dimensional process behavior using real-time and historical datasets (Pereira et al., 2020; Shenbagamuthuraman & Kasianantham, 2023). For instance, ANN-based frameworks have been shown to outperform traditional response surface methodology (RSM) by more accurately capturing non-linear relationships between pretreatment severity, enzymatic hydrolysis efficiency, and inhibitor formation, leading to enhanced sugar recovery and fermentation performance (Shenbagamuthuraman & Kasianantham, 2023).

Consequently, the AI platforms dynamically adjust operational conditions in response to fluctuations in feedstock quality, thereby stabilizing fermentation performance and reducing energy consumption (Pereira et al., 2020). Evidence from enzymatic saccharification and fermentation modeling shows that ANN-based optimization can significantly enhance yield while reducing experimental iterations, highlighting its potential to accelerate process development and scale-up (Gitifar et al., 2013; Sewsynker-Sukai et al., 2017). However, this potential is contingent on the availability of high-quality datasets, as model performance remains sensitive to data variability and limitations inherent in industrial bioenergy systems (Grahovac et al., 2016).

Similarly, in bioethanol production, the most useful role of AI/ML is to connect pretreatment decisions with downstream hydrolysis and fermentation performance rather than to optimize each unit operation in isolation. Studies show that machine-learning models can successfully relate biomass characteristics, solvent or catalyst identity, pretreatment severity, solids recovery, and sugar-release outcomes, which means that the most meaningful optimum for corn stover is not simply maximum cellulose accessibility, but the best overall balance among deconstruction efficiency, inhibitor minimization, fermentability, and process economics (Haldar et al., 2023; Phromphithak et al., 2021; C. Wang et al., 2023).

4.4. Comparison with Other Biomass Feedstocks: Advantages and Limitations

Several authors (Adler et al., 2015; Nelson, 2002; Perlack et al., 2005) have suggested that harvesting CS can be a "win-win" management practice, often stating that CS is an underutilized resource that could be used as a feedstock while simultaneously reducing residue management costs, which currently range from \$49–74 ha^{-1} (Karlen et al., 2015). However, the decision to harvest CS stover for bioenergy or any other use is not that simple, as CS (plant residue) also supports many ecosystem services (Johnson et al., 2009; Wilhelm et al., 2007, 2010).

5. Environmental and Economic Considerations

5.1. Sustainability and Carbon Footprint: Life Cycle Assessment (LCA)

Uncertainties surrounding the life cycle assessment (LCA) of corn stover (CS) utilization for lignocellulosic ethanol remain substantial and are closely tied to methodological choices, which significantly influence reported sustainability outcomes. A major source of variability lies in the definition of system boundaries, as many studies either restrict analysis to the ethanol production stage or apply expanded boundaries that incorporate displacement effects such as fossil fuel substitution and coproduct credits, yet still fail to capture broader system-level constraints, including competition for biomass and land resources (Hedegaard et al., 2008). This limitation is particularly relevant for CS, where residue removal may interfere with essential ecological functions such as soil carbon retention and nutrient cycling. Empirical evidence, however, suggests that these impacts are context-dependent; for instance, a long-term field study in central Pennsylvania found that partial removal (50%) of corn stover had minimal effects on soil organic carbon and nutrient levels, with only a reduction in surface phosphorus observed, while maintaining adequate soil cover to mitigate erosion risks (Adler et al., 2015; Wilhelm et al., 2010).

In addition to system boundary issues, inconsistencies in the choice of functional units, ranging from energy output to land area or transport service, further complicate cross-study comparisons and can significantly alter interpretations of environmental performance (Hedegaard et al., 2008; von Blottnitz & Curran, 2007). Carbon accounting also introduces considerable uncertainty, particularly in relation to soil carbon dynamics and indirect land-use change (ILUC), both of which depend heavily on assumptions regarding crop yields, land availability, and market-mediated responses. These uncertainties are compounded by limited understanding of the ethanol conversion stage, as most LCA studies emphasize feedstock production or compare bioethanol systems with fossil fuels using generalized assumptions about conversion technologies (Davis et al., 2009; González-García et al., 2010). Consequently, critical environmental burdens associated with conversion processes—such as energy-intensive pretreatment, hydrolysis, and distillation—are often underrepresented, largely due to process uncertainties and the lack of commercial-scale lignocellulosic ethanol facilities (Borrion et al., 2012).

Methodological differences in LCA approaches further contribute to divergent findings. Attributional LCAs, typically characterized by narrower system boundaries, tend to report favorable greenhouse gas (GHG) reductions and positive energy balances. In contrast, consequential LCA approaches, which account for alternative biomass uses and broader market interactions, often yield less optimistic conclusions, indicating that diverting biomass to ethanol production may displace more efficient energy pathways and reduce overall environmental benefits (Hedegaard et al., 2008). Furthermore, variations in coproduct allocation methods and the frequent exclusion of wider environmental impact categories—such as acidification, eutrophication, and toxicity—lead to inconsistent assessments, with some studies reporting unfavorable impacts despite climate-related benefits (von Blottnitz & Curran, 2007).

5.2. Soil Health and Agricultural Impacts: Soil Response and Management Strategies

The main agronomic concern of CS removal is its long-term impact on soil fertility and productivity. CS contains essential nutrients (e.g., C, N, P, K, Ca, Mg), and its return to the soil is critical for nutrient recycling and sustaining yields. Removing 40% of CS can reduce soil N by 20%, P by 14%, and K by 110% (Fixen, 2007). In tropical Mexico, CS removal led to declines in soil organic carbon (SOC), total N, and extractable P (Salinas-Garcia et al., 2001). Historically, conventional tillage and residue removal have depleted up to 70% of SOC in agricultural soils (Wilhelm et al., 2007). Nutrient loss varies with removal rate and site conditions, while some studies show minimal impact (Karlen et al., 1984), others report significant SOC and N losses (Blanco-Canqui & Lal, 2007; Karlen et al., 1994). Threshold removal rates of 30–50% are often proposed based on erosion control (Graham et al., 2007), yet higher stover retention is likely needed to sustain soil fertility (Wilhelm et al., 2007). Effects on soil structure are inconsistent: Karlen et al. (1994) found no change in aggregate stability, but Blanco-Canqui et al. (2006) observed rapid declines. Partial harvests can mitigate pest pressures and reduce N fertilizer needs (Coulter & Nafziger, 2008). Guidelines recommend harvesting only in high-yield areas with sufficient residue left to prevent erosion and maintain SOC (Adler et al., 2015; Wilhelm et al., 2007).

CS removal's effects on nitrogen are less studied than on SOC. Full removal under no-till may reduce soil N by 10–20% (Blanco-Canqui & Lal, 2009), though minimum tillage can improve N retention when at least 33% of residue remains (Salinas-Garcia et al., 2001). Results vary depending on fertilizer use (Karlen et al., 1994). Moreso, some studies show no yield decline (Adler et al., 2015), while others observe small positive or negative effects (Karlen, 2014). However, high residue levels can lower soil temperatures, harbor pests, and reduce germination (Sindelar et al., 2013), but also conserve moisture during droughts (Baumhardt et al., 2013). Excessive removal depletes SOC and nutrients (Blanco-Canqui et al., 2006; Wilhelm et al., 2004), though some findings suggest more nuanced impacts (Clapp et al., 2000).

CS plays a historical and modern role in agriculture and bioenergy (H. Xu et al., 2019). Sustainable removal depends on site-specific factors, including soil type, slope, climate, tillage, and crop rotation (Aghaei et al., 2022). Harvest efficiency is under 40% using conventional dry methods, while wet harvests reduce field passes (Sokhansanj et al., 2002). One-pass systems are being developed to improve efficiency and retain sufficient residue (Luo et al., 2009; Shinners et al., 2007). CS harvest can enhance seedbed conditions and reduce disease pressure, but may also reduce soil C and N, and increase erosion risks (Mann et al., 2002). A synthesis of 409 data points showed SOC stocks declined by 8% in soils with stover removal, regardless of tillage, but depending on removal rate, soil depth, and rotation (Anderson-Teixeira et al., 2009; Johnson et al., 2006). Though CS is a viable biofuel feedstock, long-term sustainability hinges on maintaining SOC and soil function. Residues regulate temperature and moisture, protect against erosion, and support biological processes (Blanco-Canqui & Lal, 2009). In high-residue or sensitive systems, stover can hinder crop establishment or increase agrochemical needs. Effective management is critical to balance energy use with soil health (Wilhelm et al., 2004).

5.3. Economic Viability: Cost of Collection, Processing, and Commercialization

Feedstock cost accounts for approximately 35–50% of the total cost of producing ethanol or power. The exact proportion depends on biomass species, yield, location, climate, local economic factors, and the type of systems used for harvesting, packaging, processing, storing, and transporting the biomass (Sokhansanj & Fenton, 2006). Harvesting and collection involve gathering and removing biomass from the field, depending on its condition at harvest. These operations vary based on biomass type (e.g., grass, wood, or crop residue), moisture content, and intended end use. For crop residues, harvest operations must be coordinated with grain harvest, whereas dedicated crops (grass and wood) can be harvested in separate, biomass-only operations. Collection refers to picking up biomass, packaging it, and transporting it to a nearby storage site. The most common collection method is baling, typically into either round or square bales. Round bales are popular on many U.S. farms (Cundiff, 1996). However, their use in large-scale biomass applications is limited because round bales tend to deform under static loads and are difficult to transport efficiently, especially if not perfectly round (Sokhansanj & Fenton, 2006).

Estimating nutrient removal through CS harvest is complicated by the translocation and leaching of soluble nutrients, such as potassium (K), from the upper plant parts between physiological maturity and harvest. Combined with operational variability, this leads to inconsistent nutrient composition data (Avila-Segura et al., 2011; Birrell et al., 2014; Karlen et al., 2014), making it difficult to make precise, field-specific decisions regarding CS harvest and marketing (Karlen et al., 2015). Bio-based energy has the potential to enhance energy independence, drive rural economic growth, and provide environmental benefits. However, supplying the large volumes of low-density biomass needed for industrial-scale biorefineries presents logistical challenges for transport, handling, and storage.

Despite interest in uniform feedstock and better supply logistics, densified biomass has not been widely adopted by biorefineries (Nahar et al., 2021). Earlier studies (Sokhansanj & Fenton, 2006; Sultana et al., 2010) assumed that the costs of densification would primarily offset transportation savings. However, they often overlooked interactions between densification and downstream conversion processes.

5.4. Macro-Level Socio-Economic Impacts: Energy Security and Carbon Markets

Beyond plant-level performance, corn stover bioenergy also has macro-level significance as a residue-based energy option that can diversify domestic supply without requiring additional land area. In this sense, its importance is not limited to greenhouse-gas mitigation: authoritative energy assessments increasingly frame sustainable biofuels as relevant to energy security and rural job creation, while broader renewables literature emphasizes that locally embedded renewable-energy systems can strengthen economic development and livelihoods when value chains are retained within producing regions (IEA, 2023, 2024; REN21, 2024)

The socio-economic contribution of corn stover is also shaped by carbon-trading and low-carbon credit markets through which low-carbon attributes are measured and monetized. Under the U.S. Renewable Fuel Standard, RINs operate as tradable compliance credits and post-2022 volume-setting explicitly considers rural economic development and agricultural supply effects, while the Low Carbon Fuel Standard uses declining carbon-intensity benchmarks to generate tradable credits for fuels that outperform the benchmark; in parallel, public low-carbon agriculture programs are beginning to pair biomass recycling with MRV systems, local agro-processing, job creation, and even exploration of voluntary carbon-market access (California Air Resources Board, n.d.; IEA, 2024; Y. Li et al., 2024; Uuited States Environmental Protection Agency, 2025; World Bank, 2024)

6. Challenges

Pretreatment remains one of the most significant technical barriers to efficiently converting lignocellulosic biomass like CS into biofuels. This critical step, necessary to overcome biomass recalcitrance, is economically and environmentally challenging due to its high energy requirements, long reaction times, and chemical usage (Yang & Wyman, 2008). Pretreatment alone accounts for 20%–30% of total biofuel production costs, yet little research has addressed energy consumption across different pretreatment methods. Although pelleting has been shown to improve hydrolysis yields (Guragain et al., 2013; Nahar & Pryor, 2014, 2017; Rijal et al., 2012). Its effects on energy savings and greenhouse gas (GHG) reductions remain underexplored (Nahar et al., 2021). High production costs continue to limit the commercial viability of lignocellulosic ethanol at scale (Shadbahr et al., 2015). Infrastructure and logistics also pose major challenges across the biomass supply chain, particularly regarding storage, transport, and processing. Transportation costs are highly dependent on biomass bulk density, which influences the overall economics of biofuel production (Sokhansanj & Fenton, 2006).

7. Barriers and Strategies

Corn stover utilization significantly reduces lifecycle carbon emissions and enables diverse products such as ethanol, biogas, lignin-derived materials, in a circular bioeconomy (Zabed et al., 2023; Zhang et al., 2025). Table 1 outlines major barriers with some evidence/impacts, and remedies/mitigation strategies.

Table 1. Barriers and Strategies of Corn Stover Utilization.

Barrier	Evidence & Impact	Remedy
Biomass recalcitrance to enzymatic conversion	Lignin-carbohydrate complexes inhibit hydrolysis efficiency (Himmel et al., 2007).	Deploy integrated pretreatment pathways such as alkali-assisted steam explosion or ionic liquid fractionation to disrupt lignin structure and improve cellulose accessibility. Combine with genetically engineered low-lignin feedstocks to enhance digestibility and reduce severity requirements. (J. Wang et al., 2026)
High pretreatment/enzyme costs.	Pretreatment is the most energy- and cost-intensive stage, significantly influencing economic and environmental performance (Baral & Shah, 2017).	Implement process-integrated pretreatment systems (e.g., NaOH/ethanol organosolv with solvent recovery, or steam explosion coupled with heat integration). Adopt on-site enzyme production, enzyme recycling, or immobilized enzyme systems to reduce operational cost (Pei et al., 2026)
Low bulk density and high logistics cost	Low density increases transportation and handling costs, limiting supply chain efficiency (Sokhansanj & Turhollow, 2004).	Introduce densification technologies such as pelleting, briquetting, or baling combined with decentralized preprocessing depots to reduce transport distances and stabilize supply chains (Manandhar & Shah, 2018).
Moisture variability & storage	Fresh stover often exceeds 40% moisture in-season, causing aerobic spoilage (dry matter losses approximately 8–28% at 20–50% MC).	Optimize harvest timing within low-moisture windows, apply high-moisture densification (e.g., pelleting) or forced-air drying systems, and utilize covered or anaerobic storage systems (e.g., silage, sealed bags) to minimize spoilage losses (Smith et al., 2020).
Soil erosion & SOC loss	Long-term SOC depletion reduces fertility and increases erosion (Wilhelm et al., 2010).	Apply precision residue harvesting with site-specific removal thresholds, integrate soil monitoring tools, and return nutrients via digestate or biochar application to maintain soil health (Zhang et al., 2025).
Nutrient depletion/ Feedstock heterogeneity	CS removal extracts essential nutrients (N, P, K), and variability in composition affects conversion efficiency (Hess et al., 2009).	Develop standardized feedstock quality specifications, integrate advanced preprocessing (size reduction, fractionation), and implement nutrient recycling systems to offset fertilizer demand and stabilize feedstock quality (Battaglia, 2018).
Limited industrial scale-up and sustainability data	Few commercial cellulosic biorefineries exist due to supply-chain and economic hurdles (Pei et al., 2026). As such, most LCAs rely on modeled or pilot systems (Cherubini & Ulgiati, 2010).	Promote integrated biorefinery models with co-located preprocessing, conversion, and valorization units (e.g., lignin-to-chemicals, biogas integration). Support through policy incentives (e.g., Renewable Identification Numbers, carbon credits) and public–private partnerships. Advance process-specific LCAs at demonstration scale incorporating co-product valorization and real operational data.

8. Future Perspectives and Research Directions

The future of CS utilization in bioenergy production hinges on innovative strategies that enhance yields, integrate circular economy principles, scale industrial applications, and foster interdisciplinary collaboration. Improving biogas and bioethanol yields requires both genetic and process innovations. Genetically engineered corn varieties with lower lignin content and higher cellulose concentrations can significantly enhance digestibility, thereby increasing biogas and ethanol yields (M. Li et al., 2016). Concurrently, advancements in pretreatment techniques—such as steam explosion, alkaline treatments, and enzymatic hydrolysis—can improve efficiency of lignocellulosic breakdown, making more fermentable sugars available for conversion (Alvira et al., 2010). The integration of these innovations can substantially boost overall energy yields while maintaining economic feasibility.

Circular economy approaches further enhance the sustainability of CS stover utilization by promoting waste valorization and co-product development (Beckham et al., 2016). These strategies not only reduce environmental impacts but also improve the economic viability of biorefineries. Scalability and industrial application are critical for commercial success. Decentralized biorefinery models can reduce transportation costs and logistical challenges in corn-producing regions (Sokhansanj et al., 2008), while biomass densification techniques, such as pelletization and briquetting, enhance storage and transport efficiency (Kaliyan & Morey, 2006).

Policy incentives, including subsidies and carbon credits, will be vital in encouraging industrial adoption and driving investments in large-scale bioenergy projects. Moreso, a particularly promising direction is the use of interpretable AI/ML frameworks that link feedstock properties, pretreatment severity, sugar release, anaerobic-digestion stability, lifecycle carbon intensity, and low-carbon credit eligibility, so that process optimization is aligned with commercial viability and broader socio-economic performance rather than yield alone (R. Gupta et al., 2024; IEA, 2024; Khan et al., 2023; C. Wang et al., 2023). Collectively, these future directions promise to elevate CS's role as a sustainable bioenergy feedstock while advancing broader circular economy and climate objectives.

9. Conclusion

Corn stover (CS) generated at approximately 1.66×10^9 t yr^{-1} globally, nearly half of total maize biomass, represents an abundant, low-cost, non-food lignocellulosic resource with minimal land-use change implications. The high structural carbohydrate content provides a substantial reservoir of fermentable sugars, supporting efficient second-generation bioethanol and biogas production without exacerbating food–fuel competition. Moreover, CS-derived fuels exhibit low carbon intensity (21 g CO_2e MJ^{-1}) and significant greenhouse gas mitigation potential relative to fossil fuels, positioning CS as a strategic feedstock for compliance with renewable fuel standards and broader climate objectives. Its versatility across value chains, including fermentative biofuels, anaerobic digestion, and lignin valorization into advanced materials, further aligns CS with circular bioeconomy principles. The existence of large, underutilized residue pools, such as the 27–111 Mt yr^{-1} potentially available in the U.S. Corn Belt, underscores its scalability and the feasibility of decentralized biorefineries that can reduce rural energy poverty and transport emissions.

However, the realization of these resource advantages is inherently constrained by agronomic, environmental, and logistical considerations that must be carefully managed to ensure long-term sustainability. Residue harvest increases annual nutrient removal compared to grain-only systems, necessitating site-specific retention thresholds to safeguard soil organic carbon (SOC) and erosion control. Evidence suggests maintaining between 6 and 9.25 Mg ha^{-1} of residue, depending on agroecological conditions, reinforcing the need for adaptive, yield-based harvesting guidelines. Simultaneously, logistical and preprocessing constraints, particularly the deformation and transport inefficiencies associated with round bales, highlight the importance of densification strategies. Emerging evidence indicates that pelleting and granulation not only enhance bulk density and handling efficiency but may also improve pretreatment performance and reduce downstream costs. Ensuring feedstock cleanliness, uniform moisture, and controlled particle size remains essential for stable biorefinery operations.

In this context, advancing from constraint management to system optimization requires an integrated approach that maximizes value across the entire biomass conversion. Beyond fuels, coproduct valorization strengthens system-level sustainability. Lignin upgrading to bioplastics, adhesives, and carbon fibers, alongside nutrient recycling through anaerobic digestate application, closes material loops and enhances economic resilience. However, optimizing environmental performance at commercial scale demands further research into process design and technological improvements in ethanol plants. Ultimately, progress will depend on interdisciplinary collaboration across agronomy, engineering, and policy to harmonize residue management, conversion efficiency, and supply chain integration, thereby enabling the sustainable and scalable deployment of CS within the evolving economy.

CRediT Author Statement: Anthony Amotoe-Bondzie: Writing – original draft; **Isaac Slaven:** Conceptualization, Supervision, Writing – review & editing, and Funding; **Nichole Hugo:** Writing – review & editing, Conceptualization, Supervision, and Writing – review & editing; **Kelly J. Best:** Writing – review & editing.

Data Availability Statement: Not applicable.

Funding: This research was funded by the U.S. Department of Energy, Grant No. DE-EE-0010679.

Conflicts of Interest: The authors declare no conflicts of interest.

IRB Statement: Not applicable.

Informed Consent Statement: Not applicable.

Acknowledgments: The authors of this paper wish to acknowledge Mara Roach and Robert Davies of the Jo-Carroll Local Redevelopment Authority in Savanna, Illinois, USA for orchestrating the financial support for this research. Additionally, the funders had no role in the study design; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

CS	Corn Stover
VFAs	Volatile Fatty Acids
WAS	Waste-Activated Sludge
WWTPs	Wastewater Treatment Plants
AD	Anaerobic Digestion
L-AD	Liquid Anaerobic Digestion
SS-AD	Solid-State Anaerobic Digestion
TS	Total Solids
VC	Vermicompost
AFEX	Ammonia Fiber Explosion
LHW	Liquid Hot Water
DESs	Deep Eutectic Solvents
LCA	Life Cycle Assessment
SOC	Soil Organic Carbon
GHG	Greenhouse Gas
EU	European Union
RINs	Renewable Identification Numbers
K_2O	Potassium Oxide (potash fertilizer equivalent)
P_2O_5	Phosphorus Pentoxide (phosphate fertilizer equivalent)
CO_2e	Carbon Dioxide Equivalent
USA	United States of America
MgO	Magnesium Oxide
NaOH	Sodium Hydroxide
Na_2CO_3	Sodium Carbonate
SO_2	Sulfur Dioxide
IEA	International Energy Agency

References

Aditiya, H. B., Mahlia, T. M. I., Chong, W. T., Nur, H., & Sebayang, A. H. (2016). Second generation bioethanol production: A critical review. *Renewable and Sustainable Energy Reviews*, *66*, 631–653. https://doi.org/10.1016/j.rser.2016.07.015

Adler, P. R., Rau, B. M., & Roth, G. W. (2015). Sustainability of corn stover harvest strategies in Pennsylvania. *BioEnergy Research*, *8*, 1310–1320. https://doi.org/10.1007/s12155-015-9593-2

Ağdağ, O. N., & Sponza, D. T. (2007). Co-digestion of mixed industrial sludge with municipal solid wastes in anaerobic simulated landfilling bioreactors. *Journal of Hazardous Materials*, *140*(1–2), 75–85. https://doi.org/10.1016/j.jhazmat.2006.06.059

Aghaei, S., Karimi Alavijeh, M., Shafiei, M., & Karimi, K. (2022). A comprehensive review on bioethanol production from corn stover: Worldwide potential, environmental importance, and perspectives. *Biomass and Bioenergy*, *161*, 106447. https://doi.org/10.1016/j.biombioe.2022.106447

Agostinho, F., Bertaglia, A. B. B., Almeida, C. M. V. B., & Giannetti, B. F. (2015). Influence of cellulase enzyme production on the energetic–environmental performance of lignocellulosic ethanol. *Ecological Modelling*, *315*, 46–56. https://doi.org/10.1016/j.ecolmodel.2014.09.005

Alvira, P., Tomas-Pejo, E., Ballesteros, M., & Negro, M. J. (2010). Pretreatment technologies for an efficient bioethanol production process based on enzymatic hydrolysis: A review. *Bioresource Technology*, *101*(13), 4851–4861. https://doi.org/10.1016/j.biortech.2009.11.093
Anal, A. K. (2019). Quality ingredients and safety concerns for traditional fermented foods and beverages from Asia: A review. *Fermentation*, *5*(1), 8. https://doi.org/10.3390/fermentation5010008
Anderson-Teixeira, K. J., Davis, S. C., Masters, M. D., & Delucia, E. H. (2009). Changes in soil organic carbon under biofuel crops. *Global Change Biology and Bioenergy*, *1*(1), 75–96. https://doi.org/10.1111/j.1757-1707.2008.01001.x
Avila-Segura, M., Barak, P., Hedtcke, J. L., & Posner, J. L. (2011). Nutrient and alkalinity removal by corn grain, stover and cob harvest in Upper Midwest USA. *Biomass and Bioenergy*, *35*(3), 1190–1195. https://doi.org/10.1016/j.biombioe.2010.12.010
Balat, M., Balat, H., & Öz, C. (2008). Progress in bioethanol processing. *Progress in Energy and Combustion Science*, *34*(5), 551–573. https://doi.org/10.1016/j.pecs.2007.11.001
Baral, N. R., & Shah, A. (2017). Comparative techno-economic analysis of steam explosion, dilute sulfuric acid, ammonia fiber explosion and biological pretreatments of corn stover. *Bioresource Technology*, *232*, 331–343. https://doi.org/10.1016/j.biortech.2017.02.068
Battaglia, M. (2018). Harvesting and nutrient replacement costs associated with corn stover removal in Virginia. Virginia Tech. https://vtechworks.lib.vt.edu/server/api/core/bitstreams/4c8e9d45-6e6d-4964-9cdf-a0ba84467f22/content
Bauer, A., Bösch, P., Friedl, A., & Amon, T. (2009). Analysis of methane potentials of steam-exploded wheat straw and estimation of energy uields of combined ethanol and methane production. *Journal of Biotechnology*, *142*(1), 50–55. https://doi.org/10.1016/j.jbiotec.2009.01.017
Baumhardt, R. L., Schwartz, R., Howell, T., Evett, S. R., & Colaizzi, P. (2013). Residue management effects on water use and yield of deficit irrigated corn. *Agronomy Journal*, *105*(4), 1035–1044. https://doi.org/10.2134/agronj2012.0362
Beckham, G. T., Johnson, C. W., Karp, E. M., Salvachúa, D., & Vardon, D. R. (2016). Opportunities and challenges in biological lignin valorization. *Current Opinion in Biotechnology*, *42*, 40–53. https://doi.org/10.1016/j.copbio.2016.02.030
Birrell, S. J., Karlen, D. L., & Wirt, A. (2014). Development of sustainable corn stover harvest strategies for cellulosic ethanol production. *BioEnergy Research*, *7*, 509–516. https://doi.org/10.1007/s12155-014-9418-8
Blanco-Canqui, H., & Lal, R. (2007). Soil and crop response to harvesting corn residues for biofuel production. *Geoderma*, *141*(3–4), 355–362. https://www.sciencedirect.com/science/article/pii/S0016706107001851
Blanco-Canqui, H., & Lal, R. (2009). Corn stover removal for expanded uses reduces soil fertility and structural stability. *Soil Science Society of America Journal*, *73*(2), 418–426. https://doi.org/10.2136/sssaj2008.0141
Blanco-Canqui, H., Lal, R., Post, W. M., Izaurralde, R. C., & Owens, L. B. (2006). Rapid changes in soil carbon and structural properties due to stover removal from no-till corn plots. *Soil Science*, *171*(6), 468–482. https://doi.org/10.1097/01.ss.0000209364.85816.1b
Borrion, A. L., McManus, M. C., & Hammond, G. P. (2012). Environmental life cycle assessment of lignocellulosic conversion to ethanol: A review. *Renewable and Sustainable Energy Reviews*, *16*(7), 4638–4650. https://doi.org/10.1016/j.rser.2012.04.016
Brown, D., & Li, Y. (2013). Solid state anaerobic co-digestion of yard waste and food waste for biogas production. *Bioresource Technology*, *127*, 275–280. https://doi.org/10.1016/j.biortech.2012.09.081
California Air Resources Board. (n.d.). *Low carbon fuel standard*. https://ww2.arb.ca.gov/our-work/programs/low-carbon-fuel-standard/about
Capolupo, L., & Faraco, V. (2016). Green methods of lignocellulose pretreatment for biorefinery development. *Applied Microbiology and Biotechnology*, *100*, 9451–9467. https://doi.org/10.1007/s00253-016-7884-y
Chang, J., Cheng, W., Yin, Q., Zuo, R., Song, A., Zheng, Q., Wang, P., Wang, X., & Liu, J. (2012). Effect of steam explosion and microbial fermentation on cellulose and lignin degradation of corn stover. *Bioresource Technology*, *104*, 587–592. https://doi.org/10.1016/j.biortech.2011.10.070
Chen, G., Zheng, Z., Yang, S., Fang, C., Zou, X., & Luo, Y. (2010). Experimental co-digestion of corn stalk and vermicompost to improve biogas production. *Waste Management*, *30*(10), 1834–1840. https://doi.org/10.1016/j.wasman.2010.03.014
Chen, H., & Fu, X. (2016). Industrial technologies for bioethanol production from lignocellulosic biomass. *Renewable and Sustainable Energy Reviews*, *57*, 468–478. https://doi.org/10.1016/j.rser.2015.12.069
Chen, Z., & Wan, C. (2018). Ultrafast fractionation of lignocellulosic biomass by microwave-assisted deep eutectic solvent pretreatment. *Bioresource Technology*, *250*, 532–537. https://doi.org/10.1016/j.biortech.2017.11.066
Cherubini, F., & Ulgiati, S. (2010). Crop residues as raw materials for biorefinery systems – A LCA case study. *Applied Energy*, *87*(1), 47–57. https://doi.org/10.1016/j.apenergy.2009.08.024
Clapp, C. E., Allmaras, R. R., Layese, M. F., Linden, D. R., & Dowdy, R. H. (2000). Soil organic carbon and ^{13}C abundance as related to tillage, crop residue, and nitrogen fertilization under continuous corn management in Minnesota. *Soil and Tillage Research*, *55*(3–4), 127–142. https://doi.org/10.1016/S0167-1987(00)00110-0
Coulter, J. A., & Nafziger, E. D. (2008). Continuous corn response to residue management and nitrogen fertilization. *Agronomy Journal*, *100*(6), 1774–1780. https://doi.org/10.2134/agronj2008.0170
Cundiff, J. S. (1996). Simulation of five large round bale harvesting systems for biomass. *Bioresource Technology*, *56*(1), 77–82. https://doi.org/10.1016/0960-8524(95)00168-9
Cui, Z., Wan, C., Shi, J., Sykes, R. W., & Li, Y. (2012). Enzymatic digestibility of corn stover fractions in response to fungal pretreatment. *Industrial & Engineering Chemistry Research*, *51*(21), 7153–7159. https://doi.org/10.1021/ie300487z
Czajkowski, Ł., Wojcieszak, D., Olek, W., & Przybył, J. (2019). Thermal properties of fractions of corn stover. *Construction and Building Materials*, *210*, 709–712. https://doi.org/10.1016/j.conbuildmat.2019.03.092
Da Costa Sousa, L., Chundawat, S. P. S., Balan, V., & Dale, B. E. (2009). 'Cradle-to-grave' assessment of existing lignocellulose pretreatment technologies. *Current Opinion in Biotechnology*, *20*(3), 339–347. https://doi.org/10.1016/j.copbio.2009.05.003
Darr, M., Schon, B., & Schau, D. (2014). *Nutrient removal from corn Stover harvesting*. Iowa State University Extension and Outreach.
Davis, S. C., Anderson-Teixeira, K. J., & DeLucia, E. H. (2009). Life-cycle analysis and the ecology of biofuels. *Trends in Plant Science*, *14*(3), 140–146. https://doi.org/10.1016/j.tplants.2008.12.006
El-Mashad, H. M., & Zhang, R. (2010). Biogas production from co-digestion of dairy manure and food waste. *Bioresource Technology*, *101*(11), 4021–4028. https://doi.org/10.1016/j.biortech.2010.01.027

Farheen, V. H., Abdullah, M. A., & Moorthy, I. G. (2026). Lignocellulosic bioethanol production: A review on pretreatment strategies, biofuel separation, and artificial intelligence/machine learning − based sustainable optimization. *Current Research in Biotechnology*, *11*, 100355. https://doi.org/10.1016/j.crbiot.2025.100355
Fernandes, T. V., Klaasse Bos, G. J., Zeeman, G., Sanders, J. P. M., & Van Lier, J. B. (2009). Effects of thermo-chemical pre-treatment on anaerobic biodegradability and hydrolysis of lignocellulosic biomass. *Bioresource Technology*, *100*(9), 2575–2579. https://doi.org/10.1016/j.biortech.2008.12.012
Fixen, P. E. (2007). Potential biofuels influence on nutrient use and removal in the US. *Better Crops*, *91*(2), 12–14. https://www.researchgate.net/profile/Paul-Fixen/publication/284312614
Fotidis, I. A., Kougias, P. G., Zaganas, I. D., Kotsopoulos, T. A., & Martzopoulos, G. G. (2014). Inoculum and zeolite synergistic effect on anaerobic digestion of poultry manure. *Environmental Technology*, *35*(10), 1219–1225. https://doi.org/10.1080/09593330.2013.865083
Ganeshan, P., Bose, A., Lee, J., Barathi, S., & Rajendran, K. (2024). Machine learning for high solid anaerobic digestion: Performance prediction and optimization. *Bioresource Technology*, *400*, 130665. https://doi.org/10.1016/j.biortech.2024.130665
Gao, Y., Lipton, A. S., Wittmer, Y., Murray, D. T., & Mortimer, J. C. (2020). A grass-specific cellulose–xylan interaction dominates in sorghum secondary cell walls. *Nature Communications*, *11*, 6081. https://doi.org/10.1038/s41467-020-19837-z
Geng, X., & Henderson, W. A. (2012). Pretreatment of corn stover by combining ionic liquid dissolution with alkali extraction. *Biotechnology and Bioengineering*, *109*(1), 84–91. https://doi.org/10.1002/bit.23281
Ge, X., Matsumoto, T., Keith, L., & Li, Y. (2014). Biogas energy production from tropical biomass wastes by anaerobic digestion. *Bioresource Technology*, *169*, 38–44. https://doi.org/10.1016/j.biortech.2014.06.067
Gitifar, V., Eslamloueyan, R., & Sarshar, M. (2013). Experimental study and neural network modeling of sugarcane bagasse pretreatment with H_2SO_4 and O^3 for cellulosic material conversion to sugar. *Bioresource Technology*, *148*, 47–52. https://doi.org/10.1016/j.biortech.2013.08.060
González-García, S., Moreira, M. T., & Feijoo, G. (2010). Comparative environmental performance of lignocellulosic ethanol from different feedstocks. *Renewable and Sustainable Energy Reviews*, *14*(7), 2077–2085. https://doi.org/10.1016/j.rser.2010.03.035
Graham, R. L., Nelson, R., Sheehan, J., Perlack, R. D., & Wright, L. L. (2007). Current and potential U.S. corn stover supplies. *Agronomy Journal*, *99*(1), 1–11. https://doi.org/10.2134/agronj2005.0222
Grahovac, J., Jokić, A., Dodić, J., Vučurović, D., &Dodić, S. (2016). Modelling and prediction of bioethanol production from intermediates and byproduct of sugar beet processing using neural networks. *Renewable Energy*, *85*, 953–958. https://doi.org/10.1016/j.renene.2015.07.054
Guan, Y., Wu, J., Gao, Y., Zheng, Y., Zheng, J., Xia, T., Li, G., Zhang, L., Shi, Y., Huo, M., Yang, X., & Wang, X. (2025). Achieve full utilization of lignin, cellulose and hemicellulose from corn stover with amphiphilic polyoxometalate catalysts in a one-pot method. *International Journal of Biological Macromolecules*, *309*, 142892. https://doi.org/10.1016/j.ijbiomac.2025.142892
Guo, Z., Zhou, A., Yang, C., Liang, B., Sangeetha, T., He, Z., Wang, L., Cai, W., Wang, A., & Liu, W. (2015). Enhanced short chain fatty acids production from waste activated sludge conditioning with typical agricultural residues: Carbon source composition regulates community functions. *Biotechnology for Biofuels*, *8*, 192. https://doi.org/10.1186/s13068-015-0369-x
Gupta, R., Ouderji, Z. H., Uzma, Yu, Z., Sloan, W. T., & You, S. (2024). Machine learning for sustainable organic waste treatment: A critical review. *npj Materials Sustainability*, *2*, 5. https://doi.org/10.1038/s44296-024-00009-9
Guragain, Y. N., Wilson, J., Staggenborg, S., McKinney, L., Wang, D., & Vadlani, P. V. (2013). Evaluation of pelleting as a pre-processing step for effective biomass deconstruction and fermentation. *Biochemical Engineering Journal*, *77*, 198–207. https://doi.org/10.1016/j.bej.2013.05.014
Haldar, D., Shabbirahmed, A. M., & Mahanty, B. (2023). Multivariate regression and artificial neural network modelling of sugar yields from acid pretreatment and enzymatic hydrolysis of lignocellulosic biomass. *Bioresource Technology*, *370*, 128519. https://doi.org/10.1016/j.biortech.2022.128519
Hedegaard, K., Thyø, K. A., & Wenzel, H. (2008). Life cycle assessment of an advanced bioethanol technology in the perspective of constrained biomass availability. *Environmental Science & Technology*, *42*(21), 7992–7999. https://doi.org/10.1021/es800358d
Hess, J. R., Kenney, K. L., Wright, C. T., Perlack, R., & Turhollow, A. (2009). Corn stover availability for biomass conversion: Situation analysis. *Cellulose*, *16*, 599–619. https://doi.org/10.1007/s10570-009-9323-z
Himmel, M. E., Ding, S.-Y., Johnson, D. K., Adney, W. S., Nimlos, M. R., Brady, J. W., & Foust, T. D. (2007). Biomass recalcitrance: Engineering plants and enzymes for biofuels production. *Science*, *315*(5813), 804–807. https://www.jstor.org/stable/20038950
Houfani, A. A., Anders, N., Spiess, A. C., Baldrian, P., & Benallaoua, S. (2020). Insights from enzymatic degradation of cellulose and hemicellulose to fermentable sugars– a review. *Biomass and Bioenergy*, *134*, 105481. https://doi.org/10.1016/j.biombioe.2020.105481
International Energy Agency. (2023). *Bioenergy supply globally in the Net Zero Scenario, 2010-2030*. https://www.iea.org/data-and-statistics/charts/bioenergy-supply-globally-in-the-net-zero-scenario-2010-2030
International Energy Agency. (2024). *Carbon accounting for sustainable biofuels*. https://iea.blob.core.windows.net/assets/79f31c02-0efe-41ca-ac15-9d076bf2cd29/CarbonAccountingforSustainableBiofuels.pdf
Jameel, M. K., Mustafa, M. A., Ahmed, H. S., Mohammed, A. J., Ghazy, H., Shakir, M. N., Lawas, A. M., Mohammed, S. K., Idan, A. H., Mahmoud, Z. H., Sayadi, H., & Kianfar, E. (2024). Biogas: Production, properties, applications, economic and challenges: A review. *Results in Chemistry*, *7*, 101549. https://doi.org/10.1016/j.rechem.2024.101549
Johnson, J. M. F., Papiernik, S. K., Mikha, M. M., Spokas, K., Tomer, M. D., & Weyers, S. L. (2009). Soil processes and residue harvest management. In R. Lal, & B. A. Stewart (Eds.), *Soil quality and biofuel production* (pp. 1–44). Taylor & Francis Group.
Johnson, J. M. F., Allmaras, R. R., & Reicosky, D. C. (2006). Estimating source carbon from crop residues, roots and rhizodeposits using the national grain-yield database. *Agronomy Journal*, *98*(3), 622–636. https://doi.org/10.2134/agronj2005.0179
Kakihana, T., Nagata, K., & Sitia, R. (2012). Peroxides and peroxidases in the endoplasmic reticulum: Integrating redox homeostasis and oxidative folding. *Antioxidants & Redox Signaling*, *16*(8), 763–771. https://doi.org/10.1089/ars.2011.4238
Kaliyan, N., & Morey, R. V. (2006). Densification characteristics of corn stover and switchgrass. *Transactions of the ASABE*, *52*(3), 907–920. https://doi.org/10.13031/2013.21202

Kangle, K. M., Kore, S. V., Kore, V. S., Kulkarni, G. S. (2012). Recent trends in anaerobic codigestion: A review. *Universal Journal of Environmental Research and Technology*, *2*(4), 210–219. https://www.environmentaljournal.org/2-4/ujert-2-4-3.pdf

Karimi, K., Shafiei, M., & Kumar, R. (2013). Progress in physical and chemical pretreatment of lignocellulosic biomass. In V. K. Gupta & M. G. Tuohy (Eds.), *Biofuel technologies: Recent developments* (pp. 53–96). Springer Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-34519-7_3

Karlen, D. L. (2014). *Cellulosic energy cropping systems*. John Wiley & Sons Ltd.

Karlen, D. L., Birrell, S. J., Johnson, J. M. F., Osborne, S. L., Schumacher, T. E., Varvel, G. E., Ferguson, R. B., Novak, J. M., Fredrick, J. R., Baker, J. M., Lamb, J. A., Adler, P. R., Roth, G. W., & Nafziger, E. D. (2014). Multilocation corn stover harvest effects on crop Yields and nutrient removal. *BioEnergy Research*, *7*, 528–539. https://doi.org/10.1007/s12155-014-9419-7

Karlen, D. L., Hunt, P. G., & Campbell, R. B. (1984). Crop residue removal effects on corn yield and fertility of a Norfolk sandy loam. *Soil Science Society of America Journal*, *48*(4), 868–872. https://doi.org/10.2136/sssaj1984.03615995004800040034x

Karlen, D. L., Kovar, J. L., & Birrell, S. J. (2015). Corn stover nutrient removal estimates for central Iowa, USA. *Sustainability*, *7*(7), 8621–8634. https://doi.org/10.3390/su7078621

Karlen, D. L., Wollenhaupt, N. C., Erbach, D. C., Berry, E. C., Swan, J. B., Eash, N. S., & Jordahl, J. L. (1994). Crop residue effects on soil quality following 10-years of no-till corn. *Soil and Tillage Research*, *31*(2–3), 149–167. https://doi.org/10.1016/0167-1987(94)90077-9

Khan, M., Chuenchart, W., Surendra, K. C., & Kumar Khanal, S. (2023). Applications of artificial intelligence in anaerobic co-digestion: Recent advances and prospects. *Bioresource Technology*, *370*, 128501. https://doi.org/10.1016/j.biortech.2022.128501

Kougias, P. G., Boe, K., Tsapekos, P., & Angelidaki, I. (2014). Foam suppression in overloaded manure-based biogas reactors using antifoaming agents. *Bioresource Technology*, *153*, 198–205. https://doi.org/10.1016/j.biortech.2013.11.083

Kumar, A. K., & Sharma, S. (2017). Recent updates on different methods of pretreatment of lignocellulosic feedstocks: A review. *Bioresources and Bioprocessing*, *4*, 7. https://doi.org/10.1186/s40643-017-0137-9

Kumar, R., Tabatabaei, M., Karimi, K., & Sárvári Horváth, I. (2016). Recent updates on lignocellulosic biomass derived ethanol – A review. *Biofuel Research Journal*, *3*(1), 347–356. https://doi.org/10.18331/BRJ2016.3.1.4

Labatut, R. A., Angenent, L. T., & Scott, N. R. (2011). Biochemical methane potential and biodegradability of complex organic substrates. *Bioresource Technology*, *102*(3), 2255–2264. https://doi.org/10.1016/j.biortech.2010.10.035

Lászlok, A., Takács-György, K., & Takács, I. (2020). Examination of first generation biofuel production in some selected biofuel producing countries in Europe: A case study. *Agricultural Economics – Czech*, *66*(10), 469–476. https://doi.org/10.17221/237/2020-AGRICECON

Li, H. Y., Xu, L., Liu, W. J., Fang, M. Q., & Wang, N. (2014). Assessment of the nutritive value of whole corn stover and its morphological fractions. *Asian-Australasian Journal of Animal Sciences*, *27*(2), 194–200. https://doi.org/10.5713/ajas.2013.13446

Li, M., Pu, Y., & Ragauskas, A. J. (2016). Current understanding of the correlation of lignin structure with biomass recalcitrance. *Frontiers in Chemistry*, *4*, *45*. https://doi.org/10.3389/fchem.2016.00045

Li, Y., Sohail, M. T., Zhang, Y., & Ullah, S. (2024). Bioenergy for sustainable rural development: Elevating government governance with environmental policy in China. *Land*, *13*(12), 2147. https://doi.org/10.3390/land13122147

Li, Y., Zhang, R., Liu, G., Chen, C., He, Y., & Liu, X. (2013). Comparison of methane production potential, biodegradability, and kinetics of different organic substrates. *Bioresource Technology*, *149*, 565–569. https://doi.org/10.1016/j.biortech.2013.09.063

Lindkvist, E. (2020). System studies of biogas production-comparisons and performance [Doctoral dissertation, Linköping University]. ProQuest Dissertations and Theses Global. https://search.proquest.com/openview/42ebbd636abe9f7b8bb09efb69793e8e/1?pq-origsite=gscholar&cbl=2026366&diss=y

Liu, W., He, Z., Yang, C., Zhou, A., Guo, Z., Liang, B., Varrone, C., & Wang, A.-J. (2016). Microbial network for waste activated sludge cascade utilization in an integrated system of microbial electrolysis and anaerobic fermentation. *Biotechnology for Biofuels*, *9*, 83. https://doi.org/10.1186/s13068-016-0493-2

Liu, W., Wang, A., Cheng, S., Logan, B. E., Yu, H., Deng, Y., Van Nostrand, J. D., Wu, L., He, Z., & Zhou, J. (2010). Geochip-based functional gene analysis of anodophilic communities in microbial electrolysis cells under different operational modes. *Environmental Science & Technology*, *44*(19), 7729–7735. https://doi.org/10.1021/es100608a

Loow, Y.-L., Wu, T. Y., Md. Jahim, J., Mohammad, A. W., & Teoh, W. H. (2016). Typical conversion of lignocellulosic biomass into reducing sugars using dilute acid hydrolysis and alkaline pretreatment. *Cellulose*, *23*, 1491–1520. https://doi.org/10.1007/s10570-016-0936-8

Luo, L., van der Voet, E., Huppes, G., & Udo de Haes, H. A. (2009). Allocation issues in LCA methodology: A case study of corn stover-based fuel ethanol. *The International Journal of Life Cycle Assessment*, *14*, 529–539. https://doi.org/10.1007/s11367-009-0112-6

Manandhar, A., & Shah, A. (2018). Feedstock logistics for agricultural residues and energy crops: Moving biomass from the field to biorefinery gate. Ohioline. https://ohioline.osu.edu/factsheet/fabe-6604

Mann, L., Tolbert, V., & Cushman, J. (2002). Potential environmental effects of corn (*Zea mays* L.) stover removal with emphasis on soil organic matter and erosion. *Agriculture, Ecosystems & Environment*, *89*(3), 149–166. https://doi.org/10.1016/S0167-8809(01)00166-9

Menardo, S., Airoldi, G., Cacciatore, V., & Balsari, P. (2015). Potential biogas and methane yield of maize stover fractions and evaluation of some possible stover harvest chains. *Biosystems Engineering*, *129*, 352–359. https://doi.org/10.1016/j.biosystemseng.2014.11.010

Mirmohamadsadeghi, S., Chen, Z., & Wan, C. (2016). Reducing biomass recalcitrance via mild sodium carbonate pretreatment. *Bioresource Technology*, *209*, 386–390. https://doi.org/10.1016/j.biortech.2016.02.096

Molaverdi, M., Karimi, K., & Mirmohamadsadeghi, S. (2019). Improvement of dry simultaneous saccharification and fermentation of rice straw to high concentration ethanol by sodium carbonate pretreatment. *Energy*, *167*, 654–660. https://doi.org/10.1016/j.energy.2018.11.017

Morales, M., Quintero, J., Conejeros, R., & Aroca, G. (2015). Life cycle assessment of lignocellulosic bioethanol: Environmental impacts and energy balance. *Renewable and Sustainable Energy Reviews*, *42*, 1349–1361. https://doi.org/10.1016/j.rser.2014.10.097

Mosier, N., Hendrickson, R., Ho, N., Sedlak, M., & Ladisch, M. R. (2005). Optimization of pH controlled liquid hot water pretreatment of corn stover. *Bioresource Technology*, *96*(18), 1986–1993. https://doi.org/10.1016/j.biortech.2005.01.013

Myint, M., Nirmalakhandan, N., & Speece, R. E. (2007). Anaerobic fermentation of cattle manure: Modeling of hydrolysis and acidogenesis. *Water Research*, *41*(2), 323–332. https://doi.org/10.1016/j.watres.2006.10.026
Nahar, N., Pandey, R., Pourhashem, G., Ripplinger, D., & Pryor, S. W. (2021). Life cycle perspectives of using non-pelleted vs. pelleted corn stover in a cellulosic biorefinery. *Energies*, *14*(9), 2518. https://doi.org/10.3390/en14092518
Nahar, N., & Pryor, S. W. (2014). Reduced pretreatment severity and enzyme loading enabled through switchgrass pelleting. *Biomass and Bioenergy*, *67*, 46–52. https://doi.org/10.1016/j.biombioe.2014.04.027
Nahar, N., & Pryor, S. W. (2017). Effects of reduced severity ammonia pretreatment on pelleted corn stover. *Industrial Crops and Products*, *109*, 163–172. https://doi.org/10.1016/j.indcrop.2017.08.024
Nelson, R. G. (2002). Resource assessment and removal analysis for corn stover and wheat straw in the Eastern and Midwestern United States—Rainfall and wind-induced soil erosion methodology. *Biomass and Bioenergy*, *22*(5), 349–363. https://www.sciencedirect.com/science/article/pii/S0961953402000065
Oltmans, R. R., & Mallarino, A. P. (2011, November 16–17). Phosphorus and potassium removal and leaching from residue in corn and soybean. North Central Extension-Industry Soil Fertility Conference, Des Moines, IA. https://www.agronext.iastate.edu/soilfertility/info/Mallarino-Oltmans_PKrecycling_NCExt-IndSFConf2011.pdf
Parawira, W., Murto, M., Zvauya, R., & Mattiasson, B. (2004). Anaerobic batch digestion of solid potato waste alone and in combination with sugar beet leaves. *Renewable Energy*, *29*(11), 1811–1823. https://doi.org/10.1016/j.renene.2004.02.005
Parawira, W., Read, J. S., Mattiasson, B., & Björnsson, L. (2008). Energy production from agricultural residues: High methane yields in pilot-scale two-stage anaerobic digestion. *Biomass and Bioenergy*, *32*(1), 44–50. https://doi.org/10.1016/j.biombioe.2007.06.003
Pei, X., Xie, J., & Fan, M. (2026). NaOH-ethanol pretreatment for cellulosic ethanol: A techno-economic comparison with Dilute acid and Steam explosion processes. *Biomass and Bioenergy*, *208*, 108807. https://doi.org/10.1016/j.biombioe.2025.108807
Pereira, R. D., Badino, A. C., & Cruz, A. J. G. (2020). Framework based on artificial intelligence to increase industrial bioethanol production. *Energy & Fuels*, *34*(4), 4670–4677. https://doi.org/10.1021/acs.energyfuels.0c00033
Perlack, R. D., Wright, L. L., Turhollow, A. F., Graham, R. L., Stokes, B. J., & Erbach, D. C. (2005). *Biomass as feedstock for a bioenergy and bioproducts industry: The technical feasibility of a billion-ton annual supply*. United States Department of Agriculture & United States Department of Energy. https://doi.org/10.2172/885984
Perrot, T., Pauly, M., & Ramírez, V. (2022). Emerging Roles of β-glucanases in plant development and adaptative responses. *Plants*, *11*(9), 1119. https://doi.org/10.3390/plants11091119
Phromphithak, S., Onsree, T., & Tippayawong, N. (2021). Machine learning prediction of cellulose-rich materials from biomass pretreatment with ionic liquid solvents. *Bioresource Technology*, *323*, 124642. https://doi.org/10.1016/j.biortech.2020.124642
Qing, Q., Zhou, L., Guo, Q., Gao, X., Zhang, Y., He, Y., & Zhang, Y. (2017). Mild alkaline presoaking and organosolv pretreatment of corn stover and their impacts on corn stover composition, structure, and digestibility. *Bioresource Technology*, *233*, 284–290. https://doi.org/10.1016/j.biortech.2017.02.106
Radwan, A. M., Sebak, H. A., Mitry, N. R., El-Zanati, E. A., & Hamad, M. A. (1993). Dry anaerobic fermentation of agricultural residues. *Biomass and Bioenergy*, *5*(6), 495–499. https://doi.org/10.1016/0961-9534(93)90045-6
REN21. (2024). *Renewables 2024 global status report collection*. https://www.ren21.net/gsr-2024/
Rijal, B., Igathinathane, C., Karki, B., Yu, M., & Pryor, S. W. (2012). Combined effect of pelleting and pretreatment on enzymatic hydrolysis of switchgrass. *Bioresource Technology*, *116*, 36–41. https://doi.org/10.1016/j.biortech.2012.04.054
Roy, R., Rahman, M. S., & Raynie, D. E. (2020). Recent advances of greener pretreatment technologies of lignocellulose. *Current Research in Green and Sustainable Chemistry*, *3*, 100035. https://doi.org/10.1016/j.crgsc.2020.100035
Saha, B. C., Qureshi, N., Kennedy, G. J., & Cotta, M. A. (2016). Biological pretreatment of corn stover with white-rot fungus for improved enzymatic hydrolysis. *International Biodeterioration & Biodegradation*, *109*, 29–35. https://doi.org/10.1016/j.ibiod.2015.12.020
Salinas-Garcia, J. R., Báez-González, A. D., Tiscareño-López, M., & Rosales-Robles, E. (2001). Residue removal and tillage interaction effects on soil properties under rain-fed corn production in Central Mexico. *Soil and Tillage Research*, *59*(1–2), 67–79. https://doi.org/10.1016/S0167-1987(00)00187-2
Sawyer, J. E. (n.d.). *Nutrient removal when harvesting corn stover*. Iowa State University Extension and Outreach.
Sawyer, J. E., & Mallarino, A. P. (2014). *Nutrient Considerations with corn stover harvest (PM 3052C)*. Iowa State University Extension and Outreach. https://store.extension.iastate.edu/product/14052
Sawyer, J. E., & Mallarino, A. P. (2012). Nutrient considerations with corn silage and stover harvest. 2012 Integrated Crop Management Conference, Ames, USA. https://doi.org/10.31274/icm-180809-101
Schober, G., & Trösch, W. (2000). Degradation of digestion residues by lignolytic fungi. *Water Research*, *34*(13), 3424–3430. https://doi.org/10.1016/S0043-1354(00)00076-2
Sewsynker-Sukai, Y., Faloye, F., & Kana, E. B. G. (2017). Artificial neural networks: An efficient tool for modelling and optimization of biofuel production (a mini review). *Biotechnology & Biotechnological Equipment*, *31*(2), 221–235. https://doi.org/10.1080/13102818.2016.1269616
Shadbahr, J., Zhang, Y., & Khan, F. (2015). Life cycle assessment of bioethanol production from woodchips with modifications in the pretreatment process. *Applied Biochemistry and Biotechnology*, *175*, 1080–1091. https://doi.org/10.1007/s12010-014-1293-4
Shenbagamuthuraman, V., & Kasianantham, N. (2023). Microwave irradiation pretreated fermentation of bioethanol production from Chlorella vulgaris Biomasses: Comparative analysis of response surface methodology and artificial neural network techniques. *Bioresource Technology*, *390*, 129867. https://doi.org/10.1016/j.biortech.2023.129867
Shinners, K. J., Binversie, B. N., Muck, R. E., & Weimer, P. J. (2007). Comparison of wet and dry corn stover harvest and storage. *Biomass and Bioenergy*, *31*(4), 211–221. https://doi.org/10.1016/j.biombioe.2006.04.007
Sindelar, A. J., Coulter, J. A., Lamb, J. A., & Vetsch, J. A. (2013). Agronomic responses of continuous corn to stover, tillage, and nitrogen management. *Agronomy Journal*, *105*(6), 1498–1506. https://doi.org/10.2134/agronj2013.0181
Singh, T., Gangil, B., Patnaik, A., Kumar, S., Rishiraj, A., & Fekete, G. (2018). Physico-mechanical, thermal and dynamic mechanical behaviour of natural-synthetic fiber reinforced vinylester based homogenous and functionally graded composites. *Materials Research Express*, *6*(2), 025704. https://doi.org/10.1088/2053-1591/aaee30

Sluiter, J. B., Ruiz, R. O., Scarlata, C. J., Sluiter, A. D., & Templeton., D. W. (2010). Compositional analysis of lignocellulosic feedstocks. 1. Review and description of methods. *Journal of Agricultural and Food Chemistry*, *58*(16), 9043–9053. https://doi.org/10.1021/jf1008023

Smith, W. A., Wendt, L. M., Bonner, I. J., & Murphy, J. A. (2020). Effects of storage moisture content on corn stover biomass stability, composition, and conversion efficacy. *Frontiers in Bioengineering and Biotechnology*, *8*, 716. https://doi.org/10.3389/fbioe.2020.00716

Smullen, E., Finnan, J., Dowling, D., & Mulcahy, P. (2019). The environmental performance of pretreatment technologies for the bioconversion of lignocellulosic biomass to ethanol. *Renewable Energy*, *142*, 527–534. https://doi.org/10.1016/j.renene.2019.04.082

Soam, S., Kapoor, M., Kumar, R., Borjesson, P., Gupta, R. P., & Tuli, D. K. (2016). Global warming potential and energy analysis of second generation ethanol production from rice straw in India. *Applied Energy*, *184*, 353–364. https://doi.org/10.1016/j.apenergy.2016.10.034

Sokhansanj, S., & Fenton, J. (2006). *Cost benefit of biomass supply and pre-processing*. BIOCAP Canada. https://www.biocap.ca/rif/report/Sokhansanj_S.pdf

Sokhansanj, S., Turhollow, A., Cushman, J., & Cundiff, J. (2002). Engineering aspects of collecting corn stover for bioenergy. *Biomass and Bioenergy*, *23*(5), 347–355. https://doi.org/10.1016/S0961-9534(02)00063-6

Sokhansanj, S., & Turhollow, A. F. (2004). Biomass densification–cubing operations and costs for corn stover. *Applied Engineering in Agriculture*, *20*(4), 495–499. https://doi.org/10.13031/2013.16480

Sokhansanj, S., Turhollow, A., Wilkerson, E. (2008). *Development of the Integrated Biomass Supply Analysis and Logistics Model (IBSAL)*. Oak Ridge National Laboratory. https://doi.org/10.2172/932647

Sokhansanj, S., Mani, S., Tagore, S., & Turhollow, A. F. (2010). Techno-economic analysis of using corn stover to supply heat and power to a corn ethanol plant – Part 1: Cost of feedstock supply logistics. *Biomass and Bioenergy*, *34*(1), 75–81. https://doi.org/10.1016/j.biombioe.2009.10.001

Søndergaard, M. M., Fotidis, I. A., Kovalovszki, A., & Angelidaki, I. (2015). Anaerobic co-digestion of agricultural byproducts with manure for enhanced biogas production. *Energy & Fuels*, *29*(12), 8088–8094. https://doi.org/10.1021/acs.energyfuels.5b02373

Song, L., Ma, F., Zeng, Y., Zhang, X., & Yu, H. (2013). The promoting effects of manganese on biological pretreatment with *Irpex lacteus* and enzymatic hydrolysis of corn stover. *Bioresource Technology*, *135*, 89–92. https://doi.org/10.1016/j.biortech.2012.09.004

Sonwai, A., Pholchan, P., & Tippayawong, N. (2023). Machine learning approach for determining and optimizing influential factors of biogas production from lignocellulosic biomass. *Bioresource Technology*, *383*, 129235. https://doi.org/10.1016/j.biortech.2023.129235

Sultana, A., Kumar, A., & Harfield, D. (2010). Development of agri-pellet production cost and optimum size. *Bioresource Technology*, *101*(14), 5609–5621. https://doi.org/10.1016/j.biortech.2010.02.011

Sun, F.-h., Li, J., Yuan, Y.-x., Yan, Z.-y., & Liu, X.-f. (2011). Effect of biological pretreatment with *Trametes hirsuta* yj9 on enzymatic hydrolysis of corn stover. *International Biodeterioration & Biodegradation*, *65*(7), 931–938. https://doi.org/10.1016/j.ibiod.2011.07.001

Sun, L., Li, C., Xue, Z., Simmons, B. A., & Singh, S. (2013). Unveiling high-resolution, tissue specific dynamic changes in corn stover during ionic liquid pretreatment. *RSC Advances*, *3*(6), 2017–2027. https://doi.org/10.1039/C2RA20706K

Swan, J. B., Kaspar, T. C., & Erbach, D. C. (1996). Seed-row residue management for corn establishment in the northern US Corn Belt. *Soil and Tillage Research*, *40*(1–2), 55–72. https://doi.org/10.1016/S0167-1987(96)80006-7

Tabatabaei, M., Aghbashlo, M., Valijanian, E., Kazemi Shariat Panahi, H., Nizami, A.-S., Ghanavati, H., Sulaiman, A., Mirmohamadsadeghi, S., & Karimi, K. (2020). A comprehensive review on recent biological innovations to improve biogas production, Part 1: Upstream strategies. *Renewable Energy*, *146*, 1204–1220. https://doi.org/10.1016/j.renene.2019.07.037

Teng, Z., Hua, J., Wang, C., & Lu, X. (2014). Design and optimization principles of biogas reactors in large scale applications. In F. Shi (Ed.), *Reactor and process design in sustainable energy technology* (pp. 99–134). Elsevier. https://doi.org/10.1016/B978-0-444-59566-9.00004-1

Teymouri, F., Laureano-Perez, L., Alizadeh, H., & Dale, B. E. (2005). Optimization of the ammonia fiber explosion (AFEX) treatment parameters for enzymatic hydrolysis of corn stover. *Bioresource Technology*, *96*(18), 2014–2018. https://doi.org/10.1016/j.biortech.2005.01.016

Tsapekos, P., Kougias, P. G., Treu, L., Campanaro, S., & Angelidaki, I. (2017). Process performance and comparative metagenomic analysis during co-digestion of manure and lignocellulosic biomass for biogas production. *Applied Energy*, *185*, 126–135. https://doi.org/10.1016/j.apenergy.2016.10.081

Um, B.-H., Karim, M. N., & Henk, L. L. (2003). Effect of sulfuric and phosphoric acid pretreatments on enzymatic hydrolysis of corn stover. In B. H. Davison, J. W. Lee, M. Finkelstein, & J. D. McMillan (Eds.), *Biotechnology for fuels and chemicals: The twenty-fourth symposium* (pp. 115–125). Humana Press. https://doi.org/10.1007/978-1-4612-0057-4_9

Uppugundla, N., Da Costa Sousa, L., Chundawat, S. P. S., Yu, X., Simmons, B., Singh, S., Gao, X., Kumar, R., Wyman, C. E., Dale, B. E., & Balan, V. (2014). A comparative study of ethanol production using dilute acid, ionic liquid and AFEX^{TM} pretreated corn stover. *Biotechnology for Biofuels*, *7*, 72. https://doi.org/10.1186/1754-6834-7-72

United States Environmental Protection Agency. (2025). *Overview of the renewable fuel standard program*. https://www.epa.gov/renewable-fuel-standard/overview-renewable-fuel-standard-program

Vedrenne, F., Béline, F., Dabert, P., & Bernet, N. (2008). The effect of incubation conditions on the laboratory measurement of the methane producing capacity of livestock wastes. *Bioresource Technology*, *99*(1), 146–155. https://doi.org/10.1016/j.biortech.2006.11.043

von Blottnitz, H., & Curran, M. A. (2007). A review of assessments conducted on bio-ethanol as a transportation fuel from a net energy, greenhouse gas, and environmental life cycle perspective. *Journal of Cleaner Production*, *15*(7), 607–619. https://doi.org/10.1016/j.jclepro.2006.03.002

Wan, C., & Li, Y. (2010). Microbial pretreatment of corn stover with *Ceriporiopsis subvermispora* for enzymatic hydrolysis and ethanol production. *Bioresource Technology*, *101*(16), 6398–6403. https://doi.org/10.1016/j.biortech.2010.03.070

Wang, C., Zhang, X., Zhao, G., & Chen, Y. (2023). Mechanisms, methods and applications of machine learning in bio-alcohol production and utilization: A review. *Chemosphere*, *342*, 140191. https://doi.org/10.1016/j.chemosphere.2023.140191
Wang, H., Lehtomäki, A., Tolvanen, K., Puhakka, J., & Rintala, J. (2009). Impact of crop species on bacterial community structure during anaerobic co-digestion of crops and cow manure. *Bioresource Technology*, *100*(7), 2311–2315. https://doi.org/10.1016/j.biortech.2008.10.040
Wang, J., Tuo, H., Cheng, P., Chen, F., Li, Y., Yuan, M., Lin, W., Bao, X., Sun, J., Guo, Y., Wang, Z., Wang, Y., Li, H., Mu, X., Zhang, Q., Yu, Y., & Gou, M. (2026). Boosting lignocellulosic bioethanol production: Fermentation of corn stalk using a lignin-reduced *brown midrib* 2 maize mutant. *Biotechnology for Biofuels and Bioproducts*, *19*, 1. https://doi.org/10.1186/s13068-025-02722-6
Wang, K., Chen, J., Sun, S.-N., & Sun, R.-C. (2015). Steam explosion. In A. Pandey, S. Negi, P. Binod, & C. Larroche (Eds.), Pretreatment of biomass (pp. 75–104). Elsevier. https://doi.org/10.1016/B978-0-12-800080-9.00006-2
Wilhelm, W. W., Hess, J. R., Karlen, D. L., Johnson, J. M. F., Muth, D. J., Baker, J. M., Gollany, H. T., Novak, J. M., Stott, D. E., & Varvel, G. E. (2010). REVIEW: Balancing limiting factors & economic drivers for sustainable Midwestern US agricultural residue feedstock supplies. *Industrial Biotechnology*, *6*(5), 271–287. https://doi.org/10.1089/ind.2010.6.271
Wilhelm, W. W., Johnson, J. M. F., Hatfield, J. L., Voorhees, W. B., & Linden, D. R. (2004). Crop and soil productivity response to corn residue removal: A literature review. *Agronomy Journal*, *96*(1), 1–17. https://doi.org/10.2134/agronj2004.1000a
Wilhelm, W. W., Johnson, J. M. F., Karlen, D. L., & Lightle, D. T. (2007). Corn stover to sustain soil organic carbon further constrains biomass supply. *Agronomy Journal*, *99*(6), 1665–1667. https://doi.org/10.2134/agronj2007.0150
Wojtusik, M., Zurita, M., Villar, J. C., Ladero, M., & Garcia-Ochoa, F. (2016). Influence of fluid dynamic conditions on enzymatic hydrolysis of lignocellulosic biomass: Effect of mass transfer rate. *Bioresource Technology*, *216*, 28–35. https://doi.org/10.1016/j.biortech.2016.05.042
World Bank. (2024). *Maximizing finance for development: Impact bonds for water and sanitation in Latin America and the Caribbean*. https://doi.org/10.1596/37100
Xu, G.-C., Ding, J.-C., Han, R.-Z., Dong, J.-J., & Ni, Y. (2016). Enhancing cellulose accessibility of corn stover by deep eutectic solvent pretreatment for butanol fermentation. *Bioresource Technology*, *203*, 364–369. https://doi.org/10.1016/j.biortech.2015.11.002
Xu, H., Sieverding, H., Kwon, H., Clay, D., Stewart, C., Johnson, J. M. F., Qin, Z., Karlen, D. L., & Wang, M. (2019). A global meta-analysis of soil organic carbon response to corn stover removal. *GCB Bioenergy*, *11*(10), 1215–1233. https://doi.org/10.1111/gcbb.12631
Yang, B., & Wyman, C. E. (2008). Pretreatment: The key to unlocking low-cost cellulosic ethanol. *Biofuels, Bioproducts and Biorefining*, *2*(1), 26–40. https://doi.org/10.1002/bbb.49
Zabed, H. M., Akter, S., Yun, J., Zhang, G., Zhao, M., Mofijur, M., Awasthi, M. K., Kalam, M. A., Ragauskas, A., & Qi, X. (2023). Towards the sustainable conversion of corn stover into bioenergy and bioproducts through biochemical route: Technical, economic and strategic perspectives. *Journal of Cleaner Production*, *400*, 136699. https://doi.org/10.1016/j.jclepro.2023.136699
Zhang, X., LeDuc, S. D., Kim, S., Dale, B. E., Zhao, K., Zhou, Y., McCarty, G. W., & Moglen, G. E. (2025). Soil erosion and lateral carbon fluxes from corn stover-derived biofuel. *Scientific Reports*, *15*, 18315. https://doi.org/10.1038/s41598-025-99218-y
Zhao, X., Cheng, K., & Liu, D. (2009). Organosolv pretreatment of lignocellulosic biomass for enzymatic hydrolysis. *Applied Microbiology and Biotechnology*, *82*, 815–827. https://doi.org/10.1007/s00253-009-1883-1
Zhao, Y., Damgaard, A., & Christensen, T. H. (2018). Bioethanol from corn stover – a review and technical assessment of alternative biotechnologies. *Progress in Energy and Combustion Science*, *67*, 275–291. https://doi.org/10.1016/j.pecs.2018.03.004
Zheng, M., Li, X., Li, L., Yang, X., & He, Y. (2009). Enhancing anaerobic biogasification of corn stover through wet state NaOH pretreatment. *Bioresource Technology*, *100*(21), 5140–5145. https://doi.org/10.1016/j.biortech.2009.05.045
Zhong, X., Yuan, R., Zhang, B., Wang, B., Chu, Y., & Wang, Z. (2021). Full fractionation of cellulose, hemicellulose, and lignin in pith-leaf containing corn stover by one-step treatment using aqueous formic acid. *Industrial Crops and Products*, *172*, 113962. https://doi.org/10.1016/j.indcrop.2021.113962
Zhou, A., Guo, Z., Yang, C., Kong, F., Liu, W., & Wang, A. (2013). Volatile fatty acids productivity by anaerobic co-digesting waste activated sludge and corn straw: Effect of feedstock proportion. *Journal of Biotechnology*, *168*(2), 234–239. https://doi.org/10.1016/j.jbiotec.2013.05.015
Zhou, A., Zhang, J., Wen, K., Liu, Z., Wang, G., Liu, W., Wang, A., & Yue, X. (2016). What could the entire cornstover contribute to the enhancement of waste activated sludge acidification? Performance assessment and microbial community analysis. *Biotechnology for Biofuels*, *9*, 241. https://doi.org/10.1186/s13068-016-0659-y
Zhou, J., Liu, W., Deng, Y., Jiang, Y.-H., Xue, K., He, Z., Van Nostrand, J. D., Wu, L., Yang, Y., & Wang, A. (2013). Stochastic assembly leads to alternative communities with distinct functions in a bioreactor microbial community. *mBio*, *4*(2), e00584-12. https://doi.org/10.1128/mBio.00584-12
Zhu, J., Han, M., Zhang, G., & Yang, L. (2015). Co-digestion of spent mushroom substrate and corn stover for methane production via solid-state anaerobic digestion. *Journal of Renewable and Sustainable Energy*, *7*(2), 023135. https://doi.org/10.1063/1.4919404
Zhu, Y., Lee, Y. Y., & Elander, R. T. (2005). Optimization of dilute-acid pretreatment of corn stover using a high-solids percolation reactor. *Applied Biochemistry and Biotechnology*, *124*, 1045–1054. https://doi.org/10.1385/ABAB:124:1-3:1045
Zou, J., Lü, F., Chen, L., Zhang, H., & He, P. (2024). Machine learning for enhancing prediction of biogas production and building a VFA/ALK soft sensor in full-scale dry anaerobic digestion of kitchen food waste. *Journal of Environmental Management*, *371*, 123190. https://doi.org/10.1016/j.jenvman.2024.123190
Zych, D. (2008). *The viability of corn cobs as a bioenergy feedstock*. West Central Research and Outreach Center.

Agricultural & Rural Studies

Article

A Cluster-Analytic Approach to Constraint Typologies and Technical Efficiency among Maize Farmers in Himachal Pradesh

Sudha Kumari [1,*] and Rakesh Singh [1]

1 Department of Economics, Himachal Pradesh University, Shimla 171005, India; rakeshsingh@hpuniv.ac.in
* Correspondence: sharmasudhgarg22@gmail.com

Abstract: The study examines how farmer-perceived production constraints cluster influence technical efficiency and maize output in the low-hill zone of Himachal Pradesh. This paper analyses 432 maize-farming households from Kangra, Mandi, and Hamirpur districts using hierarchical clustering based on the Salama–Quade weighted rank correlation. The study identifies six distinct constraint typologies and evaluates differences in technical efficiency and maize output across clusters. The results show that Cluster C5 (Mechanization-constraint salient) records the highest mean technical efficiency (0.86), while Cluster C4 (Land Fragmentation) has the lowest (0.52), with the overall mean efficiency around 0.60. The study finds that larger landholdings and improved seed adoption are associated with lower inefficiency, while institutional variables do not have a significant effect in the full model. Substantial spatial variation in cluster profiles is observed across districts. Climate variability and seed-access cluster are more prominent in Mandi, while wildlife and pest and disease pressures are more acute in Kangra and Hamirpur. These results indicate that farm performance in low hill maize systems is shaped by locally specific constraint environments. Overall, the findings suggest that performance differences across maize systems are closely linked to the specific constraints faced by farmers, underscoring the need for more targeted policy support.

Keywords: technical efficiency; production constraints; hierarchical clustering; stochastic frontier analysis; ranked data; maize-farming; low-hill agriculture

Check for updates

Citation: Kumari, S., & Singh, R. (2026). A Cluster-Analytic Approach to Constraint Typologies and Technical Efficiency among Maize Farmers in Himachal Pradesh. *Agricultural & Rural Studies*, *4*(2), 26. https://doi.org/10.59978/ar04020008

Received: 12 December 2025
Revised: 4 March 2026
Accepted: 9 March 2026
Published: 22 May 2026

1. Introduction

Maize (*Zea mays* L.) is one of the world's three dominant cereals, alongside rice and wheat, and is an essential component of global agri-food systems (Erenstein et al., 2022; Ranum et al., 2014). Beyond its role as a staple food, maize increasingly drives the food–feed–fuel nexus, supplying raw materials for livestock feed, starch-based industries, and emerging bio-ethanol markets (OECD & FAO, 2025; USDA-FAS, 2025). India is the world's sixth-largest maize producer, cultivating about 9.9 million hectares and contributing nearly 3 percent of global output. However, national yields average only 3.1 t/ha, far below the global mean of 5.8 t/ha (Erenstein et al., 2022; Economics, Statistics, and Evaluation Division, 2023).

Maize production conditions differ substantially across India's agro-ecological regions. In the irrigated plains of the Indo-Gangetic, maize cultivation operates within a relatively input-intensive production environment, supported by extensive irrigation infrastructure, greater mechanization, and stronger market linkages (Jat et al., 2025). By contrast, low-hill production systems are predominantly rain-fed and characterized by small, fragmented landholdings, with limited mechanization and increasing reliance on household labor. Evidence from South Asia shows that mechanization intensity and technology adoption are strongly conditioned by farm size and topography (Aryal et al., 2021). These structural differences suggest that the drivers of maize productivity and technical efficiency vary across agro-ecological contexts, highlighting the importance of context-specific analysis in low-hill production systems.

In India, maize holds particular significance in Himachal Pradesh, where geoclimatic conditions favor its cultivation in the low- and mid-hill agroecological zones. Maize occupies nearly 300 thousand hectares, around one-fifth of the state's gross cropped area, and is predominantly grown as a rain-fed kharif crop on small and highly fragmented holdings (Department of Economics & Statistics, 2022; ICAR–Indian Institute of Maize Research, 2022).These structural characteristics shape the maize cultivation environment in the state, where farmers face both climatic risks and persistent structural and institutional constraints.

As a predominantly rain-fed crop, maize production is particularly sensitive to climatic variability. Recent evidence documents rising rainfall variability, extended dry periods, and irregular monsoon patterns across the state, increasing yield variability and production risk (KC et al., 2022; S. Kumar & Singh, 2023; Srivastava et al., 2021). In addition to rainfall variability, climate change is also altering pest and disease pressure in maize-based farming systems. Shifting temperature and humidity conditions have increased the incidence of insect pests and plant diseases, constraining crop productivity (Ma et al., 2024). Recent field evidence from Himachal Pradesh documents measurable pest-related yield losses under local agro-climatic conditions, highlighting the economic significance of biotic constraints in maize cultivation (Kharwal et al., 2025). These pressures contribute to yield reductions and often require additional pest-management inputs, thereby raising production costs and potentially reducing technical efficiency (Chatterjee et al., 2023). Such interactions underscore the importance of examining pest and disease pressure as a distinct production constraint within location-specific farming systems.

Furthermore, socio-economic and institutional constraints, combined with agronomic risks, place additional burdens on smallholder farmers. Rising input costs and persistently low maize returns compress farm margins, weakening the economic viability of maize cultivation and limiting reinvestment capacity. Within this constrained economic environment, limited access to quality agricultural inputs, uneven coverage of public provisioning systems, and persistent financial barriers further restrict farmers' ability to adopt improved production practices (A. N. Thakur, 2024). Labor scarcity compounds these challenges. As the availability of rural labor declines, dependence on machinery increases; however, financial and structural constraints continue to slow the adoption of mechanization (Aryal et al., 2021; Rajkhowa & Kubik, 2021). In addition, the prevalence of small and fragmented landholdings further reduces the economic feasibility of mechanization, particularly in low-hill production systems (Roy et al., 2022). Beyond economic and structural barriers, ecological pressures further compound production risks. Wild-animal incursions, particularly by monkeys and wild boars, cause frequent and severe crop losses in low- and mid-hill villages (R. K. Thakur et al., 2022).

Even when farmers adopt improved inputs, productivity outcomes often remain unstable because climatic, biological, and socio-economic pressures interact to create complex and overlapping production environments (Kumari & Sharma, 2018; S. Kumar & Singh, 2023; P. Kumar et al., 2025; Negi et al., 2012). As a result, farmers operate within multidimensional constraint systems in which the combined effects of multiple barriers shape both yield performance and technical efficiency (KC et al., 2022). Because production constraints often operate jointly, assessing them in isolation may overlook important variations in productivity and technical efficiency across farms. In the present study, fifteen farmer-ranked production constraints are analysed and subsequently grouped into six empirically derived constraint typologies using hierarchical clustering techniques. Rather than treating individual barriers independently, this approach identifies distinct multidimensional constraint environments that shape farm performance. These typologies provide the analytical basis for examining heterogeneity in technical efficiency and maize output.

This typology-based perspective is consistent with emerging global evidence showing that technical efficiency improves when production, information, and risk constraints are addressed collectively. This is demonstrated by integrated water–nutrient management in the North China Plain, which increased yields by 1.6 percent annually while reducing greenhouse-gas intensity by 17 percent (Wang et al., 2023). A randomized controlled trial in Kenya showed that a bundled package of credit, high-quality inputs, training, and insurance increased maize yields by about 26 percent (Deutschmann et al., 2025). These examples illustrate the importance of adopting a multi-constraint view of productivity shortfalls rather than evaluating individual constraints in isolation. Earlier Indian work combining analysis with frontier modeling, notably Banerjee et al. (2014) for maize systems, demonstrated that efficiency losses arise from interactions among institutional and socio-economic barriers. Similar multivariate approaches in dairy systems (Baral & Bardhan, 2016) further reinforce the role of structural heterogeneity in shaping farm performance. However, rigorous empirical evidence on how multidimensional constraints influence agricultural efficiency remains limited.

Despite extensive literature on climate risks, pest dynamics, and socio-economic barriers, two critical gaps persist. First, there is an empirical understanding of how farmer-perceived constraints cluster together where risks are highly localized and interdependent. Second, few studies have examined how these constraint typologies relate to technical efficiency and output performance. No previous work has integrated ranked-constraint data with both hierarchical clustering and stochastic frontier analysis to explain performance heterogeneity among maize growers in Himachal Pradesh. Given these gaps, the present study examines how different combinations of farmer-perceived production constraints influence technical efficiency and maize output in the low-hill zone of Himachal Pradesh. Using ranked constraint data from maize-growing households, the analysis applies the Salama–Quade weighted rank correlation and hierarchical clustering to identify empirically grounded constraint typologies. These typologies are subsequently linked to efficiency estimates from Stochastic Frontier Analysis (SFA) and to output regressions estimated using ordinary least squares. By integrating ranked-constraint data with frontier modelling, the study provides new empirical evidence on how multidimensional constraint environments shape performance heterogeneity in smallholder maize systems. The paper is organized as follows: Section 2 presents the data and variables; Section 3 explains the methods used in the analysis; Section 4 reports the results; Section 5 discusses the findings; and Section 6 concludes the study.

2. Data and Variables

Primary data were collected during 2023–24 using a multistage cluster sampling design and a structured questionnaire. At the first stage, three districts, Kangra, Mandi, and Hamirpur, were purposively selected as they together account for nearly 48% of the maize-harvested area in Himachal Pradesh. In the second stage, nine major maize-growing blocks were identified, namely Karsog, Gohar, and Sadar (Mandi); Bamson, Bhoranj, and Hamirpur (Hamirpur); and Nurpur, Fatehpur, and Dehra (Kangra). At the third stage, four panchayats were selected from each block, and one village was chosen from each selected panchayat, yielding a total of 36 panchayats and 36 corresponding villages. Finally, 12 farm households were randomly selected from each village, yielding a total sample of 432 maize-producing households. This sampling process ensured the use of diverse maize-growing conditions across the study area. The findings are therefore most directly applicable to districts with similar agro-climatic and production characteristics, and may not fully capture conditions prevailing in areas where maize plays a relatively minor role.

As part of the questionnaire, farmers were asked to rank 15 production constraints affecting maize cultivation from 1 (most severe) to 15 (least severe). The constraint items included high input costs, low maize returns, weather and climate variability, wild animal crop damage, pest and disease pressure, lack of financial assistance, labor shortage, limited access to improved seed, declining soil fertility, inadequate research and extension support, fragmented landholdings, under-mechanized agriculture, crop insurance challenges, irrigation access constraints, and land tenure issues. These constraints were identified based on prior literature, regional agronomic conditions, and pre-survey consultations with farmers and extension officials to ensure contextual relevance.

The socio-economic profile and institutional characteristics of the sampled households are summarized in Table A1. The majority of sampled households fall within the marginal and small farm categories, with approximately 80 percent cultivating less than two hectares. Education levels were limited, with nearly one-third of respondents being illiterate. Although adoption of improved maize seed was relatively high, access to institutional support remained limited. Fewer than one-fifth of sampled households possessed a Kisan Credit Card, and awareness of the minimum support price was low. Crop insurance coverage was negligible. Around half of the households reported receiving technical advice from extension agents or other sources, but participation in soil health card schemes and formal training programs remained minimal. For the efficiency and output estimations, the key production variables included total maize output (kg per household), operated land area (hectares), seed use (kg), fertilizer (kg), labor (hours), machinery use (hours), and pesticide expenditure (₹) during the 2023–24 kharif season. Fertilizer denotes the combined quantity of chemical fertilizer and farmyard manure applied by the household. Table 1 reports summary statistics of maize output and production inputs. Average maize output was 928 kg per household, with substantial dispersion across farms, reflecting heterogeneity in production scale. Operated land averaged 1.07 hectares, consistent with the smallholder structure of maize cultivation in the study area. Considerable variation is also observed in fertilizer use and labor input, suggesting differences in input intensity across farms. Machinery use and pesticide expenditure display high dispersion relative to their means, indicating uneven adoption of mechanization and crop protection practices. Such variability in input use supports the application of a stochastic frontier framework to assess efficiency differentials.

Table 1. Descriptive Statistics for Input–Output Variables.

Variable	Mean	SD	Min	Max
Output (kg)	928.48	527.63	31.00	2,000.00
Land (ha)	1.07	0.78	0.20	4.80
Seed (kg)	24.31	19.76	1.45	152.69
Fertilizer (kg)	1,600.86	1,377.81	0.00	7,495.31
Labour (hours)	137.85	81.91	18.61	435.39
Machinery use (hours)	11.31	13.91	0.00	116.51
Pesticide cost (₹)	44.52	177.12	0.00	2,468.00

Source: Author’s calculations.

3. Methods

3.1. Clustering of Ranked Production Constraints

3.1.1. Salama–Quade Similarity

The units of analysis in this study are maize-growing farm households characterized by their ranked production constraints. These rankings were used to group farmers based on how they prioritized production constraints. The focus is on identifying heterogeneity in production constraints across farmers rather than relying on average rankings. The clusters reported in the results are derived from similarity patterns in these fifteen ranked constraint items and do not represent predefined categories. To quantify the similarity of farmers' ranked constraints, the Salama–Quade (SQ) weighted rank correlation was computed for every pair of farmers. The SQ correlation is appropriate for ordinal data because it assigns greater weight to higher-ranked items, reflecting the fact that higher-ranked constraints have greater informational value than lower-ranked ones (Salama & Quade, 1982). Similarity between farmers' ranked constraint profiles was measured using the Salama–Quade (SQ) weighted rank coefficient. For a pair of farmers i and j, define the pair-specific weight for constraint k as:

$$w_{ik} = \left(\frac{1}{r_{ik}} + \frac{1}{r_{jk}}\right) \quad (1)$$

where $N = 432$ denotes the number of farmers and $K = 15$ the number of constraints and $r_{ik} \in \{1,...,K\}$ denote the rank assigned by the farmer $i(i=1,...,N)$ to constraint $k(k=1,...,K)$. After data cleaning, all farmers provided complete rankings, so K is constant across all farmer pairs. which assigns greater influence to constraints that receive lower ranks (i.e., are perceived as more severe). The Salama–Quade similarity coefficient between farmers i and j is defined as:

$$\rho_{SQ}(i,j) = 1 - \frac{2\sum_{k=1}^{K}(r_{ik}-r_{jk})^2 w_{ijk}}{(K-1)\sum_{t=1}^{K} 2t(K+1-t)} \quad (2)$$

In equation (2), t indexes rank positions from 1 to K. The numerator represents the weighted squared difference in ranks across constraints. The denominator is a normalising constant that depends only on K and ensures comparability across farmer pairs.

3.1.2. Dissimilarity Transformation

The SQ rank-correlation matrix was converted into a farmer-to-farmer dissimilarity matrix for use as input to the clustering analysis. Dissimilarity was computed using a standard linear transformation:

$$d(i,j) = 1 - \frac{\rho_{SQ}(i,j)}{2} \quad (3)$$

This transformation allows farmers with similar constraint rankings to be grouped together in the clustering stage, where $d(i,j) \in [0,1]$, denotes the dissimilarity between farmers i and j, and $\rho_{SQ}(i,j)$ is the Salama–Quade correlation. The resulting 432×432 dissimilarity matrix captures the pairwise distance between all farmers based on the weighted differences in their ranked constraints.

3.1.3. Agglomerative Hierarchical Clustering

After computing the rank-based dissimilarity matrix, farmers were grouped using agglomerative hierarchical clustering (AHC) with the complete-linkage criterion (Johnson, 1967). AHC was preferred because it allows farmer groups to emerge directly from the dissimilarity structure without pre-specifying the number of clusters (Brentari et al., 2016; Fonseca, 2013; Kaufman & Rousseeuw, 2009). Complete linkage was used to emphasize compact, well-separated clusters. Under the complete linkage criteria, the distance between two clusters C_a and C_b is defined as the maximum pairwise dissimilarity between any member of the cluster C_a and any member of C_b:

$$D(C_a, C_b) = \max_{i \epsilon C_a,\ j \epsilon C_b} d(i,j) \quad (4)$$

where $d(i,j)$ represents the dissimilarity between farmers i and j. AHC has been widely applied in social science and agricultural research to develop interpretable farmer typologies based on multidimensional data (Etumnu & Gray, 2020; Rasool & Abler, 2023).

3.1.4. Silhouette Index

In the present study, the optimal number of clusters was assessed using the silhouette index, which measures how well each observation is assigned to its cluster relative to other clusters (Kaufman & Rousseeuw, 2009). Let $C(i)$ denote the cluster containing the farmer i, and let $|\cdot|$ denote set cardinality. Define

$$a(i)=\frac{1}{|C(i)|-1}\sum_{\substack{j\in C(i)\\ j\neq i}} d(i,j) \tag{5}$$

$$b(i)=\min_{C\neq C(i)}\left\{\frac{1}{|C|}\sum_{j\in C}\mathrm{d}(i,j)\right\} \tag{6}$$

For farmer i, the silhouette value $s(i)$is defined as:

$$s(i)=\frac{b(i)-a(i)}{\max\{a(\mathrm{i}),b(\mathrm{i})\}} \tag{7}$$

where $a(i)$ denotes the average dissimilarity of farmer i with all other farmers in the same cluster, and $b(i)$is the lowest average dissimilarity of farmer i to any other cluster of which i is not a member.

In this study, the average silhouette width (ASW) was computed for alternative solutions with 2 to 20 clusters using the SQ-based dissimilarity matrix. The average silhouette width (ASW) is:

$$ASW=\frac{1}{N}\sum_{i=1}^{N}s(i) \tag{8}$$

The final number of clusters was selected based on the highest ASW, with support from a visual inspection of the dendrogram. The corresponding silhouette values for the potential cluster solutions are reported in the results section.

3.1.5. Cluster Stability (Bootstrap Jaccard Coefficient)

To assess whether the identified farmer clusters were robust to sampling variation, cluster stability was evaluated using bootstrap resampling. Cluster stability is important because bootstrap-based evaluation provides a measure of how consistently each cluster appears across repeated samples (Hennig, 2007). The Jaccard coefficient is defined as:

$$\mathrm{J}(\mathrm{C},\widehat{\mathrm{C}})=\frac{|\mathrm{C}\cap\widehat{\mathrm{C}}|}{|\mathrm{C}\cup\widehat{\mathrm{C}}|} \tag{9}$$

where $|\mathrm{C}\cap\widehat{\mathrm{C}}|$ is the number of farmers retained in both the original and bootstrap-derived clusters, and $|\mathrm{C}\cup\widehat{\mathrm{C}}|$ is the total number of unique farmers belonging to either cluster. Jaccard values range from 0 to 1, with higher values indicating stronger stability. Following established practice, higher Jaccard values were interpreted as indicating greater cluster stability, while lower values were treated as reflecting weaker or less consistent groupings (Hennig, 2007). In this study, 200 bootstrap replications were performed using the cluster boot function in the fpc package in R. The Jaccard stability values for all six clusters are reported in the results section.

3.2. Plackett–Luce Model for Aggregate Constraint Ranking

Alongside the cluster-based analysis of ranked constraints, we also estimated a Plackett–Luce (PL) model to obtain an overall ordering of production constraints across all farmers. While clustering captures heterogeneity in how constraints are prioritized, the PL model provides a summary measure of each constraint's relative severity at the aggregate level. In this application, each of the fifteen maize production constraints was treated as an item, and the estimated worth parameters reflect their relative importance as perceived by farmers. The PL model was estimated using the *Plackett-Luce* package in R, which is designed for analyzing full and partial ranking data (Turner et al., 2020). The resulting worth parameters were rescaled to sum to one and are reported along with quasi-standard errors to indicate the precision of the estimates. While the Salama–Quade correlations and agglomerative hierarchical clustering capture heterogeneity in constraint patterns across farmers, the PL model provides a complementary system-level ranking of constraints for the sample as a whole.

3.3. Stochastic Frontier Model

Following the identification of constraint-based farmer clusters, their association with technical efficiency (TE) and maize output was analyzed using a combination of Stochastic Frontier Analysis (SFA) and Ordinary Least Squares (OLS) output models. This sequential framework allows us to distinguish how constraint typologies influence (i) production efficiency and (ii) observed output levels, while controlling for farm, technology, and demographic characteristics.

Efficiency was estimated using a stochastic frontier framework rather than a deterministic Data Envelopment Analysis (DEA) approach. As a deterministic frontier method, DEA attributes deviations from the frontier to inefficiency (Charnes et al., 1978). In contrast, stochastic frontier analysis (SFA) decomposes the composite error into a symmetric noise term and a non-negative inefficiency term, thereby accounting for random shocks and measurement error (Coelli et al., 2005). Agricultural output is subject to stochastic influences such as weather variability and measurement error (Battese, 1992; Bravo-Ureta & Pinheiro, 1993). The stochastic frontier model separates statistical noise from inefficiency within a parametric likelihood framework (Aigner et al., 1977; Kumbhakar et al., 2021; Meeusen & van Den Broeck, 1977).

The stochastic frontier model is specified as:

$$\ln Y_i = f(X_i; \beta) + v_i - u_i \tag{10}$$

where Y_i denotes the maize output of the farm i, X_i is a vector of production inputs including landholding (ha), seed use (kg), fertilizer use (kg), labour (hours), machinery use (hours), and pesticide expenditure (₹), $v_i \sim N(0, \sigma_v^2)$ represents statistical noise, and $u_i \geq 0$ captures technical inefficiency (Aigner et al., 1977). This specification allows deviations from the frontier to reflect both random shocks and inefficiency within a parametric likelihood framework (Kumbhakar et al., 2021).

The production function was specified in both Cobb–Douglas and Translog forms. The Cobb–Douglas specification is written as:

$$\ln Y_i = \beta_0 + \sum_{k=1}^{K} \beta_k \ln X_{ki} + v_i - u_i - u_i \tag{11}$$

where X_{ki} denotes the k-th input for farm i, and K is the number of inputs included in the production function. The Translog specification extends this as:

$$\ln Y_i = \beta_0 + \sum_{k=1}^{K} \beta_k \ln X_{ki} + \frac{1}{2} \sum_{k=1}^{K} \sum_{j=1}^{K} \beta_{kj} \ln X_{ki} \ln X_{ji} + v_i - u_i \tag{12}$$

The Translog provides a flexible second-order approximation to an arbitrary production function, whereas the Cobb–Douglas is a restricted nested specification (Christensen et al., 1973). To assess the appropriate functional form, a likelihood ratio (LR) test was conducted:

$$LR = -2[\ln L(H_0) - \ln L(H_1)] \tag{13}$$

where H_0 represents the restricted Cobb–Douglas model and H_1 denotes the unrestricted Translog specification. The statistic follows a chi-square distribution with degrees of freedom equal to the number of imposed restrictions.

Both specifications were estimated using maximum likelihood methods consistent with standard stochastic frontier practice (Greene, 2003; Kumbhakar et al., 2021).

Alternative distributional assumptions were considered for the inefficiency term. The half-normal specification assumes that inefficiency follows a one-sided normal distribution with zero mean, whereas the truncated-normal specification allows a non-zero mean and greater flexibility (Aigner et al., 1977; Stevenson, 1980). Because efficiency estimates may be sensitive to the assumed distribution, both specifications were estimated and compared using information criteria consistent with standard stochastic frontier practice (Kumbhakar et al., 2021). Model selection was guided by likelihood-ratio tests for nested specifications and by information criteria, including the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC; Akaike, 1974; Schwarz, 1978), with preference given to the more parsimonious specification when competing models provided comparable fit (Burnham & Anderson, 2004).

Technical inefficiency was modeled within the stochastic frontier following the one-step approach of Battese and Coelli (1995). In this framework, the non-negative inefficiency term u_i is specified as:

$$u_i = z_i' \delta + w_i \tag{14}$$

where z_i is a vector of variables influencing inefficiency, δ is a parameter vector, and w_i follows a normal distribution truncated at zero. The vector z_i includes farm, household characteristics, as well as cluster indicators entered as fixed effects in the inefficiency component. Cluster indicators are incorporated in the inefficiency equation rather than the production frontier to allow systematic variation in technical inefficiency across constraint typologies while maintaining a common production function for all farms. All parameters were estimated jointly by maximum likelihood.

Technical efficiency scores were obtained from the estimated frontier using the conditional expectation approach of Jondrow et al. (1982). In this method, inefficiency is estimated as $E(u_i|\varepsilon_i)$, where $\varepsilon_i = v_i - u_i$ is the composed error term. Farm-level technical efficiency is then calculated as:

$$TE_i = exp\,(-u_i) \tag{15}$$

with values bounded between 0 and 1.

3.4. OLS Output Models

In addition to efficiency outcomes, cluster differences in production levels were examined using the logarithm of total maize output as the dependent variable:

$$\ln Y_i = \alpha_0 + \sum_{c=2}^{6} \alpha_c D_{ic} + \gamma' Wi + \varepsilon_i \tag{16}$$

where D_{ic} is a binary indicator equal to 1 if the household i belongs to a cluster c (for $c = 2, \ldots, 6$) and 0 otherwise, with Cluster 1 serving as the reference category. Wi denotes additional covariates introduced sequentially across specifications. These OLS models capture reduced-form output differentials associated with constraint typologies without imposing the stochastic frontier structure. All econometric analyses were implemented in R (R Core Team, 2025).

4. Result

4.1. Cluster Identification and Validation

Agglomerative hierarchical clustering applied to the SQ-based dissimilarity matrix resulted in a six-cluster solution. The average silhouette width (ASW) was 0.43, corresponding to a moderate level of cluster structure (Figure 1). Although this reflects only moderate separation, the distribution of silhouette values provides further insight into the structure of the solution. Approximately 91% of observations exhibit positive silhouette scores, and the median silhouette width is 0.50 (Figure 2), suggesting that most farmers are more similar to members of their assigned cluster than to those in neighboring groups. Further evidence of internal coherence is provided by dispersion diagnostics. As shown in Table 2, the mean within-cluster dissimilarity (0.0042) is substantially lower than the mean between-cluster dissimilarity (0.0114), yielding an overall within-to-between ratio of 0.37. This pattern is observed across clusters and indicates that farmers grouped together share relatively similar constraint profiles compared with those in other clusters. To assess sensitivity to methodological choices, the clustering procedure was repeated using alternative linkage criteria. The resulting partitions show strong agreement with the complete-linkage solution, with adjusted Rand indices of 0.75 (average linkage) and 0.78 (Ward D2; Table A2). These results suggest that the six-cluster structure is not driven by the specific linkage rule employed. The dendrogram structure was preserved across resamples, confirming that the six-cluster solution reflects an underlying and reproducible segmentation of farmers based on constraint patterns (Figure 3).

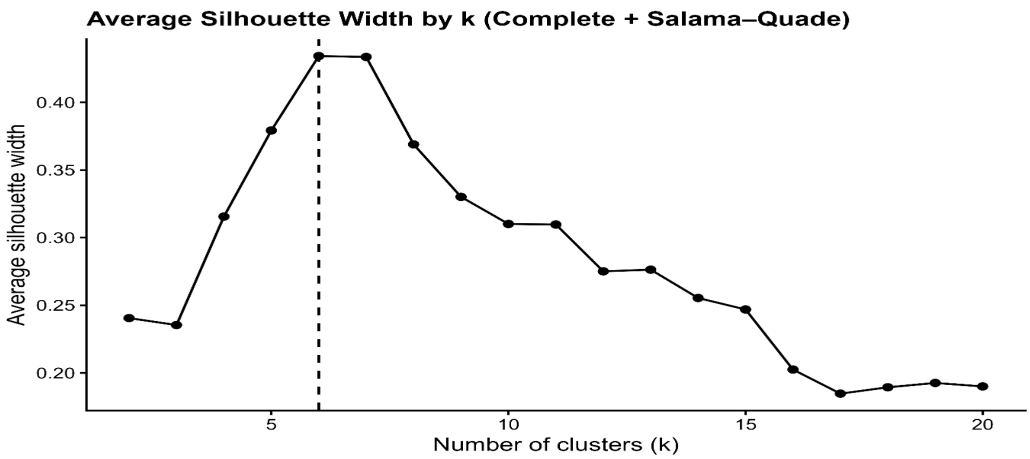


Figure 1. Average silhouette width (ASW) across alternative cluster solutions (k = 2–20).
Note: Clustering performed using agglomerative hierarchical clustering with complete linkage on the rank-based dissimilarity matrix (N = 432). Vertical line indicates the six-cluster solution.

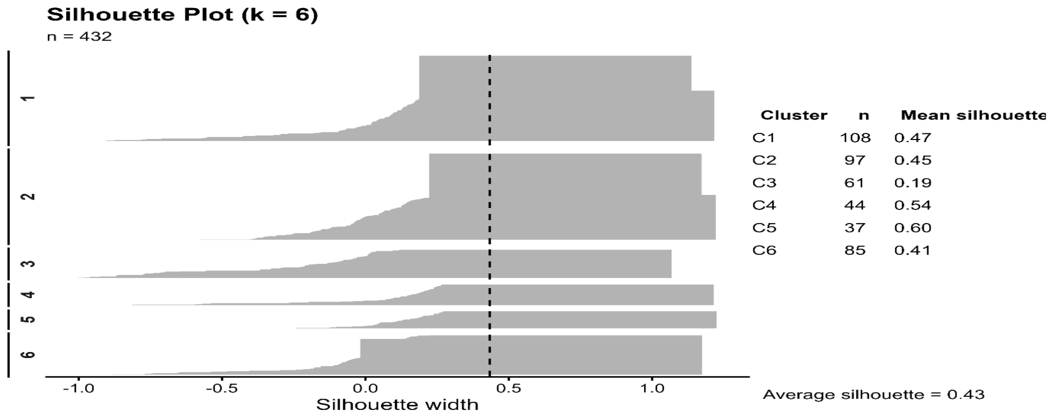


Figure 2. Silhouette widths for the six-cluster solution.
Note: Average silhouette width (ASW = 0.43) indicates moderate separation and internal consistency across clusters.

Table 2. Within- and Between-Cluster Salama–Quade Dissimilarity (k = 6).

Cluster	N	Mean Within-Cluster Dissimilarity	Mean Between-Cluster Dissimilarity	Within/Between Ratio
C1	108	0.0041	0.0115	0.36
C2	97	0.0035	0.0101	0.35
C3	61	0.0061	0.0117	0.52
C4	44	0.0037	0.0119	0.31
C5	37	0.0037	0.0130	0.29
C6	85	0.0044	0.0116	0.38
Overall	432	0.0042	0.0114	0.37

Note: Dissimilarity values are based on the Salama–Quade rank-based distance metric. Ratios below 1 indicate that within-cluster dissimilarity is substantially lower than between-cluster dissimilarity, supporting internal coherence of the identified groups.

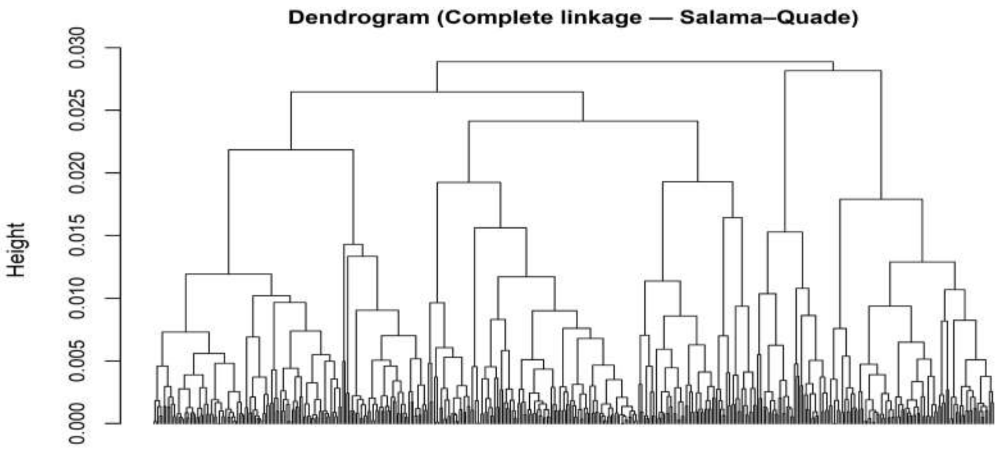


Figure 3. Complete-linkage dendrogram based on rank-dissimilarity matrix.
Note: Clustering performed using agglomerative hierarchical clustering with complete linkage on the rank-based dissimilarity matrix (N = 432). Vertical line indicates the six-cluster solution.

Table 3 presents the Jaccard similarity coefficients, which indicate variation in stability across clusters. Cluster C5 shows the highest stability (0.80), followed closely by C1 (0.80) and C6 (0.76). Cluster C2 exhibits moderate stability (0.74), while C4 (0.67) and C3 (0.65) display comparatively lower, though acceptable, stability levels. Across bootstrap samples, recovered clusters consistently outnumbered dissolved clusters, indicating that the six-cluster structure is reproducible. Overall, the bootstrap results support the robustness of the identified segmentation. The dendrogram structure was preserved across resamples, confirming that the six-cluster solution reflects an underlying and reproducible segmentation of farmers based on constraint patterns.

Table 3. Cluster Stability Measures (Jaccard, Dissolved, Recovered).

Cluster	Jaccard (mean)	Dissolved (B)	Recovered (B)	Stability class
C1	0.80	5	140	High
C2	0.73	21	113	Moderate
C3	0.65	18	39	Low
C4	0.67	32	72	Low
C5	0.80	14	140	High
C6	0.76	24	127	High

Note: Stability based on 200 bootstrap replications; Jaccard = mean similarity between original and resampled clusters. Stability classification follows Hennig (2007).

4.2. Cluster Profiles and Spatial Distribution

The six clusters identified through hierarchical clustering exhibit distinct patterns in production constraint rankings (Figure 4). The six clusters are labelled as follows: C1 (Climate Variability), C2 (Wildlife Damage), C3 (Seed Access), C4 (Land Fragmentation), C5 (Mechanization-Constraint Salient), and C6 (Pest and Disease Pressure). These differences are not only statistical but also associated with variations in farm size, institutional participation, and district-level distribution. Evidence from Table 4, Figure 5, and Figure 6 shows that each cluster reflects a distinct maize production setting. The clusters are therefore discussed below by linking their constraint profiles with socio-economic characteristics and spatial concentration. Cluster C1 is characterized by the prominence of weather and climate variability (mean rank 1.27), followed by declining soil fertility (3.07) and low returns (4.69). The ranking pattern points to environmental instability as the central concern within this group, distinguishing it from clusters shaped primarily by land structure or input access. Most farmers in C1 operate marginal holdings (64%), with a further 22% classified as small farmers, limiting their ability to buffer rainfall fluctuations through scale advantages. Improved seed adoption is moderate (57%), and participation in formal mechanisms such as Kisan Credit Cards (16%) and technical advice (43%) remains modest. Crop insurance coverage is extremely low (3.7%), and with marginal and small holdings dominating this cluster, formal financial protection against climatic shocks remains limited. The cluster is heavily concentrated in Mandi (65%), where exposure to rainfall variability during the kharif season increases production uncertainty. In such settings, climatic risk interacts with small farm size and weak insurance coverage, reinforcing vulnerability to weather shocks.

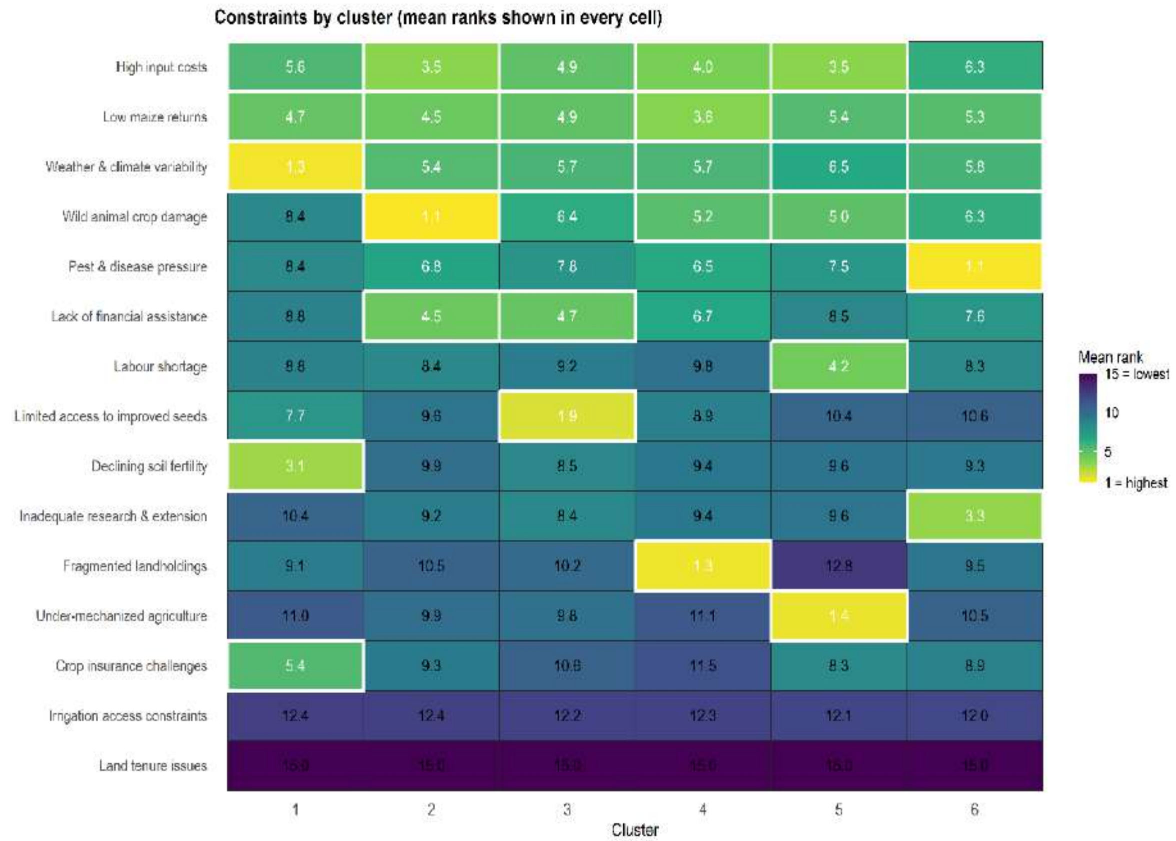


Figure 4. Cluster-wise rank profiles of constraints (C1–C6).
Note: Lower ranks indicate more severe constraints. C1–C6 correspond to the six identified constraint typologies.

Table 4. Farmers' characteristics by cluster (categorical variables; χ^2 tests with Holm-adjusted p-values).

Variable	C1 N = 108	C2 N = 97	C3 N = 61	C4 N = 44	C5 N = 37	C6 N = 85	Overall N = 432	p-value (Holm-adjusted)
Age								0.026
< 35	8 (7.4)	8 (8.2)	8 (13)	9 (20)	5 (14)	4 (4.7)	42 (9.7)	
35–44	24 (22)	18 (19)	8 (13)	9 (20)	6 (16)	15 (18)	80 (19)	
45–54	26 (24)	18 (19)	16 (26)	14 (32)	5 (14)	13 (15)	92 (21)	
55–64	26 (24)	23 (24)	11 (18)	5 (11)	5 (14)	29 (34)	99 (23)	
≥ 65	24 (22)	30 (31)	18 (30)	7 (16)	16 (43)	24 (28)	119 (28)	
Household size								0.2
1–3	15 (14)	16 (16)	7 (11)	3 (6.8)	5 (14)	17 (20)	63 (15)	
4–5	26 (24)	33 (34)	14 (23)	17 (39)	12 (32)	32 (38)	134 (31)	
6–7	44 (41)	31 (32)	21 (34)	17 (39)	10 (27)	25 (29)	148 (34)	
8+	23 (21)	17 (18)	19 (31)	7 (16)	10 (27)	11 (13)	87 (20)	
Education level								0.3
Illiterate	31 (29)	35 (36)	27 (44)	11 (25)	16 (43)	26 (31)	146 (34)	
Primary	16 (15)	16 (16)	13 (21)	11 (25)	6 (16)	19 (22)	81 (19)	
Secondary/High School	50 (46)	38 (39)	19 (31)	22 (50)	13 (35)	37 (44)	179 (41)	
UG	8 (7.4)	5 (5.2)	2 (3.3)	0 (0)	1 (2.7)	3 (3.5)	19 (4.4)	
Postgraduate & above	3 (2.8)	3 (3.1)	0 (0)	0 (0)	1 (2.7)	0 (0)	7 (1.6)	
Holding size (GOI, ha)								< 0.001
Marginal (< 1 ha)	69 (64)	52 (54)	41 (67)	43 (98)	5 (14)	45 (53)	255 (59)	
Small (1–2 ha)	24 (22)	22 (23)	15 (25)	1 (2.3)	15 (41)	24 (28)	101 (23)	

Table 4. *Cont.*

Variable	C1 N = 108	C2 N = 97	C3 N = 61	C4 N = 44	C5 N = 37	C6 N = 85	Overall N = 432	p-value (Holm-adjusted)
Semi-medium (2–4 ha)	15 (14)	20 (21)	5 (8.2)	0 (0)	16 (43)	15 (18)	71 (16)	
Medium (4–10 ha)	0 (0)	2 (2.1)	0 (0)	0 (0)	1 (2.7)	1 (1.2)	4 (0.9)	
Large (≥ 10 ha)	0 (0)	1 (1.0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (0.2)	
Uses improved seed								< 0.001
No	46 (43)	11 (11)	53 (87)	12 (27)	3 (8.1)	4 (4.7)	129 (30)	
Yes	62 (57)	86 (89)	8 (13)	32 (73)	34 (92)	81 (95)	303 (70)	
Aware of MSP								0.2
No	98 (91)	80 (82)	55 (90)	37 (84)	35 (95)	71 (84)	376 (87)	
Yes	10 (9.3)	17 (18)	6 (9.8)	7 (16)	2 (5.4)	14 (16)	56 (13)	
Household insured								0.3
No	104 (96)	91 (94)	61 (100)	43 (98)	37 (100)	82 (96)	418 (97)	
Yes	4 (3.7)	6 (6.2)	0 (0)	1 (2.3)	0 (0)	3 (3.5)	14 (3.2)	
Technical advice adoption								< 0.001
No	62 (57)	36 (37)	35 (57)	16 (36)	8 (22)	52 (61)	209 (48)	
Yes	46 (43)	61 (63)	26 (43)	28 (64)	29 (78)	33 (39)	223 (52)	
Kisan credit card								0.030
No	91 (84)	69 (71)	55 (90)	38 (86)	27 (73)	68 (80)	348 (81)	
Yes	17 (16)	28 (29)	6 (9.8)	6 (14)	10 (27)	17 (20)	84 (19)	
Soil health card								0.034
No	106 (98)	95 (98)	56 (92)	44 (100)	37 (100)	84 (99)	422 (98)	
Yes	2 (1.9)	2 (2.1)	5 (8.2)	0 (0)	0 (0)	1 (1.2)	10 (2.3)	
Agriculture training								0.5
No	105 (97)	91 (94)	59 (97)	43 (98)	37 (100)	83 (98)	418 (97)	
Yes	3 (2.8)	6 (6.2)	2 (3.3)	1 (2.3)	0 (0)	2 (2.4)	14 (3.2)	
District								< 0.001
Kangra	27 (25)	44 (45)	14 (23)	19 (43)	16 (43)	24 (28)	144 (33)	
Mandi	70 (65)	11 (11)	31 (51)	8 (18)	8 (22)	16 (19)	144 (33)	
Hamirpur	11 (10)	42 (43)	16 (26)	17 (39)	13 (35)	45 (53)	144 (33)	

Note: Values denote the number of farmers within each category, with percentages shown in parentheses. P-values are adjusted using the Holm method for multiple comparisons.

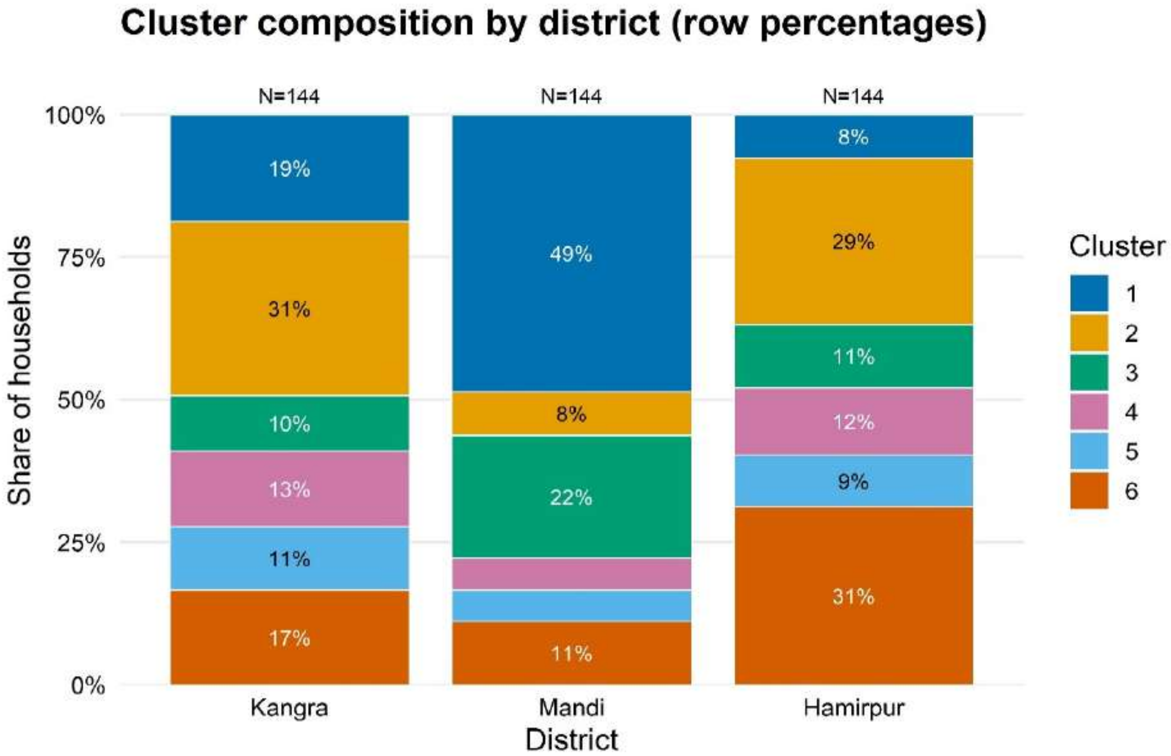


Figure 5. District-wise distribution of constraint clusters (C1–C6).
Note: Each bar represents the proportion of farmers in each constraint cluster across Kangra, Mandi, and Hamirpur districts.

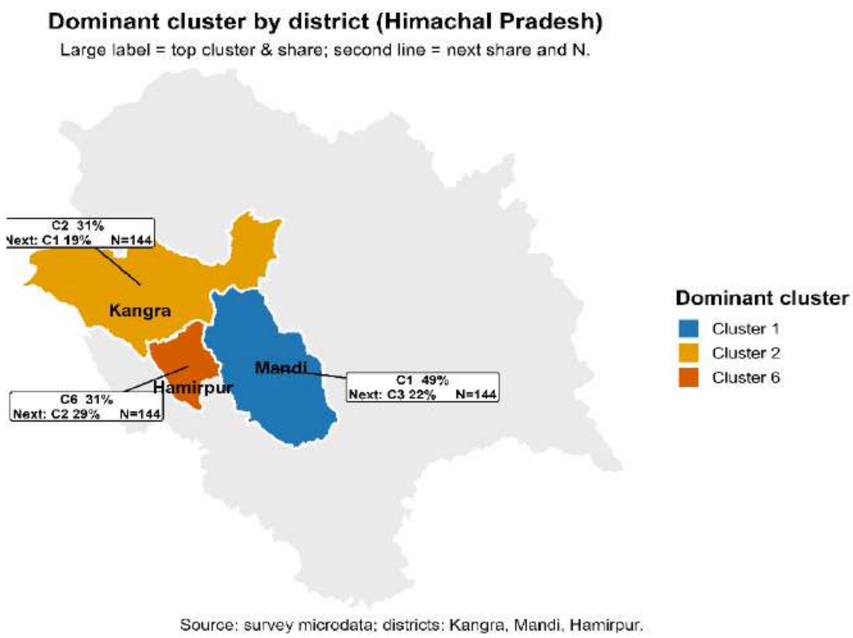


Figure 6. Dominant constraint cluster by district in Himachal Pradesh.

Cluster C2 is dominated by wild-animal crop damage (mean rank 1.08), followed by high input costs (3.51) and constraints in accessing financial support (4.45). This profile reflects a production context in which exposure to wildlife damage, rather than land structure or input scarcity, shapes perceptions of constraint. Farm characteristics indicate relatively active engagement in production. Improved seed adoption (89%) and KCC participation (29%) indicate that farmers in this cluster have access to institutional support. However, formal financial access remains limited for a substantial share of farmers. This combination suggests that crop losses from wildlife occur despite moderate technological and institutional participation. The cluster is concentrated in Kangra (45%) and Hamirpur (43%), with limited presence in Mandi (11%). Reported damage largely involves stray cattle and monkeys rather than forest wildlife, pointing to recurring crop disturbance in village peripheries. Under such conditions, wildlife exposure remains a persistent risk that undermines output stability even where input use is relatively high.

Cluster C3 is characterized by restricted access to improved seed (mean rank 1.85), followed by financial constraints (4.67) and low returns (4.89). This profile identifies input access as the central constraint within this group. Improved seed use is low: 87% of farmers in C3 do not use improved varieties, representing the highest non-adoption rate across clusters. Credit participation is limited, with only 9.8% holding a Kisan Credit Card, and MSP awareness is similarly low. Landholdings are largely marginal (67%) or small (25%), which may restrict purchasing capacity and regular access to quality seed. C3 is most prevalent in Mandi (51%), with smaller shares in Hamirpur (26%) and Kangra (23%). In this context, limited seed adoption appears to be closely linked to farm size and restricted institutional access rather than to environmental conditions. The cluster farm-level conditions in which access to improved seed technology remains the primary constraint. C4 is defined by very small landholdings. Nearly all households (98%) operate marginal farms, making this the most land-constrained group in the sample. Land fragmentation receives the lowest mean rank (1.32), followed by low returns (3.59) and rising input costs (3.95), indicating that operational scale, rather than climatic or input-access factors, defines this typology. Although participation in improved seed use and technical advice is present, the extremely small land base limits opportunities for mechanization and scale efficiency. In this cluster, production capacity is shaped primarily by farm size rather than by access to technology or institutions. C4 is largely present in Kangra (43%) and Hamirpur (39%), with a smaller share in Mandi (18%). Where marginal holdings dominate, limited operational scale continues to constrain productivity over time. Cluster C5 is characterized by mechanization as the most salient constraint (mean rank 1.43), followed by input costs (3.51) and labor shortages (4.22), while wildlife damage ranks lower (5.0). The ranking pattern points to operational scale rather than basic resource scarcity as the defining feature of this group. Farm characteristics indicate comparatively greater production capacity. A substantial share of farmers operate small (41%) and semi-medium holdings (43%), and improved seed adoption is widespread (92%). Participation in technical advice is also high (78%), indicating active engagement with formal agricultural services. In this farm structure, mechanization becomes a scale-related requirement, particularly where timely field operations and labor management become more important as farm size increases. C5 is distributed across Kangra (43%), Hamirpur (35%), and Mandi (22%), suggesting that the pattern is linked to farm structure rather than district-specific conditions. Therefore, mechanization appears as a constraint associated with scale-intensive maize production.

Cluster C6 is characterized by the prominence of pest and disease pressure (mean rank 1.12), followed by research and extension gaps (3.33) and low returns (5.27). The ranking pattern identifies biological stress as the central constraint within this group. The cluster is concentrated in Hamirpur (53%), with smaller shares in Kangra (28%) and Mandi (19%), indicating localized exposure to pest-related risks. Although improved seed use is widespread (95%), recurrent pest incidence appears to undermine yield stability. The simultaneous prominence of low returns suggests that production outcomes are influenced less by input availability and more by biological vulnerability. C6 represents a maize production environment in which pest pressure remains the dominant factor influencing farm performance.

This pattern is consistent with the Plackett–Luce worth estimates, where low returns (15.5%), climate risk (15.3%), input costs (13.7%), wild-animal damage (10.9%), and pest pressure (8.6%) together account for more than 64% of total constraint severity in the sample (Figure 7). These results indicate that while certain pressures are widely shared across farmers, the cluster analysis reveals how different constraints become locally dominant depending on farm structure and district context.

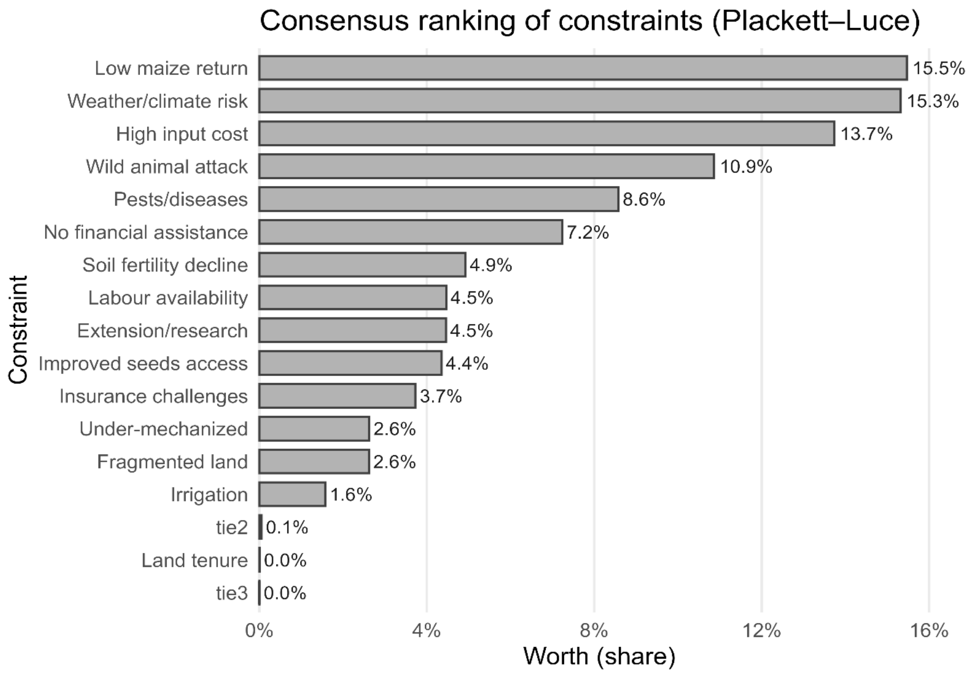


Figure 7. Plackett–Luce worth estimates for ranked production constraints.
Note: A higher value indicates a greater probability that a constraint will be ranked as most severe.

4.3. Model Selection and Robustness

We first compare the Cobb–Douglas and translog specifications using likelihood ratio (LR) tests. Since the Cobb–Douglas model is a restricted version of the translog specification, the LR test evaluates whether the additional squared and interaction terms are jointly significant. As reported in Table A3, the LR statistics reject the restrictions implied by the Cobb–Douglas form under both distributional assumptions, indicating that the translog model provides a statistically better fit in a nested comparison. However, once model complexity is penalized through information criteria, the incremental improvement in fit becomes negligible. The translog specification substantially increases the number of estimated parameters, while yielding only a marginal gain in log-likelihood. The information criteria reported in Table A4 show only minor differences in AIC values across specifications, and the BIC does not provide clear support for the more parameter-intensive translog model. Given the limited increase in log-likelihood relative to the substantial rise in the number of estimated parameters, the simpler Cobb–Douglas specification offers a more stable and interpretable representation of production technology in this sample. This choice is consistent with the principle of parsimony in model selection, which favors the simplest specification that adequately captures the data-generating process (Burnham & Anderson, 2004).

We also compare half-normal and truncated-normal distributions for the inefficiency term. The truncated-normal specification demonstrates a slightly better fit and stable convergence within the one-step framework and is therefore adopted as the preferred distributional assumption. Finally, efficiency estimates are highly consistent across alternative specifications. As shown in Table A5, correlations between efficiency scores exceed 0.95, while Table A6 indicates that mean efficiency levels vary only marginally across models. This stability suggests that the study's substantive conclusions are not sensitive to the choice of functional form or the inefficiency distribution. Accordingly, the Cobb–Douglas truncated-normal model is selected as the preferred specification, with alternative models used for robustness analyses.

Table 5 presents the estimates of the alternative stochastic frontier specifications. The estimated gamma ($\gamma = 0.953$) suggests that a substantial share of the unexplained variation in maize output is attributable to technical inefficiency rather than random shocks. The sum of the input elasticities is approximately 0.48, indicating decreasing returns to scale. This implies that a proportional increase in all inputs would lead to a less-than-proportional increase in total maize output. The estimated parameters for the production inputs are statistically significant, except for fertilizer. Among the production inputs, labor has the largest elasticity (0.172), indicating that labor plays the most prominent role in maize production. A 1% increase in labor is associated with a 0.17% increase in output, highlighting the labor-intensive nature of cultivation practices in the study region.

Table 5. Stochastic Production Frontier Estimates (One-Step CD_TN Model).

Variable	Coefficient	Std. Error	z-value
Intercept	5.985***	(0.251)	23.826
ln(Land)	0.118**	(0.050)	2.349
ln(Seed)	0.118***	(0.041)	2.874
ln(Labour)	0.172***	(0.037)	4.653
ln(Fertilizer)	0.007	(0.013)	0.540
ln(Machinery)	0.029**	(0.013)	2.139
ln(Pesticide)	0.036***	(0.009)	3.861
Statistic		Value	
Sigma²		0.7137	
Gamma		0.953	
Log Likelihood		−268.278	
AIC		588.6	
BIC		694.3	
N		432	

Note: Dependent variable is ln(output). The model is estimated using a truncated-normal stochastic frontier specification. Standard errors are in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Land and seeds are also positively associated with output, with elasticities of 0.118, suggesting that expanding cultivated area and improving seed use increase production. Machinery and pesticides have smaller effects, whereas fertilizer is not statistically significant. Maize production remains labor-intensive, with substantial variation in efficiency across farmers.

4.4 Technical Inefficiency and Cluster Effects

The cluster fixed effects in the inefficiency component reveal meaningful differences across typologies (Table 6). In the baseline specification (M1), C5 is associated with substantially lower inefficiency relative to the reference group (C1), and this relationship remains stable across subsequent specifications. In the final specification (M4), the estimated coefficient increases in absolute value and remains statistically significant, indicating that Cluster 5's efficiency advantage persists after accounting for farm structure, technology adoption, household characteristics, and regional heterogeneity. C6 is associated with a negative, statistically significant coefficient in intermediate specifications; however, this association becomes statistically insignificant once the full set of covariates is included. C2, C3, and C4 do not differ statistically from the reference group in the final model.

Table 6. Cluster-Specific and Determinants of Inefficiency (One-Step SFA Model).

Variable	M1	M2	M3	M4
C2	−0.230	−0.274	−0.299	−0.299
	(0.234)	(0.238)	(0.249)	(0.249)
C3	−0.274	−0.316	−0.348	−0.348
	(0.242)	(0.249)	(0.260)	(0.260)
C4	−0.146	−0.164	−0.146	−0.146
	(0.252)	(0.261)	(0.272)	(0.272)
C5	−2.046*	−2.356*	−2.556**	−3.406**
	(1.181)	(1.215)	(1.275)	(1.290)
C6	−0.068	−0.456*	−0.611**	−0.002
	(0.226)	(0.237)	(0.254)	(0.244)
Small (1–2 ha)		−0.400*	−0.419*	−0.478*
		(0.233)	(0.221)	(0.223)
≥ 2 ha		−0.930**	−0.935**	−0.967**
		(0.349)	(0.350)	(0.354)
Improved seed			−0.536**	−0.641**
			(0.218)	(0.216)
KCC			0.237	0.216
			(0.197)	(0.194)
MSP awareness			−0.209	−0.223
			(0.232)	(0.232)
Technical advice			−0.106	−0.100
adoption			(0.158)	(0.156)
Literate				0.027
				(0.179)
Age				0.001
				(0.006)
Household size				−0.115**
				(0.042)
District (Mandi)				−0.997***
				(0.293)
District (Hamirpur)				0.476*
				(0.212)
Statistic				
Log Likelihood	−306.16	−301.45	−296.05	−268.28
AIC	642.3	636.9	634.1	588.6
BIC	704.0	707.5	712.5	694.3
N	432	432	432	432

Note: C1 is the reference category. Marginal holding (< 1 ha) is the base category. District Kangra is the reference district.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

In the fully specified model (M4), landholding size is negatively associated with inefficiency. Farms of 2 hectares or more show a substantial reduction in inefficiency (−0.967), while holdings of 1–2 hectares also exhibit lower inefficiency (−0.478) relative to marginal farms. This pattern indicates that a larger operational scale is associated with lower inefficiency. The use of improved seed is likewise associated with lower inefficiency (−0.641), suggesting that adoption of enhanced varieties contributes to improved production outcomes. Household size is also negatively associated with inefficiency, suggesting that farms with greater family labor availability tend to perform better. In contrast, institutional variables and demographic characteristics such as KCC access, MSP awareness, technical advice, age, and literacy do not show statistically significant effects in the fully specified model. Regional differences remain pronounced. Farmers in Mandi district exhibit significantly lower inefficiency (−0.997), whereas those in Hamirpur display higher inefficiency (0.476) relative to the reference district. The results indicate that efficiency differences across clusters are largely explained by farm structure and regional characteristics, with C5 remaining statistically distinct in the fully specified model.

Table 7 and Figure 8 present the cluster-wise mean technical efficiency scores based on the final one-step CD–TN model. The overall mean efficiency is 0.604, indicating that maize farmers operate, on average, at about 60% of the potential frontier output given existing inputs. C5 records the highest mean efficiency (0.857; 95% CI: 0.825–0.888), clearly above the sample average. This level is substantially higher than the overall sample mean (0.604), indicating a pronounced efficiency differential. In contrast, Cluster 4 (0.52) and Cluster 6 (0.538) show comparatively lower efficiency levels, while Clusters 1–3 are concentrated around 0.60–0.61. These descriptive differences align with the regression results, where Cluster 5 remains the only group with a statistically distinct inefficiency effect in the fully specified model.

Table 7. Mean Technical Efficiency by Cluster (One-Step CD_TN Model).

Cluster	N	Mean TE	SD	95% CI
1	108	0.60	0.23	[0.56, 0.64]
2	97	0.61	0.23	[0.56, 0.65]
3	61	0.60	0.25	[0.54, 0.66]
4	44	0.52	0.24	[0.45, 0.59]
5	37	0.86	0.10	[0.83, 0.89]
6	85	0.54	0.23	[0.49, 0.59]
Total	432	0.60	0.24	[0.58, 0.63]

Note: Technical efficiency scores are derived from the preferred one-step truncated-normal stochastic frontier model. Values are rounded to 2 decimal places and represent the mean efficiency by cluster, with 95 confidence intervals.

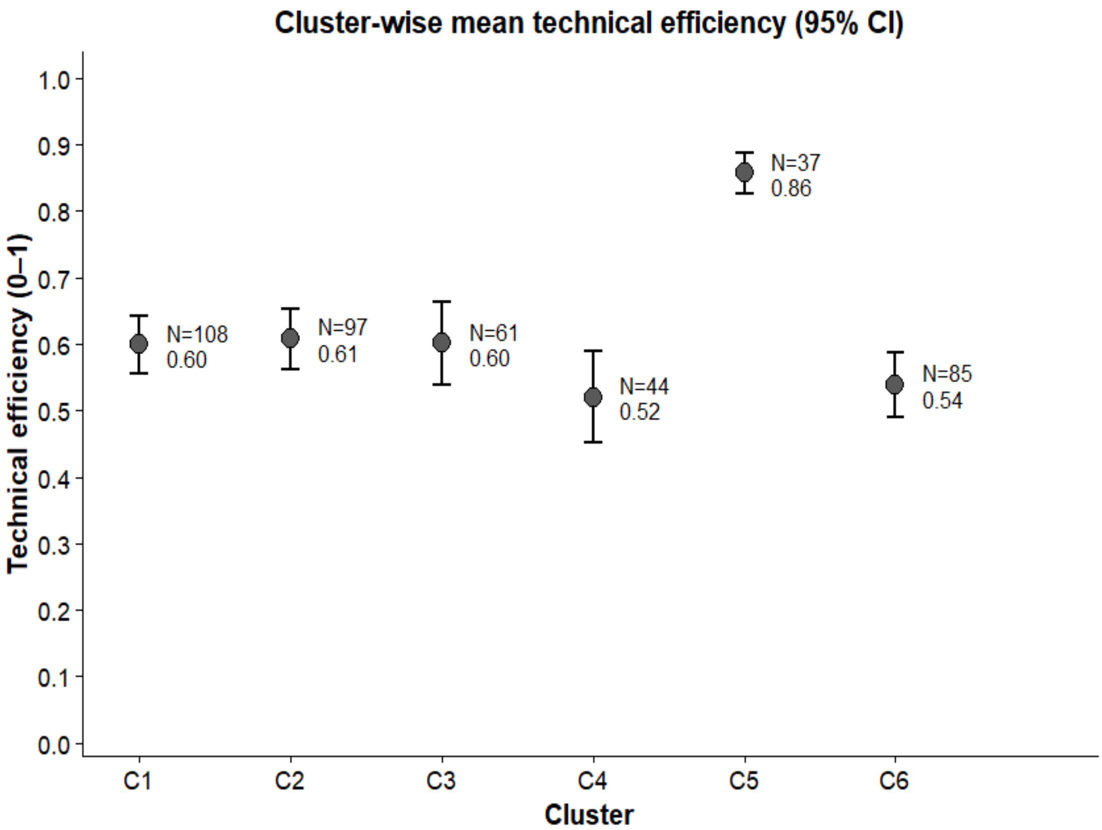


Figure 8. Cluster-specific technical efficiency scores (Model 4).

4.5. Output Effects and Cluster Differences (OLS Estimates)

Table 8 presents OLS estimates of ln(total output) from sequential model specifications. In the initial specification, C5 shows significantly higher output than the reference group (C1), whereas C4 shows significantly lower output. C6 also shows lower output in Models 2 and 3. In the specification including all covariates (M4), only Cluster 5 remains statistically significant (0.496, $p < 0.01$) after controlling for farm structure, household characteristics, and district effects. The coefficients for the remaining clusters are not statistically significant in this specification. Farm size remains positively and significantly associated with output across models. Relative to marginal holdings, farms operating 1–2 hectares show significantly higher output (0.448, $p < 0.01$), while holdings above 2 hectares exhibit an even larger effect (0.703, $p < 0.01$). Improved seed adoption is also positively associated with output (0.315, $p < 0.01$). Household size is also positively and significantly associated with output. District effects remain substantial: Mandi records higher output than Kangra, whereas Hamirpur records lower output. The results indicate that output differences across clusters are largely explained by structural factors such as farm size and district conditions, with only C5 retaining a significant output advantage in the full specification. This finding is consistent with the efficiency estimates, where C5 also remains statistically distinct in the fully specified model.

Table 8. OLS Estimates for ln (Total Output) with HC1 Robust Standard Errors.

Variables	M1	M2	M3	M4
Intercept	6.601^{***} (0.061)	6.404^{***} (0.068)	6.241^{***} (0.087)	5.881^{***} (0.168)
C2	0.107 (0.099)	0.038 (0.095)	−0.052 (0.101)	0.154 (0.097)
C3	−0.035 (0.119)	−0.008 (0.111)	0.118 (0.123)	0.178 (0.115)
C4	-0.433^{***} (0.121)	-0.246^{*} (0.124)	-0.286^{*} (0.120)	−0.127 (0.118)
C5	0.738^{***} (0.079)	0.436^{***} (0.080)	0.340^{***} (0.094)	0.496^{***} (0.100)
C6	−0.096 (0.096)	-0.157^{*} (0.089)	-0.274^{***} (0.096)	−0.031 (0.097)
Holding: Small (1–2 ha)		0.462^{***} (0.071)	0.469^{***} (0.069)	0.448^{***} (0.066)
Holding: ≥2 ha		0.678^{***} (0.072)	0.718^{***} (0.072)	0.703^{***} (0.071)
Improved seed			0.286^{***} (0.093)	0.315^{***} (0.089)
KCC			−0.050 (0.071)	−0.015 (0.072)
MSP awareness			0.097 (0.096)	0.164 (0.100)
Technical advice adoption			−0.022 (0.067)	0.011 (0.065)
Literate				−0.006 (0.070)
Age				−0.002 (0.002)
Household Size				0.053^{***} (0.012)
District (Mandi)				0.289^{***} (0.070)
District (Hamirpur)				-0.191^{**} (0.089)
Observations	432	432	432	432
R^2	0.130	0.255	0.278	0.358
Adjusted R^2	0.120	0.243	0.259	0.333

Note: C1 is the reference category. Marginal holding (<1 ha) is the base category. District Kangra is the reference district. Robust standard errors (HC1) are reported in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

5. Discussion

The results indicate that maize farmers in the low-hill regions of Himachal Pradesh face different sets of production constraints, rather than a single common set of problems. The six clusters identified include climate variability (C1), wildlife damage (C2), seed access (C3), land fragmentation (C4), mechanization constraint salient (C5), and pest and disease pressure (C6), each representing a distinct production situation. These cluster profiles help explain why farmers experience unequal efficiency and output outcomes in the region. This comparison highlights that whether constraints are within or beyond farmers' control plays a key role in shaping production performance. As a result, examining technical efficiency and output through cluster-specific constraint profiles provides a clearer understanding of performance differences than analyses based only on average farm characteristics. In this context, results show that maize-growing households in the low-hill zone operate in distinct constraint environments that shape TE and output in consistent ways.

The highest mean technical efficiency (0.86) is observed in C5 (Mechanization-Constraint Salient). This indicates that farms in this cluster benefit from relatively stronger structural conditions, particularly regarding operational scale and input organization. Although mechanization is identified as a salient constraint, the comparatively high efficiency level suggests that production processes are otherwise well aligned with available resources. Evidence from India indicates that labor and machinery shortages continue to limit the adoption of mechanization (Awachat & Sharma, 2024), while improved access to mechanization services enhances timeliness and reduces labor dependence, contributing to higher technical efficiency (Zhu et al., 2025). These findings suggest that the salience of mechanization in this cluster arises at higher levels of operational intensity rather than from basic resource scarcity.

In contrast to the higher efficiency observed in C5, Cluster C2 (Wildlife Damage) exhibits a lower mean technical efficiency (0.61). This efficiency gap is closely associated with wildlife-related disruptions, which raise labor costs through field guarding and lead to production losses. At the farm level, field evidence from Himachal Pradesh indicates that wild boar and monkeys are the dominant crop-raiding species, with farmers reporting frequent incursions and substantial reductions in harvested produce and household income (Mehta et al., 2018). In low-hill regions, nearly 46% of the maize area is exposed to wildlife damage, resulting in a 16–24% decline in cereal productivity and increased production costs due to fencing and watch-and-ward requirements (R. K. Thakur et al., 2022). Government conflict-mapping exercises and regional Himalayan evidence similarly identify wild boar and monkeys as major threats to maize cultivation, with economic losses increasing as wildlife encounter rates rise (Adhikari et al., 2024; Zoological Survey of India, 2020). Wildlife damage represents a largely exogenous pressure. These recurring production losses help explain why farmers in this cluster struggle to consistently operate near their potential.

C1 (Climate Variability) shows low TE (0.60), a pattern also observed in the Western Himalayan region. KC et al. (2022) document that climate change has increased the irregularity of rainfall, snowfall, and precipitation timing across Himalayan farming systems, disrupting field operations and nutrient uptake in major crops. Recent evidence from the Western Himalaya shows that higher Kharif-season temperatures significantly reduce maize yield (Choudhary & Gupta, 2023). District-level analyses from Himachal Pradesh further demonstrate that increased variability in monsoon rainfall and rising maximum temperatures significantly reduce maize productivity, particularly in Kangra, one of the key maize-producing hill districts (Singh et al., 2024). Such timing disruptions make it more difficult for farmers to achieve potential efficiency, even when input management is adequate.

C3 (Seed Access) shows relatively low technical efficiency (0.60) compared to C5, which appears to be linked to difficulties in securing good-quality seed at the appropriate time. Delayed access to preferred seed varieties affects crop establishment and yield realization, with implications for production performance (Gairhe et al., 2021). Related studies emphasize the importance of timely access to quality seed for improved productivity outcomes (Atreya et al., 2025; Nandi, 2024). These findings suggest that unreliable or delayed access to improved seed may limit farmers' ability to reach potential output levels

The lowest technical efficiency is observed in C4 (Land Fragmentation), with a mean TE of 0.52, followed by C6 (Pest and Disease Pressure) at 0.54. For C6, biological risks such as pest outbreaks and fall armyworm infestation have been shown to generate significant yield losses in maize systems (Ma et al., 2024; Srinivasan et al., 2022), which is consistent with the comparatively lower efficiency observed in this cluster.

In the case of C4, lower efficiency appears linked to structural land fragmentation. Managing multiple small and scattered plots increases supervision requirements and limits the effective use of machinery. Empirical evidence from India consistently shows that land fragmentation raises cultivation costs and reduces farm efficiency (Deininger et al., 2017; Manjunatha et al., 2013), with similar constraints reported in the Western Himalayas (Shukla et al., 2018). These clusters, therefore, operate further from the production frontier relative to other groups, reflecting constraints related to biological risk in C6 and structural scale limitations in C4. While efficiency captures how effectively inputs are converted into output, realized production outcomes remain influenced by additional structural and environmental factors.

The regression results further support these findings. C5 households also record higher output levels, while C6 and C4 exhibit comparatively lower output levels, consistent with the role of constraint environments in shaping outcomes. The results suggest two broad types of production contexts. One is shaped primarily by biophysical and external risks, such as climate variability, wildlife damage, and pest pressure, where short-run adjustment is limited, and efficiency gaps are more evident. The other is characterized by constraints linked to farm structure and management, including seed access, mechanization, and land fragmentation, where performance outcomes appear more responsive to farm-level decisions and institutional support. This distinction helps clarify why efficiency and output differences vary across constraint environments.

The Plackett–Luce ranking results offer complementary evidence on the constraint environment identified through the cluster analysis. They indicate that low returns and rising input costs are widely ranked as the most severe constraints across clusters, while climate-related risks, pest pressure, and wildlife damage also appear among the leading limitations. The concentration of importance around these factors supports the cluster-based finding that production outcomes are shaped primarily by risk- and cost-related pressures rather than by purely farm-specific limitations.

The inefficiency estimates underscore the importance of structural capacity and technology adoption. Larger landholdings are associated with lower inefficiency, consistent with evidence from India showing that land fragmentation and smaller operational scale increase cultivation costs and reduce farm efficiency (Dagar et al., 2021; Deininger et al., 2017). Improved seed use is likewise linked to lower inefficiency, aligning with findings from maize-based systems in India that emphasize the role of input quality in shaping technical performance (Guha & Mandal, 2021). Household size also shows a negative association with inefficiency, suggesting that family labour availability supports more efficient farm management. In contrast, institutional variables do not exhibit statistically significant effects in the full specification, indicating that efficiency differences in this sample are more closely tied to structural and technological factors than to formal program participation.

District patterns reinforce these mechanisms. Kangra and Hamirpur contain more C2 and C6 households, consistent with known wildlife pressure and pest prevalence in these districts, while Mandi's more balanced constraint mix corresponds to its moderate TE and output. The mean technical efficiency in this study is about 0.60. These values indicate substantial unused production potential among maize-growing households. Comparable patterns are reported for hill maize systems in Sikkim, where Guha and Mandal (2021) find an average TE of around 0.55 and wide variation across agro-climatic zones, suggesting that low-hill maize cultivation typically exhibits considerable heterogeneity in efficiency.

These patterns indicate that differences in technical efficiency and output depend on the dominant constraint environment that farmers face. In clusters affected mainly by climatic risk, wildlife damage, and pest and disease pressure, improvements in efficiency and output are more likely to come from better risk management and timely advisory support than from greater input use. By contrast, clusters characterized by operational constraints—such as delayed or poor-quality seed access, land fragmentation, and limited mechanization—may benefit more from improvements in input timing, access to machinery, and coordination of farm operations. Overall, the findings suggest that institutional support is most effective when aligned with the specific constraints faced by farmers, rather than applied uniformly across different production contexts.

The findings indicate that TE differences in the low-hill zone arise from distinct combinations of climatic risk, wildlife interference, seed-timing conditions, land structure, mechanization access, and pest pressure. These constraints function as production environments that shape farmers' ability to convert inputs into outputs, and the alignment among empirical clusters, district patterns, and the existing literature reinforces the credibility of this interpretation.

6. Conclusion

The findings from this study indicate that maize-growing households in the low-hill zone operate in structurally distinct constraint environments, which significantly influence their technical efficiency and production outcomes. These findings point to the importance of constraints beyond household characteristics in shaping agricultural performance.

By combining constraint ranking with efficiency estimation, this study provides additional insight into how different production stresses are associated with productivity differences. Rather than assuming homogeneous production conditions, the analysis highlights how heterogeneity in constraints shapes efficiency and output outcomes. The alignment between empirical clusters and known agronomic, ecological, and institutional challenges enhances the credibility of these patterns and helps clarify where policy attention may be most effective. Together, these findings highlight that differences in technical efficiency and output are closely tied to heterogeneity in production constraints rather than to uniform input use alone. From a policy perspective, the results suggest that interventions are likely to be more effective when tailored to the dominant constraint environment farmers face, rather than applied uniformly across heterogeneous production contexts.

As with any empirical analysis, this study has limitations. The data represent one agricultural season, which limits the ability to capture year-to-year climatic variation or long-term adjustment. Ranked constraints reflect farmers' perceptions, which may differ from measured biophysical conditions. In addition, the study covers three districts, limiting broader generalization. Future research could examine multi-season data, integrate remote-sensing-based agronomic indicators, explore how constraint clusters evolve over time, and assess how targeted interventions reshape efficiency patterns.

CRediT Author Statement: Sudha Kumari: Conceptualization, Methodology, Data Curation, Formal Analysis, Software, Visualization, Writing–original draft, and Writing–review & editing; **Rakesh Singh:** Supervision, Methodology, Validation, and Writing–review & editing.

Data Availability Statement: The data supporting this study is available from the corresponding author upon reasonable request.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflicts of interest.

IRB Statement: This study was conducted in accordance with institutional ethical guidelines for research involving human participants.

Informed Consent Statement: Informed consent was obtained from all participants involved in the study.

Acknowledgments: The authors would like to thank the anonymous reviewers and the editor for their valuable comments and suggestions, which helped improve the manuscript. The authors also gratefully acknowledge the farmers who participated in the survey and generously shared their time, experiences, and insights for this research.

Abbreviations

The following abbreviations are used in this manuscript:

AHC	Agglomerative hierarchical clustering
AIC	Akaike Information Criterion
ARI	Adjusted Rand Index
ASW	Average Silhouette Width
BIC	Bayesian Information Criterion
CD	Cobb–Douglas
DEA	Data Envelopment Analysis
GOI	Government of India
HC1	Heteroskedasticity-consistent standard errors (type 1)
HN	Half-normal
KCC	Kisan Credit Card
LR	Likelihood ratio
MSP	Minimum Support Price
OECD	Organization for Economic Co-operation and Development
OLS	Ordinary least squares
PL	Plackett–Luce
SFA	Stochastic Frontier Analysis
SQ	Salama–Quade
TE	Technical efficiency
TN	Truncated-normal

Appendix A

Table A1. Farmers' characteristics.

Characteristic	n (%)
Age (years, grouped)	
18–35	42 (9.7)
35–44	80 (19)
45–54	92 (21)
55–64	99 (23)
≥65	119 (28)
Household size (grouped)	
1–3	63 (15)
4–5	134 (31)
6–7	148 (34)
8+	87 (20)
Education level	
Illiterate	146 (34)
Primary	81 (19)
Secondary/High School	179 (41)
UG	19 (4.4)
Postgraduate & above	7 (1.6)
Holding size (Government of India (GOI), ha)	
Marginal (<1 ha)	255 (59)
Small (1–2 ha)	101 (23)
Semi-medium (2–4 ha)	71 (16)
Medium (4–10 ha)	4 (0.9)
Large (≥10 ha)	1 (0.2)
Uses improved seed	
No	129 (30)
Yes	303 (70)
Aware of Minimum Support Price (MSP)	
No	376 (87)
Yes	56 (13)
Household insured	
No	418 (97)
Yes	14 (3.2)
Technical advice adoption	
No	209 (48)
Yes	223 (52)
Kisan credit card (KCC)	
No	348 (81)
Yes	84 (19)
Soil health card	
No	422 (98)
Yes	10 (2.3)
Agriculture training	
No	418 (97)
Yes	14 (3.2)
District	
Kangra	144 (33)
Mandi	144 (33)
Hamirpur	144 (33)

Note: n (%) denotes frequency and percentage, respectively.

Table A2. Sensitivity of the Six-Cluster Solution to Linkage Method.

Linkage Method	k	Adjusted Rand Index (vs. Complete Linkage)
Average	6	0.75
Ward.D2	6	0.78

Note. The adjusted Rand index (ARI) measures agreement between cluster partitions. Values between 0.60 and 0.80 indicate strong agreement, suggesting that the six-cluster solution is robust to alternative linkage criteria.

Table A3. Likelihood Ratio Tests (TL vs CD).

Distribution	LogLik (CD)	LogLik (TL)	LR Stat	df	p-value	Decision
Half-normal	−325.446	−304.175	42.543	21	0.004	Reject H0
Truncated−normal	−322.987	−301.906	42.162	21	0.004	Reject H0

Note: LR = 2[LogL (TL) – LogL (CD)]. H0: Cobb–Douglas is sufficient.

Table A4. Model Selection Criteria.

Model	Log Likelihood	AIC	BIC	HQIC
CD_HN	−325.446	668.892	705.508	683.348
CD_TN	−322.987	665.973	706.658	682.035
TL_HN	−304.175	668.349	790.402	716.535
TL_TN	−301.906	665.812	791.933	715.604

Note: Lower AIC, BIC, and HQIC indicate better model fit. TL = Translog; CD = Cobb–Douglas; HN = Half-normal; TN = Truncated-normal.

Table A5. Correlation Matrix of Technical Efficiency Scores.

	TE_CD_HN	TE_CD_TN	TE_TL_HN	TE_TL_TN
TE_CD_HN	1.000	0.997	0.956	0.962
TE_CD_TN	0.997	1.000	0.954	0.964
TE_TL_HN	0.956	0.954	1.000	0.997
TE_TL_TN	0.962	0.964	0.997	1.000

Note: Correlations computed using model-specific TE predictions (N = 432).

Table A6. Distribution of Technical Efficiency by Model.

Model	Mean	SD	Min	Max
TE_CD_HN	0.575	0.216	0.042	0.938
TE_CD_TN	0.612	0.220	0.043	0.938
TE_TL_HN	0.580	0.220	0.043	0.945
TE_TL_TN	0.613	0.223	0.043	0.945

Note: TE is bounded between 0 and 1.

References

Adhikari, J. N., Bhattarai, B. P., & Thapa, T. B. (2024). Correlates and impacts of human–mammal conflict in the central part of Chitwan Annapurna Landscape, Nepal. *Heliyon*, *10*(4), e26386. https://doi.org/10.1016/j.heliyon.2024.e26386

Aigner, D., Lovell, C. A. K., & Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, *6*(1), 21–37. https://doi.org/10.1016/0304-4076(77)90052-5

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

Aryal, J. P., Rahut, D. B., Thapa, G., & Simtowe, F. (2021). Mechanisation of small-scale farms in South Asia: Empirical evidence derived from farm households survey. *Technology in Society*, *65*, 101591. https://doi.org/10.1016/j.techsoc.2021.101591

Atreya, K., Gartaula, H. N., & Kattel, K. (2025). Household seed security: A case of maize and wheat seed systems in the mountains of Nepal. *Agricultural Systems*, *229*, 104419. https://doi.org/10.1016/j.agsy.2025.104419

Awachat, S., & Sharma, D. (2024, November 18–20). A review on impact of agricultural mechanization on farm productivity and factors facilitating the growth of mechanization. 2024 5th International Conference on Data Intelligence and Cognitive Informatics (ICDICI), Tirunelveli, India. https://doi.org/10.1109/ICDICI62993.2024.10810850

Banerjee, H., Goswami, R., Chakraborty, S., Dutta, S., Majumdar, K., Satyanarayana, T., Jat, M. L., & Zingore, S. (2014). Understanding biophysical and socio-economic determinants of maize (*Zea mays* L.) yield variability in eastern India. *NJAS: Wageningen Journal of Life Sciences*, *70–71*(1), 79–93. https://doi.org/10.1016/j.njas.2014.08.001

Baral, S., & Bardhan, D. (2016). Multivariate typology of milk producing households in Uttarakhand hills: Explaining profitability in dairy farming. *Indian Journal of Agricultural Economics*, *71*(2), 160–175. https://doi.org/10.22004/ag.econ.302202

Battese, G. E. (1992). Frontier production functions and technical efficiency: A survey of empirical applications in agricultural economics. *Agricultural Economics*, *7*(3–4), 185–208. https://doi.org/10.1111/j.1574-0862.1992.tb00213.x

Battese, G. E., & Coelli, T. J. (1995). A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics*, *20*, 325–332. https://doi.org/10.1007/BF01205442

Bravo-Ureta, B. E., & Pinheiro, A. E. (1993). Efficiency analysis of developing country agriculture: A review of the frontier function literature. *Agricultural and Resource Economics Review*, *22*(1), 88–101. https://doi.org/10.1017/S1068280500000320

Brentari, E., Dancelli, L., & Manisera, M. (2016). Clustering ranking data in market segmentation: A case study on Italian McDonald's customers' preferences. *Journal of Applied Statistics*, *43*(11), 1959–1976. https://doi.org/10.1080/02664763.2015.1125864

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research*, *33*(2), 261–304. https://doi.org/10.1177/0049124104268644

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, *2*(6), 429–444. https://doi.org/10.1016/0377-2217(78)90138-8

Chatterjee, S., Dhole, A., Krishnan, A. A., & Banerjee, K. (2023). Mycotoxin monitoring, regulation and analysis in India: A success story. *Foods*, *12*(4), 705. https://doi.org/10.3390/foods12040705

Choudhary, T. F., & Gupta, M. (2023). Analyzing long-run and short-run impacts of climate change on wheat and maize yield in the Western Himalayan Region of India. *Climate Change Economics*, *14*(2), 23500197. https://doi.org/10.1142/S2010007823500197

Christensen, L. R., Jorgenson, D. W., & Lau, L. J. (1973). Transcendental logarithmic production frontiers. *The Review of Economics and Statistics*, *55*(1), 28–45. https://doi.org/10.2307/1927992

Coelli, T. J., Rao, D. S. P., O'Donnell, C. J., & Battese, G. E. (2005). *An introduction to efficiency and productivity analysis* (2nd ed.). Springer New York. https://doi.org/10.1007/b136381

Dagar, V., Khan, M. K., Alvarado, R., Usman, M., Zakari, A., Rehman, A., Murshed, M., & Tillaguango, B. (2021). Variations in technical efficiency of farmers with distinct land size across agro-climatic zones: Evidence from India. *Journal of Cleaner Production*, *315*, 128109. https://doi.org/10.1016/j.jclepro.2021.128109

Deininger, K., Monchuk, D., Nagarajan, H. K., & Singh, S. K. (2017). Does land fragmentation increase the cost of cultivation? Evidence from India. *The Journal of Development Studies*, *53*(1), 82–98. https://doi.org/10.1080/00220388.2016.1166210

Department of Economics & Statistics. (2022). *State statistical abstract 2021–22*. Government of Himachal Pradesh. https://himachalservices.nic.in/economics/pdf/State%20Statistical%20Abstract%20-%202021-22.pdf

Deutschmann, J. W., Duru, M., Siegal, K., & Tjernström, E. (2025). Relaxing multiple agricultural productivity constraints at scale. *Journal of Development Economics*, *174*, 103409. https://doi.org/10.1016/j.jdeveco.2024.103409

Economics, Statistics, and Evaluation Division. (2023). *Agricultural statistics at a glance 2023*. Government of India, Ministry of Agriculture and Farmers Welfare. Department of Agriculture and Farmers Welfare. https://nccd.gov.in/uploads/Agricultural_Statistics_at_a_Glance_2023_64c3ac54bf.pdf

Erenstein, O., Jaleta, M., Sonder, K., Mottaleb, K., & Prasanna, B. M. (2022). Global maize production, consumption and trade: Trends and R&D implications. *Food Security*, *14*, 1295–1319. https://doi.org/10.1007/s12571-022-01288-7

Etumnu, C., & Gray, A. W. (2020). A clustering approach to understanding farmers' success strategies. *Journal of Agricultural and Applied Economics*, *52*(3), 335–351. https://doi.org/10.1017/aae.2020.4

Fonseca, J. R. S. (2013). Clustering in the field of social sciences: That is your choice. *International Journal of Social Research Methodology*, *16*(5), 403–428. https://doi.org/10.1080/13645579.2012.716973

Gairhe, S., Timsina, K. P., Ghimire, Y. N., Lamichhane, J., Subedi, S., & Shrestha, J. (2021). Production and distribution system of maize seed in Nepal. *Heliyon*, *7*(4), e06775. https://doi.org/10.1016/j.heliyon.2021.e06775

Greene, W. H. (2003). *Econometric analysis* (5th ed.). Pearson Education.

Guha, P., & Mandal, R. (2021). Technical inefficiency of maize farming and its determinants in different agro-climatic regions of Sikkim, India. *Indian Journal of Agricultural Economics*, *76*(2), 225–244. https://ageconsearch.umn.edu/record/345164/?v=pdf

Hennig, C. (2007). Cluster-wise assessment of cluster stability. *Computational Statistics & Data Analysis*, *52*(1), 258–271. https://doi.org/10.1016/j.csda.2006.11.025

ICAR–Indian Institute of Maize Research. (2022). *Annual report 2022*. Indian Council of Agricultural Research. https://iimr.icar.gov.in

Jat, S. L., Jat, H. S., Rakshit, S., Sharma, P. R., Kumar, B., Kakraliya, M., Gathala, M. K., Bijarniya, D., Kalwania, K. C., Singh, Y., Choudhary, M., & Jat, M. L. (2025). Maize as an alternative to resource-intensive rice: Empirical insights from on-farm participatory study under diverse agricultural scenarios in the Indo-Gangetic Plains of Northwestern India. *Frontiers in Sustainable Food Systems*, *9*, 1700854. https://doi.org/10.3389/fsufs.2025.1700854

Johnson, S. C. (1967). Hierarchical clustering schemes. *Psychometrika*, *32*(3), 241–254. https://doi.org/10.1007/BF02289588

Jondrow, J., Lovell, C. A. K., Materov, I. S., & Schmidt, P. (1982). On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics*, *19*(2–3), 233–238. https://doi.org/10.1016/0304-4076(82)90004-5

Kaufman, L., & Rousseeuw, P. J. (2009). *Finding groups in data: An introduction to cluster analysis* (2nd ed.). Wiley.

KC, K. B., Tzadok, E., & Pant, L. (2022). Himalayan ecosystem services and climate change driven agricultural frontiers: A scoping review. *Discover Sustainability*, *3*, 35. https://doi.org/10.1007/s43621-022-00103-9

Kharwal, D., Sharma, P. K., & Kumar, A. (2025). Yield losses caused by lepidopteran pest complex infesting maize in Himachal Pradesh, India. *Asian Research Journal of Agriculture*, *18*(4),158–165. https://doi.org/10.9734/arja/2025/v18i4775

Kumar, P., Sarda, R., Yadav, A., Ashwani, Gonencgil, B., & Rai, A. (2025). Farmer's perception of climate change and factors determining the adaptation strategies to ensure sustainable agriculture in the cold desert region of Himachal Himalayas, India. *Sustainability*, *17*(6), 2548. https://doi.org/10.3390/su17062548

Kumari, S., & Sharma, H. R. (2018). Farmers' perception on environmental effects of pesticide use, climate change and strategies used in mountain of Western Himalaya. *International Journal of Agricultural Science and Research*, *8*(1), 57–68. https://ssrn.com/abstract=3172041

Kumar, S., & Singh, A. K. (2023). Modeling the effects of climate change on agricultural productivity: Evidence from Himachal Pradesh, India. *Asia-Pacific Journal of Regional Science*, *7*, 521–548. https://doi.org/10.1007/s41685-023-00291-w

Kumbhakar, S. C., Parmeter, C. F., & Zelenyuk, V. (2021). Stochastic frontier analysis: Foundations and advances. In S. C. Ray, R. G. Chambers, & S. C. Kumbhakar (Eds.), *Handbook of production economics* (pp. 1–40). Springer Singapore. https://doi.org/10.1007/978-981-10-3450-3_9-2

Manjunatha, A. V., Anik, A. R., Speelman, S., & Nuppenau, E. A. (2013). Impact of land fragmentation, farm size, land ownership and crop diversity on profit and efficiency of irrigated farms in India. *Land Use Policy*, *31*, 397–405. https://doi.org/10.1016/j.landusepol.2012.08.005

Ma, Z., Wang, W., Chen, X., Gehman, K., Yang, H., & Yang, Y. (2024). Prediction of the global occurrence of maize diseases and estimation of yield loss under climate change. *Pest Management Science*, *80*(11), 5759–5770. https://doi.org/10.1002/ps.8309

Meeusen, W., & van Den Broeck, J. (1977). Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review*, *18*(2), 435–444. https://doi.org/10.2307/2525757

Mehta, P., Negi, A., Chaudhary, R., Janjhua, Y., & Thakur, P. (2018). A study on managing crop damage by wild animals in Himachal Pradesh. *International Journal of Agriculture Sciences*, *10*(12), 6438–6442. https://bioinfopublication.org/pages/article.php?id=BIA0004324

Nandi, S. (2024). Impact of formal seed sources on smallholder farming in India: Evidence from NSS survey using propensity score matching. *Indian Journal of Agricultural Economics*, *79*(3), 709–724. https://doi.org/10.63040/25827510.2024.03.026

Negi, G. C. S., Samal, P. K., Kuniyal, J. C., Kothyari, B. P., Sharma, R. K., & Dhyani, P. P. (2012). Impact of climate change on the western Himalayan mountain ecosystems: An overview. *Tropical Ecology*, *53*(3), 345–356.

Organisation for Economic Co-operation and Development, & Food and Agriculture Organization of the United Nations. (2025). *OECD-FAO Agricultural outlook 2025–2034*. https://doi.org/10.1787/601276cd-en

Rajkhowa, P., & Kubik, Z. (2021). Revisiting the relationship between farm mechanization and labour requirement in India. *Indian Economic Review*, *56*, 487–513. https://doi.org/10.1007/s41775-021-00120-x

Ranum, P., Peña-Rosas, J. P., & Garcia-Casal, M. N. (2014). Global maize production, utilization, and consumption. *Annals of the New York Academy of Sciences*, *1312*(1), 105–112. https://doi.org/10.1111/nyas.12396

Rasool, A., & Abler, D. (2023). Heterogeneity in US farms: A new clustering by production potentials. *Agriculture*, *13*(2), 258. https://doi.org/10.3390/agriculture13020258

R Core Team. (2025). *R: A language and environment for statistical computing* (Version 4.5.1). R Foundation for Statistical Computing. https://www.r-project.org/

Roy, P., Hansra, B. S., Burman, R. R., Bhattacharyya, S., Roy, T. N., & Ahmed, R. (2022). Can farm mechanization enhance small farmers' income? Lessons from Lower Shivalik Hills of the Indian Himalayan Region. *Current Science*, *123*(5), 667–676. https://doi.org/10.18520/cs/v123/i5/667-676

Salama, I. A., & Quade, D. (1982). A nonparametric comparison of two multiple regressions by means of a weighted measure of correlation. *Communications in Statistics - Theory and Methods*, *11*(11), 1185–1195. https://doi.org/10.1080/03610928208828304

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461–464. https://doi.org/10.1214/aos/1176344136

Shukla, R., Chakraborty, A., Sachdeva, K., & Joshi, P. K. (2018). Agriculture in the Western Himalayas – An asset turning into a liability. *Development in Practice*, *28*(2), 318–324. https://doi.org/10.1080/09614524.2018.1420140

Singh, P., Guleria, A., Guleria, C., & Vaidya, M. K. (2024). The Climatic variability and its impact on maize and wheat yield in Himachal Pradesh. *MAUSAM*, *75*(3), 669–678. https://doi.org/10.54302/mausam.v75i3.5882

Srinivasan, T., Shanmugam, P. S., Baskaran, V., Kavitha, Z., Yasodha, P., Ravi, M., Sathiah, N., Rabindra, R. J., Krishnamoorthy, S. V., Vinothkumar, B., Suganthy, A., Balakrishnan, N., Backiyaraj, S., Arulkumar, G., Jeyarani, S., Shanthi, M. Srinivasan, M. R., Justin, G. L., & Prabakar, K. (2022). Estimation of avoidable yield loss in maize (*Zea mays* L.) caused by the fall armyworm *Spodoptera frugiperda* (J.E. Smith) (Noctuidae: Lepidoptera). *Ecology, Environment & Conservation*, *28*(4), 1946–1957. https://doi.org/10.53550/EEC.2022.v28i04.044

Srivastava, R. K., Panda, R. K., & Chakraborty, A. (2021). Assessment of climate change impact on maize yield and yield attributes under different climate change scenarios in eastern India. *Ecological Indicators*, *120*, 106881. https://doi.org/10.1016/j.ecolind.2020.106881

Stevenson, R. E. (1980). Likelihood functions for generalized stochastic frontier estimation. *Journal of Econometrics*, *13*(1), 57–66. https://doi.org/10.1016/0304-4076(80)90042-1

Thakur, A. N. (2024). Input use and socio-economic status of farmers: A case of paddy cultivation in India. *Indian Journal of Agricultural Economics*, *79*(3), 686–694. https://doi.org/10.63040/25827510.2024.03.024

Thakur, R. K., Walia, A., Mehta, K., Kumar, V., & Lal, H. (2022). Economic assessment of crop damages by animal menace in mid hill regions of Himachal Pradesh. *The Indian Journal of Animal Sciences*, *92*(4), 484–491. https://doi.org/10.56093/ijans.v92i4.124173

Turner, H. L., van Etten, J., Firth, D., & Kosmidis, I. (2020). Modelling rankings in R: The PlackettLuce package. *Computational Statistics*, *35*, 1027–1057. https://doi.org/10.1007/s00180-020-00959-3

United States Department of Agriculture Foreign Agricultural Service. (2025). *Biofuels annual: India*. United States Department of Agriculture. https://apps.fas.usda.gov/newgainapi/api/Report/DownloadReportByFileName?fileName=Biofuels%20Annual_New%20Delhi_India_IN2025-0031.pdf

Wang, H., Ren, H., Han, K., Zhang, L., Zhao, Y., Liu, Y., He, Q., Li, G., Zhang, J., Zhao, B., Ren, B., & Liu, P. (2023). Experimental assessment of the yield gap associated with maize production in the North China Plain. *Field Crops Research*, *295*, 108897. https://doi.org/10.1016/j.fcr.2023.108897

Zhu, Y., Wang, G., Du, H., Liu, J., & Yang, Q. (2025). The effect of agricultural mechanization services on the technical efficiency of cotton production. *Agriculture*, *15*(11), 1233. https://doi.org/10.3390/agriculture15111233

Zoological Survey of India. (2020). *Inception report on human–wildlife conflict management baseline survey and development of comprehensive strategy for its mitigation in selected districts of Himachal Pradesh*. Ministry of Environment, Forests and Climate Change.

Agricultural & Rural Studies

Article

Identifying Key Components of Underdevelopment in Rural Bamyan, Afghanistan: An Exploratory Factor Analysis

Sayed Jalil Hashimi [1], Muhammad Asef Shaiq [2,*] and Ali Akbar Barati [3]

1 Agricultural Economics and Extension Department, Bamyan University, Bamyan 1601, Afghanistan; jalilhashimi802@gmail.com
2 Agricultural Economics and Extension Department, Baghlan University, Baghlan 3601, Afghanistan
3 Department of Agricultural Management and Development, University of Tehran, Karaj 77871-31587, Iran; aabarati@ut.ac.ir
* Correspondence: asef.shaiq@gmail.com

Abstract: Rural areas of Bamyan province, Afghanistan, epitomize the paradox of underdevelopment amidst potential. This study employs exploratory factor analysis to systematically diagnose the multidimensional drivers of this crisis. The statistical population of this study was 14,315 households in the study areas. A total of 148 questionnaires were distributed among the rural households using Cochran's formula. Multi-stage sampling was used to collect data. Data was analyzed using SPSS through exploratory factor analysis. Results indicated that the causes of underdevelopment include seven key factors: Socio-economic challenges, lack of basic facilities, insufficient infrastructure, resistance to innovation and change, communication weaknesses, managerial problems, tendency towards urbanization, and geographical challenges, which explained 59.3% of the total variance. The findings of this study can be effectively applied to policy-making and rural development planning in Afghanistan and other developing countries.

Keywords: development; rural economy; rural development; Bamyan province

Check for updates

Citation: Hashimi, S. J., Shaiq, M. A., & Barati, A. A. (2026). Identifying Key Components of Underdevelopment in Rural Bamyan, Afghanistan: An Exploratory Factor Analysis. *Agricultural & Rural Studies*, *4*(2), 11. https://doi.org/10.59978/ar04020009

Received: 12 January 2026
Revised: 17 March 2026
Accepted: 23 March 2026
Published: 25 May 2026

1. Introduction

Approximately half of the world's population resides in rural areas, and many of them suffer from the pervasive issue of poverty (Adimassu et al., 2013). Given that a significant portion of the population in developing countries resides in rural areas, the importance of rural development and its vital role in the advancement and progress of these countries is undeniable (Adamowicz & Zwolińska-Ligaj, 2020). Rural development is one of the concerns of countries and governments, especially developing countries (Badri et al., 2021; Dixon, 2015). In other words, rural development is a fundamental and essential approach for achieving sustainable development (Fayez et al., 2022). However, this process faces significant threats, including environmental instability such as drought, resource scarcity, social unrest, unemployment, and economic stagnation (Ayoo, 2022). Rural development involves adapting rural areas to social and political institutions, human behavior, and community participation in the development process (Borodina & Prokopa, 2019). Achieving sustainable rural development is crucial for fulfilling the Sustainable Development Goals and improving the quality of life for rural populations. In Afghanistan, villages are recognized as important centers of social, cultural, and economic diversity (Ghazali & Zibaei, 2018). Despite this potential, rural areas in Afghanistan face numerous challenges that exacerbate poverty, social injustices, and the unsustainable depletion of natural resources, negatively impacting rural communities. The consequences of the underdevelopment of rural areas, such as widespread poverty, inequality, poor health, unemployment, and migration, have led to attention to rural development (Shaiq et al., 2022; Yar & Yasouri, 2024). The goal of rural development is to enhance the quality of life (Long et al., 2022; Pain & Hansen, 2019; Talebpour et al., 2022a) and to achieve a healthy lifestyle by addressing all the basic needs of rural communities (Qi et al., 2017). Inequality, unemployment, poverty, migration, lack of investment, absence of skills and creativity, disregard for new ideas by the people, and inadequate rural planning are the most significant factors affecting rural underdevelopment (Ahmadikish et al., 2017; Yar et al., 2022).

Development, in its broad sense, means the enhancement of the material and spiritual level of human society and the creation of suitable conditions for a healthy life for all members of the community (Flint, 2013). The ultimate goal of development is to improve the quality of life for everyone (United Nations Digital Public Infrastructure, 2017). Hence, efforts to achieve development should be such that they encompass the interests of the majority of people. Given the importance of rural development for the overall progress of the country, examining the causes of underdevelopment in the villages of Bamyan Province is of particular significance (Javadi et al., 2013). Some previous research has addressed the dimensions of this issue. But none of the previous studies in the country have addressed the causes of the underdevelopment of rural areas from the perspective of villagers. Shaiq et al. (2021), in a study titled "Investigating Afghanistan's rural development challenges" found that economic challenges, such as lack of credit and investment in villages; social challenges, such as low participation of women in rural activities; environmental challenges, including excessive exploitation of groundwater resources; physical challenges, such as weakness or lack of urban-rural networks; and finally managerial challenges, especially the predominance of unskilled labor in rural economic activities, are among the most effective challenges to rural development in Afghanistan. Zinchuk et al. (2018) investigated the challenges of sustainable economic development in rural areas. They found that mechanisms for sustainable development policy in rural areas promise directions for developing local areas and innovative solutions for environmental and social problems. Savari and Maymand (2013), in a study titled "Barriers of sustainable rural development from the perspective of experts," found that the main obstacles to sustainable rural development in Kermanshah province were five factors: physical-structural, investment problems, poor policy-making, agricultural production risk, and skill shortage. Kalantari et al. (2008), in their research titled "Major challenges of rural development in Iran," concluded that the lack of diversity in economic activities, low income levels, rural poverty, absence of appropriate models, lack of coordination in rural development programs, inadequate use of suitable technology in agriculture, lack of awareness among villagers, small size of villages, dominance of unskilled labor in rural activities, and poor rural infrastructure are significant challenges to rural development. Although previous studies have examined various aspects of rural underdevelopment, few have explored the issue from the perspective of villagers in this specific region, and none have applied exploratory factor analysis (EFA). This study addresses this gap by identifying latent factors influencing rural underdevelopment, providing a conceptual and empirical framework that can inform both local interventions and broader research on rural development.

Although the Afghan government has tried to meet the basic humanitarian needs of its people, it is essential to provide the necessary tools and opportunities to reduce poverty, especially at the community level in rural areas. Undoubtedly, the long-term vision for rural and agricultural development in Afghanistan requires ensuring the social, economic, and political welfare of rural communities, particularly for the poor and vulnerable, and necessitates the integration of rural communities into the national economy. Bamyan province, recognized as a region rich in cultural and natural resources in Afghanistan, requires a thorough examination of the causes of underdevelopment in its rural areas. Evidence indicates that despite the abundant potential of this area, including suitable water and soil resources and a rich cultural history, issues such as poverty, unemployment, and inadequate access to essential services persist. The causes of this phenomenon may stem from various factors, including weak infrastructure and insufficient investment in the agricultural and livestock sectors. In other words, despite numerous studies in the field of rural development, no research has yet utilized exploratory factor analysis to identify the structural factors underlying underdevelopment from the perspective of rural residents in Bamyan. This is particularly significant for Bamyan province, given its unique geographical, cultural, and social characteristics. The province requires approaches that are specifically aligned with local conditions. Exploring the opinions of rural residents can lead to the design of more effective and targeted development programs. Without a thorough understanding of local perspectives, any planning and policymaking efforts are likely to face challenges and setbacks. On the other hand, the use of exploratory factor analysis enables researchers to identify and understand the complex structures that influence underdevelopment. This statistical method can reveal hidden relationships among various variables and contribute to a deeper analysis of the data. The lack of such analysis in the central rural areas of Bamyan undermines the opportunity to gain a deeper understanding of the needs and challenges of this region. Therefore, this study can assist in identifying the fundamental factors of underdevelopment through exploratory factor analysis and facilitate the improvement of the quality of life for rural residents. Additionally, the findings of this research could enrich the literature on rural development and lead to positive and sustainable changes within the Bamyan community.

Therefore, this research aims to address the following questions:

- What are the main components of the causes of underdevelopment?
- Which identified factor is the priority in terms of impact on the underdevelopment of these areas?

2. Materials and Methods

2.1. Study Area

Bamyan Province is one of the mountainous provinces in central Afghanistan, classified as a second-degree province (see Figure 1). Located 190 kilometers northwest of Kabul, it rests on the northern slopes of the Baba mountain range. Bamyan borders Samangan to the north, Sar-e Pol to the northwest, Maidan Wardak, Ghazni, and Daykundi to the south, Parwan and Baghlan to the east, and Ghor and Daykundi to the west. The total area of Bamyan is 18,029 square kilometers, representing 2.8% of Afghanistan's total area. The main crops cultivated in this province include wheat, barley, potatoes, and beans. Bamyan is one of the provinces with the least agricultural produce in the country. A significant portion of the land is barren and inaccessible due to a severe water shortage, small land parcels, severe food insecurity, and poor soil quality (Shaiq et al., 2026).

Figure 1. Study area and its location in Afghanistan.

2.2. Data Collection Method

The present study employed a quantitative approach and utilized a non-experimental (survey) design. The statistical population for this research consists of the heads of families in the villages of the center of Bamyan Province, with a total of 14,315 families in the center of Bamyan Province. The population of the villages of the center of this province in the current year is 100,204 people, and the number of households living in these villages is 14,315 (Shaiq et al., 2026). Based on Cochran's formula, 148 questionnaires were distributed randomly to heads of households from the mentioned population in various villages, and they responded to the questionnaires. In other words, given the geographical dispersion of households across different villages in the study area, a multi-stage sampling method was employed. In the first stage, several villages within the study area were selected. Subsequently, the questionnaires were distributed among households in the selected villages.

$$n_0 = \frac{Z^2 pq}{e^2} = \frac{(1.96)^2\ 0.5}{(0.08)^2} = 150 \quad (1)$$

$$n = \frac{n_0}{1 + \frac{(n_0 - 1)}{N}} = \frac{150}{1 + \frac{149}{14315}} = 148 \quad (2)$$

The primary tool used in the present research was a questionnaire. To develop the questionnaire, an initial review of the existing literature and theories related to the research topic was conducted, resulting in the creation of a preliminary questionnaire. This questionnaire comprised two main sections. The first section focused on collecting personal information from the respondents, while the second section contained questions regarding the causes of underdevelopment in rural areas. Subsequently, the questionnaire was reviewed and evaluated by the members of the agricultural economics and extension department at Bamyan University. After incorporating their feedback and suggestions, the final version of the questionnaire was designed as a closed-ended instrument, consisting of 25 questions along with demographic inquiries. The reliability of the questionnaire was confirmed by calculating Cronbach's Alpha coefficient, which yielded a value of 0.87 for the entire instrument, indicating high internal consistency. The measurement scale used for data collection varied depending on the type of data, employing nominal and ordinal scales (a five-point Likert scale; Mohd Rokeman, 2024).

2.3. Method of Data Analysis

After collecting the data, it was analyzed using inferential methods through exploratory factor analysis (Widaman & Helm, 2023) via the SPSS software version 26. The descriptive section, including frequency, standard deviation, and mean, was also analyzed using the SPSS software. It is worth noting that factor analysis is a method that considers all variables simultaneously (Goretzko et al., 2021). In this context, to identify the components of the causes of underdevelopment in the rural areas of Bamyan Province, a total of 21 factors were included in the exploratory factor analysis. To assess the adequacy of the sample size, the KMO coefficient was utilized, and to determine the correlation of the data for conducting factor analysis, Bartlett's test was employed. The results of this test are reported in the findings section of the research.

3. Findings

According to Table 1, 68 individuals (45.9%) are under 30 years old, 35 individuals (23.6%) are between the ages of 30 and 40, 31 individuals (20.9%) are between the ages of 41–50, 14 individuals (9.6%) are above 50 years old. In other words, the majority of respondents are under 30 years old, making up 45.9% of the total. The majority of respondents are married, with 96 individuals making up 69.9% of the total.

Table 1. Personal characteristics of respondents (N = 148).

Characteristics	Frequency	Percentage (%)
Age		
Less than 30	68	45.9
30–40	35	23.6
41–50	31	20.9
Above 50	14	9.6
Total	148	100.0
Marital status		
Married	52	35
Single	96	65

Source: Research findings.

As stated in the research methodology, exploratory factor analysis was employed to identify the components of the causes of underdevelopment in rural areas. In utilizing this method for data analysis, two critical assumptions must be considered, which are:

Test: KMO: This test indicates the adequacy of the sample size in the research, considering the number of variables being studied. The general rule, with an acceptable criterion in most scientific research and in statistical texts, is considered to be (KMO $\geq$ 0.7).

Bartlett's test: In this test, it indicates the degree of correlation between the variables of the study. The null hypothesis in this test states that there is no complete correlation between the studied variables, while the alternative hypothesis claims that there is a correlation among them. If the significance value (Sig) is (0.05 $\geq$ Sig), the null hypothesis (H0) is rejected, and the model is usable. Table 2 shows the KMO and Bartlett test results at a 95% confidence level.

Table 2. KMO and Bartlett's test results.

Amount of KMO	Amount of Bartlett	Degree of freedom	significant
0.766	999.835	300	0.000

Source: Research findings.

To perform factor analysis, it is first necessary to determine that the data have the minimum required correlation for factor analysis. For this purpose, KMO and Bartlett statistics are used to assess the sample size for conducting factor analysis on the causes of underdevelopment in rural areas. The KMO value obtained is 0.766, which indicates the adequacy of the sample size for performing factor analysis. The Bartlett statistic value, using the chi-square approximation, is 999.835 with 300 degrees of freedom, which is significant at a 0.000% error level. Hence, there is sufficient evidence to reject the null hypothesis of the unity of the correlation matrix among the variables intended for factor analysis. Therefore, it can be assumed that the data have the minimum required correlation for conducting factor analysis. The first finding of this model that should be noted is the total variance explained table. This table first reveals the number of latent factors based on the criterion of eigenvalues greater than or equal to one (1 ≤ Eigenvalue) and then specifies their individual and collective explanatory power concerning the observable variables.

Table 3 presents the titles related to each factor and the percentage of variance calculated for each factor. Also, Table 4 shows the factor loadings of each factor with its corresponding variables. The special value for the first factor, titled economic and social problems, is 2.822, which accounts for 11.286% of the total variance. This factor includes four variables, with two variables having the most significant impact on the underdevelopment of rural areas in the study area: the low provision of facilities to the villagers, with a factor loading of 0.738, and the low income levels of families, with a factor loading of 0.627. The second factor, titled lack of basic facilities and absence of services, has a special value of 2.410, which explains 9.639% of the total variance. This factor includes four variables, with two variables having the most significant impact on the underdevelopment of rural areas in the study area: lack of access to telecommunications and the internet, with a factor loading of 0.682, and the absence of educational and health programs for villagers, with a factor loading of 0.596.

Table 3. Special values and explained variances of each factor.

Row	Agent name	Special value	Percentage of variance (%)	Cumulative percentage (%)
1	Economic and social problems	2.822	11.286	11.286
2	Lack of basic facilities and absence of services	2.410	9.639	20.925
3	Lack of infrastructure and resistance to innovation	2.259	9.037	29.962
4	Resistance to changes and weak communication.	1.966	9.862	39.862
5	Management issues of natural resources	1.929	7.715	47.539
6	Migration and the tendency towards urbanization	1.558	6.233	53.772
7	Geographical and communication problems	5.499	5.559	59.331

Source: Research findings.

Table 4. Categorical variables in each factor and their Factorial Load.

Agent name	Categorical variables in each factor	Factorial Load
Economic and social problems	Providing low facilities to the rural people	0.738
	Low-income levels of families	0.627
	Low willingness of villagers to invest in the village	0.611
	Existence of conflicts in the village	0.551
Lack of basic facilities and absence of services	Lack of access to telecommunications and the internet	0.682
	Absence of educational and health programs for villagers	0.621
	Low level of access to educational services	0.596
	Lack of access to agricultural and livestock loans	0.576
Lack of infrastructure and opposition to innovation	Lack of access to a stable electricity network in the village	0.751
	Opposition of elders to new agricultural idea	0.614
	Lack of public transportation facilities in the village	0.614
	Lack of health and medical facilities in the village	0.541
Resistance to changes and weak communication	Low acceptance of changes and technology by the villagers	0.783
	Unsuitable environment in the village for investment	0.662
	Weak communication and collaboration of the people with government agencies	0.607
Management and natural resource	Unemployment in the village	0.677
	Water resource scarcity	0.665
	People in the village not participating in public works	0.510
Migration and the tendency towards urbanization.	The interest and migration of villagers to cities	0.710
	Resistance to new agricultural methods	0.603
Geographical and communication problems	The great distance of the village from the surrounding village	0.693

Source: Research findings.

The third factor, titled lack of infrastructure and resistance to innovation, has a special value of 2.259, which explains 9.037% of the variance. This factor includes four loaded variables, with two variables having the most significant impact on the underdevelopment of rural areas in the study area: lack of access to stable electricity networks, with a factor loading of 0.751, and the opposition of village elders to new agricultural ideas, with a factor loading of 0.624. The fourth factor, titled resistance to changes and weak communication, has a special value of 1.966, which explains 7.862% of the variance. This factor includes three loaded variables, with one variable, the villagers' reluctance to adopt changes and technology, having the most significant impact on the underdevelopment of rural areas in the study area, with a factor loading of 0.662. The fifth factor, titled management issues and natural resources, has a special value of 1.929, which explains 7.715% of the variance. This factor includes three loaded variables, with one variable, unemployment in the village, having the most significant impact on the underdevelopment of rural areas in the study area, with a factor loading of 0.677. The sixth factor, titled migration and the tendency towards urbanization, has a special value of 1.558, which explains 6.233% of the variance. This factor includes two loaded variables, with one variable, the interest and migration of villagers to cities, having the most significant impact on the underdevelopment of rural areas in the study area, with a factor loading of 0.710. The seventh factor, titled geographical and communication problems, explains 5.959% of the variance. This factor includes two loaded variables, with one variable, the great distance of the village from surrounding villages, having the most significant impact on the underdevelopment of rural areas in the study area.

Finally, as a result of using the exploratory factor analysis model, seven latent factors have been clearly identified: economic and social problems, lack of basic facilities and absence of services, lack of infrastructure and resistance to innovation, resistance to changes and weak communication, management issues of natural resources, migration and urbanization tendencies, and geographical and connectivity problems. Collectively these factors contribute to the underdevelopment of rural areas.

4. Discussion

Villages play a vital role in the economic and social development of Afghanistan as key centers for the production and supply of food (Sharifi & Karim, 2024). These areas not only house a significant portion of the population but also possess rich natural resources and diverse cultures, making them potential hubs for sustainable development. The importance of rural development lies in its capacity to improve infrastructure, enhance access to educational and health services, and boost economic activities, thereby elevating the quality of life for residents and preventing excessive migration to urban areas. The villages of Bamyan Province, with their unique natural and cultural attractions, contribute to the preservation of local cultures and promote the social and economic sustainability of both the province and the country. Consequently, investing in rural development not only benefits the inhabitants of these areas but also strengthens the national economy and enhances the overall quality of life on a larger scale (Lali et al., 2020).

The present study aims to identify the main components of the causes of underdevelopment in rural areas of Bamyan Province. This research, as the first of its kind, investigates the causes of underdevelopment in these areas from the perspective of household heads in rural communities. Despite the fact that the villages of Bamyan Province possess significant potential, including rich natural resources, a vibrant culture, and a young workforce, the findings of this study indicate the existence of multiple barriers that hinder the realization of sustainable development in these regions. Findings showed that seven hidden factors have been identified as the main contributors to the underdevelopment of these regions. These factors include economic and social issues, a lack of basic facilities and services, insufficient infrastructure, resistance to innovation, resistance to change, and weak communication, management issues regarding natural resources, trends of migration and urbanization, and geographical and connectivity challenges. The findings of this research align with the results of previous studies (Kalantari et al., 2008; Savari & Maymand, 2013; Shaiq et al. 2021; Zinchuk et al., 2018).

The study indicates that economic and social problems have the highest impact on the underdevelopment of villages. Undoubtedly, insufficient provision of facilities for rural residents exacerbates low family income levels, making it challenging for them to achieve financial stability. This economic pressure, combined with the low willingness of villagers to invest in their communities, further disrupts local development efforts. Additionally, internal conflicts create an unstable environment that undermines both investment and collaboration among community members. To address these challenges, it is essential to implement targeted interventions that improve infrastructure and access to basic services, promote conflict resolution mechanisms, and encourage community participation in development initiatives. Therefore, by facilitating access to microloans and vocational training, rural residents can be empowered to invest in local businesses, thereby enhancing economic resilience and fostering a sense of ownership towards community development.

The second factor that plays a significant role in the underdevelopment of rural areas in Bamyan Province is the lack of basic facilities and access to essential services. The absence of communication infrastructure and health services clearly has a negative impact on the quality of life for residents; specifically, the lack of access to telecommunications and the internet limits social and economic interactions and deprives villagers of informational opportunities. This issue is particularly concerning in today's world, where information technology plays a crucial role in development. Moreover, the "low level of access to educational services" and "lack of access to agricultural and livestock loans" hinder the economic and social empowerment of rural residents. The absence of educational and health programs for villagers contributes to the perpetuation of poverty and social inequalities, leading future generations to face similar challenges. These barriers significantly reduce the motivation for investment and participation in development activities, creating a cycle of economic stagnation. To address these challenges, it is essential to design and implement comprehensive development programs that specifically focus on improving communication and educational infrastructure. Establishing educational and health centers, along with facilitating access to agricultural and livestock loans, can enable villagers to invest in their economic activities. These initiatives will not only lead to increased income and economic sustainability but also enhance the sense of ownership and active participation in community development.

The lack of infrastructure and resistance to innovation have been identified as the third factor contributing to the underdevelopment of rural areas, significantly impacting the quality of life for residents. The absence of access to a stable electricity network in villages is considered one of the greatest barriers to development. This deficiency hinders the effective utilization of modern technologies and agricultural equipment. Without electricity, the use of advanced irrigation systems, agricultural machinery, and even communication devices is severely limited, resulting in the persistence of traditional and inefficient farming methods. Consequently, this situation contributes to a decline in production and income for residents, thereby weakening the local economy.

Furthermore, local residents' opposition to innovative agricultural ideas obstructs the acceptance of innovation and the improvement of production methods. This resistance is often rooted in a fear of change, a lack of awareness regarding the benefits of new ideas, and an attachment to old traditions. Additionally, the failure to embrace innovation leads to decreased productivity and economic instability in rural communities, as villagers may continue to rely on traditional methods that are significantly less efficient and profitable than newer approaches. Therefore, it is essential for officials in the agricultural promotion and rural development sectors to make considerable efforts to enhance awareness and educate residents about the benefits of innovation.

Resistance to change and weak communication are recognized as significant factors contributing to the underdevelopment of rural areas. Rural populations often exhibit low acceptance of changes and new technologies, which leads to resistance against innovations and economic improvements. This issue is particularly exacerbated in rural environments that lack the necessary infrastructure for investment. Furthermore, the poor communication and collaboration between local communities and governmental organizations result in limited access to resources and support opportunities, hindering the empowerment of rural communities and their progress towards development. Consequently, these factors are interconnected, creating a cycle of underdevelopment in rural regions.

Inefficient management of natural resources in Bamyan Province has been identified as the fifth factor contributing to the underdevelopment of rural areas. Undoubtedly, inadequate planning and improper use of existing resources have led to environmental degradation and a decline in the quality of life for rural residents. To improve the situation, it is recommended that comprehensive natural resource management programs be developed and implemented in collaboration with local communities. Additionally, educating rural residents about sustainable resource utilization can help preserve these resources and enhance their livelihoods, making it a crucial requirement.

Migration from rural areas to cities, driven by better job opportunities and improved living conditions, has also been a significant factor in the underdevelopment of Bamyan. This trend not only results in a decrease in rural populations but also leads to the erosion of local skills and expertise. To address this issue, it is essential to create job opportunities in rural areas, including the development of small industries and sustainable employment. Furthermore, supportive programs aimed at retaining youth in rural areas and strengthening social and economic infrastructure can help mitigate migration.

Geographical and communication challenges in Bamyan Province, including remoteness from urban centers and a lack of transportation infrastructure, further exacerbate the underdevelopment of rural areas. These challenges lead to limited access to essential services, markets, and educational opportunities. To tackle this problem, investment in transportation and communication infrastructure is crucial. The significance of this matter has also been emphasized in Ahmadi's (2025) research.

The findings of this study align with those of previous research. For instance, in Ahmadikish et al.'s (2017) research, management factors had the greatest impact on the causes of underdevelopment in rural areas. In contrast, in the present study, economic and social problems emerged as the most significant factors. These differences may reflect the specific geographical and cultural conditions of various regions and highlight the need for further investigations in this area. To achieve sustainable development in the rural areas of the studied region, it is recommended to focus on diversifying economic activities, promoting entrepreneurship, addressing drought challenges, developing small businesses, and distributing microloans to increase employment opportunities and strengthen rural infrastructure as essential requirements. Similar studies (Hashimi & Shaiq, 2025; Talebpour et al., 2022b; Varmazyari et al., 2022) have strongly emphasized these recommendations, and authorities should undertake the necessary efforts in this regard.

By comparing the findings of this study with the results of previous research, it can be observed that there are both similarities and differences. In some previous studies, greater emphasis has been placed on policy, institutional, and environmental dimensions. In contrast, the findings of the present study indicate that certain local factors and the socio-economic conditions of the study area play a more prominent role in shaping rural development challenges. Therefore, although the overall framework of challenges identified in most studies appears to be similar, the intensity and priority of these challenges vary depending on the spatial, economic, and social conditions of each region. The results of the exploratory factor analysis in this study further revealed that rural development challenges are not merely a set of isolated variables, but can be explained through several underlying and interrelated factors. By uncovering the latent structure among the variables, this method not only reduced the complexity of the data but also enabled the identification of the main dimensions of the challenges and the determination of the relative importance of each factor.

Although this research may have its limitations, it holds significant value due to the scarcity of scientific resources and official reports on rural development in Afghanistan. Afghanistan is a country where over 70% of the population lives in rural areas, yet these regions continue to face numerous challenges. Nevertheless, the findings of this study can play a crucial role in raising awareness among stakeholders and the government, enabling them to take the necessary actions to address the identified barriers. Furthermore, the results of this research highlight critical aspects of the causes of underdevelopment in Bamyan Province, which could potentially contribute to the country's rural development efforts. This research can serve as a clear roadmap for government policymakers, providing valuable insights to overcome known obstacles and challenges in achieving sustainable rural development in Afghanistan.

5. Conclusion

This study demonstrated that underdevelopment in the villages of Bamyan is a multidimensional and interconnected phenomenon rooted in economic-social, infrastructural, institutional, and geographical factors. This reality indicates that the existing problems are not confined to a specific dimension, and addressing the issues of these areas requires more than one-dimensional development approaches. For instance, focusing solely on job creation without considering infrastructure or local institutional capacities may lead to the failure of achieving sustainable outcomes. To break the vicious cycle of underdevelopment in this region, an integrated and comprehensive program is essential. This program must simultaneously address several key dimensions and ensure synergy among them. Job creation should occur alongside strengthening infrastructure, such as improving transportation systems and access to essential services. Additionally, local institutional capacity-building, including training and empowering local individuals to manage resources and programs, should be an integral part of this initiative.

Furthermore, adapting to geographical challenges, such as remoteness from urban centers and specific climatic conditions, necessitates tailored strategies that may include the adoption of innovative technologies and sustainable solutions. Ultimately, future research could explore the dynamics of the relationships between these factors, providing deeper insights that can inform more effective development strategies.

5.1. Implications for Policy

- It is recommended that governments and international organizations invest in communication projects, such as satellite internet, and improve rural road infrastructure to reduce isolation.
- It is recommended that comprehensive educational programs focusing on the enhancement of technical, vocational, and entrepreneurial skills be designed and implemented by relevant institutions, particularly for vulnerable groups, including low-income populations.
- It is suggested that relevant authorities prioritize the provision of essential infrastructure, such as access to clean drinking water, healthcare and educational services, transportation networks, and sustainable energy.
- It is essential to develop and implement supportive policies, such as providing financial incentives, tax exemptions, and legal guarantees, to attract and encourage both domestic and foreign investors in sectors like smart agriculture, value-added industries, and sustainable transportation.
- To mitigate the excessive migration to urban areas, it is essential to formulate and implement long-term local employment policies. These policies should aim to create sustainable jobs, improve the quality of life, and enhance human development indicators in rural areas.

5.2. Recommendations for Future Research

- It is recommended that future studies utilize Structural Equation Modeling (SEM) techniques to more rigorously test the causal relationships among the identified variables. This approach can provide a deeper understanding of the interactions between various factors and contribute to strengthening both theoretical foundations and practical recommendations.
- It is suggested that qualitative research, particularly through in-depth interviews with local stakeholders, be conducted to identify the cultural, social, and psychological roots of resistance to innovation. Such studies can provide a more nuanced understanding of people's attitudes and experiences, thereby enabling the development of effective strategies to facilitate the acceptance of new technologies.
- It is recommended that a comparative study be conducted in other mountainous provinces of Afghanistan to identify regional differences and similarities in the factors influencing development. Such research could lead to more precise policy recommendations tailored to local contexts and provide a comprehensive understanding of the challenges and opportunities in various regions.
- Despite limitations such as a relatively small sample size, cross-sectional design, and reliance on self-reported data, the findings of this study are valuable and provide new insights into the topic. Future research is recommended to use larger samples, longitudinal designs, and more objective data to achieve more accurate and generalizable results.

CRediT Author Statement: Sayed Jalil Hashimi: Data curation, Writing – original draft, Visualization, and Investigation; **Muhammad Asef Shaiq:** Conceptualization, Methodology, Writing – review & editing, Software, and Validation; **Ali Akbar Barati:** Investigation, Visualization, and Supervision.

Data Availability Statement: Primary data will be made available upon official request from the first author.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflicts of interest.

IRB Statement: Not applicable.

Informed Consent Statement: Not applicable.

Acknowledgments: Not applicable.

Abbreviations

The following abbreviations are used in this manuscript:

EFA	Exploratory Factor Analysis
KMO	Kaiser–Meyer–Olkin
SEM	Structural Equation Modeling

References

Adamowicz, M., & Zwolińska-Ligaj, M. (2020). The “smart village” as a way to achieve sustainable development in rural areas of Poland. *Sustainability*, *12*(16), 6503. https://doi.org/10.3390/su12166503

Adimassu, Z., Kessler, A., Yirga, C., & Stroosnijder, L. (2013). Farmers’ perceptions of land degradation and their investments in land management: A case study in the Central Rift Valley of Ethiopia. *Environmental Management*, *51*, 989–998. https://doi.org/10.1007/s00267-013-0030-z

Ahmadikish, A.-A., Ahmadvand, M., & Sharifzadeh, M. (2017). Comparative analysis of the causes of rural underdevelopment: The case of villages in Margoon District of Boyer-Ahmad County (In Persian). *Iranian Agricultural Extension and Education Journal*, *12*(2), 123–135.

Ahmadi, R. (2025). Examination of social and economic factors affecting agricultural development in Bamyan Province, Afghanistan. *Journal of Sociological Context*, *6*(2), 213–247.

Ayoo, C. (2022). Poverty reduction strategies in developing countries. In P. de Salvo, & M. V. Pineiro (Eds.), *Rural development - Education, sustainability, multifunctionality* (pp. 17–57). IntechOpen.

Badri, S. A., Kazemi, N., Khodadadi, P., & Mohammadnejad, A. (2021). Why rural development policies have not contributed to rural development in Iran. *Rural Society*, *30*(2–3), 84–100. https://doi.org/10.1080/10371656.2021.1997423

Borodina, O., & Prokopa, I. V. (2019). Inclusive rural development: A scientific discourse. *Economy and Forecasting*, (1), 67–82. https://savearchive.zbw.eu/bitstream/11159/6910/1/1795549629_0.pdf

Dixon, C. (2015). *Rural development in the third world*. Routledge. https://doi.org/10.4324/9781315685755

Fayez, W. M., Najafi, S. M. B., Karimi, M. S. H., & Fathollahi, J. (2022). Identification and ranking of the factors affecting sustainable rural development in Afghanistan. *Economics and Regional Development Journal*, *29*(1), 212–240. https://www.researchgate.net/publication/385272140

Flint, R. W. (2013). *Practice of sustainable community development: A participatory framework for change*. Springer New York. http://dx.doi.org/10.1007/978-1-4614-5100-6

Ghazali, S., & Zibaei, M. (2017). Determinants of livelihood strategies selecting by nomadic households of Fars province (In Persian). *Agricultural Economics*, *11*(4), 63–80 https://doi.org/10.22034/iaes.2018.27529

Goretzko, D., Pham, T. T. H., & Bühner, M. (2021). Exploratory factor analysis: Current use, methodological developments and recommendations for good practice. *Current Psychology*, *40*, 3510–3521. https://doi.org/10.1007/s12144-019-00300-2

Hashimi, S. J., & Shaiq, M. A. (2025). Factor analysis of key drivers shaping agricultural entrepreneurship in Bamyan, Afghanistan. *DYSONA–Applied Science*, *6*(2), 389–397 https://doi.org/10.30493/das.2025.507024

Javadi, F. S., Rismanchi, B., Sarraf, M., Afshar, O., Saidur, R., Ping, H. W., & Rahim, N. A. (2013). Global policy of rural electrification. *Renewable and Sustainable Energy Reviews*, *19*, 402–416. https://doi.org/10.1016/j.rser.2012.11.053

Kalantari, K., Fami, H. S., Asadi, A., Qasemi, I., & Chubchian, S. (2008). Major challenges of Iranian rural communities for achieving sustainable development. *American Journal of Agricultural and Biological Sciences*, *3*(4), 724–728. https://doi.org/10.3844/ajabssp.2008.724.728

Lali, M., Zhang, S., & Arefi, R. (2020). The historically significance of Bamyan province and the role of historical aspects in today’s situations. *American Journal of Humanities and Social Sciences Research*, *4*(5), 40–48. https://www.ajhssr.com/wp-content/uploads/2020/05/F20454048.pdf

Long, H., Ma, L., Zhang, Y., & Qu, L. (2022). Multifunctional rural development in China: Pattern, process and mechanism. *Habitat International*, *121*, 102530. https://doi.org/10.1016/j.habitatint.2022.102530

Mohd Rokeman, N. R. (2024). Likert measurement scale in education and social sciences: Explored and explained. *EDUCATUM Journal of Social Sciences*, *10*(1), 77–88. https://doi.org/10.37134/ejoss.vol10.1.7.2024

Pain, A., & Hansen, K. (2019). *Rural development*. Routledge. https://doi.org/10.4324/9781315773643

Qi, K., Fan, Z., Ng, C. N., Wang, X., & Xie, Y. (2017). Functional analysis of landscape connectivity at the landscape, component, and patch levels: A case study of Minqing County, Fuzhou City, China. *Applied Geography*, *80*, 64–77. https://doi.org/10.1016/j.apgeog.2017.01.009

Savari, M., & Maymand, R. E. (2013). Barriers of sustainable rural development from perspective of experts. *International Journal of Advanced Biological and Biomedical Research*, *1*(8), 789–794. https://www.ijabbr.com/article_7806_4bad371804214c4bd5286a25d59650b4.pdf

Shaiq, M. A., Barati, A. A., Kalantari, K., & Asadi, A. (2022). Dimensions of poverty in Kunduz province of Afghanistan. *World*, *3*(4), 979–992. https://doi.org/10.3390/world3040055

Shaiq, M. A., Kalantaari, K., Asadi, A., & Barati, A. A. (2021). Investigating Afganistan's rural development challenges and their solutions (In Persian). *Journal of Rural Research*, *12*(3), 520–535. https://doi.org/10.22059/jrur.2021.314964.1585

Shaiq, M. A., Subat, S., Naiemi, M., & Fazli, G. R. (2026). Key barriers to livestock development in Baghlan Province: Insights from livestock farmers. *Journal of Natural Sciences – Kabul University*, *8*(4), 237–253. https://doi.org/10.62810/jns.v8i4.498

Sharifi, M. S., & Karim, A. Q. (2024). Agriculture in Afghanistan: A journey from livelihood to development. *Journal of Natural Science Review*, *2*(Special. Issue), 537–549. https://doi.org/10.62810/jnsr.v2iSpecial.Issue.156

Talebpour, F., Shabanali Fami, H., Barati, A., & Shaiq, M. A. (2022a). Prerequisites and mechanisms for the development of green businesses in Rey County from the perspective of rural women and experts (In Persian). *Journal of Economic Geography Research*, *3*(7). https://jurs.znu.ac.ir/article_252532_9feab5852e94c008a82547fa7d12236c.pdf

Talebpour, F., Shabanali Fami, H., Barati, A., & Shaiq, M. A. (2022b). Investigating the green status of rural women's businesses in Rey County. *Journal of Studies in Entrepreneurship and Sustainable Agricultural Development*, *9*(1), 1–16.

United Nations Digital Public Infrastructure. (2017). *The sustainable development goals*. United Nations.

Varmazyari, H., Shaiq, M. A., Amiri Mijan, M. A., & Kalantari, K. (2022). Spending business microloans on rural employment development; Application of a logistic regression model. *Journal of Rural Research*, *13*(3), 372–389. https://doi.org/10.22059/jrur.2022.328164.1660

Widaman, K. F., & Helm, J. L. (2023). Exploratory factor analysis and confirmatory factor analysis. In H. Cooper, M. N. Coutanche, L. M. McMullen, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), *APA handbook of research methods in psychology: Data analysis and research publication* (2nd ed., pp. 379–410). American Psychological Association. https://doi.org/10.1037/0000320-017

Yar, F. G. M., & Yasouri, M. (2024). Rural development challenges in addition to effective solutions to overcome obstacles. *Zhongguo Kuangye Daxue Xuebao*, *29*(3), 79–90. https://zkdx.ch/journal/zkdx/article/view/59

Yar, F. G. M., Yasouri, M., & Astuti, A. W. (2022). Identification and analysis of economic factors affecting rural underdevelopment in Afghanistan. *Jurnal Impresi Indonesia*, *1*(12), 1349–1354. https://doi.org/10.58344/jii.v1i12.4647

Zinchuk, T., Kutsmus, N., Kovalchuk, O., & Charucka, O. (2018). Challenges of sustainable development of rural economy. *Management Theory and Studies for Rural Business and Infrastructure Development*, *40*(4), 609–619. https://portalcris.vdu.lt/server/api/core/bitstreams/3f3117ca-ea2c-464c-992c-1d3c00c76200/content

Agricultural & Rural Studies

Article

Estimating the Recreational Value of Volcanic-Slope Agricultural Landscapes Using Avidity-Corrected Travel Cost Models

Evi Irawan [1,*], Syafika Rahma Bintari [1] and Chunyu Ying [2,*]

1 Faculty of Agriculture, Universitas Sebelas Maret, Surakarta 57126, Indonesia; syafikarahmab@gmail.com
2 Jiyang College, Zhejiang A&F University, Zhuji 311800, China
* Correspondence: eirawan@staff.uns.ac.id (E.I.); 20239024@zjyc.edu.cn (C.Y.)

Abstract: Volcanic-slope agricultural landscapes provide both productive and recreational functions, yet their economic value remains underexamined, particularly in developing countries where many such areas operate as informal destinations. This study estimates recreation demand and consumer surplus for a volcanic-slope agricultural landscape in Indonesia, focusing on Selo and Cepogo on the saddle between Mount Merapi and Mount Merbabu. Using an on-site intercept survey of 100 visitors conducted in June–July 2024, annual visit frequency is modeled as a function of generalized travel cost, including time. To address the sampling problems inherent in on-site surveys, namely zero truncation and endogenous stratification, four count-data models are estimated: zero-truncated Poisson, negative binomial models, and their avidity-corrected counterparts. Across specifications, visitation declines with higher travel cost and increases strongly with perceived landscape attractiveness, while substitute sites reduce repeat visits, indicating a competitive regional recreation market. Correcting for avidity and overdispersion improves model fit and yields more conservative welfare estimates. Bootstrapped consumer surplus remains positive and substantial across specifications, with a preferred benchmark of approximately 435 thousand IDR per trip and 1,517 thousand IDR per individual annually. Although based on a modest sample, the findings show that volcanic-slope agroecosystems can generate measurable recreation benefits and that correcting on-site sampling bias is important for credible welfare estimation and local landscape governance.

Keywords: travel cost method; endogenous stratification; zero truncation; agricultural landscapes; cultural ecosystem services; volcanic slopes; consumer surplus; informal tourism

Citation: Irawan, E., Bintari, S. R., & Ying, C. (2026). Estimating the Recreational Value of Volcanic-Slope Agricultural Landscapes Using Avidity-Corrected Travel Cost Models. *Agricultural & Rural Studies*, *4*(2), 17. https://doi.org/10.59978/ar04020010

Received: 4 March 2026
Revised: 1 April 2026
Accepted: 7 April 2026
Published: 29 May 2026

1. Introduction

Agricultural landscapes are increasingly recognized not only as sites of production, but also as sources of cultural ecosystem services (CES) such as recreation, aesthetic enjoyment, and place-based meaning. Yet their economic value remains underexamined in many developing-country settings, especially where landscapes attract visitors without being planned or managed as formal tourist destinations. This gap is particularly important for volcanic-slope agricultural landscapes. In such settings, the visible qualities that attract visitors are inseparable from ecological and managerial practices that sustain farming on steep, disturbance-prone terrain. Estimating the recreational value of these landscapes therefore matters not only for tourism appraisal but also for wider debates on landscape stewardship, rural livelihoods, and multifunctional land use.

Volcanic-slope agricultural landscapes emerge from an unusual combination of high agricultural potential and recurring disturbance, producing scenery that is both dynamic and highly legible to visitors (De Bauw et al., 2016; Ligot et al., 2023; Saputra et al., 2022). Steep elevation gradients generate abrupt microclimatic transitions over short distances, enabling diversified crop portfolios such as horticulture, agroforestry, and mixed smallholder systems, while simultaneously increasing vulnerability to slope instability, erosion, and episodic ashfall (De Bauw et al., 2016; Purwaningsih et al., 2020; Saputra et al., 2022; Sarmiento-Soler et al., 2022). The resulting fine-grained mosaics of fields, hedgerows, forest patches, agroforestry, and settlement nodes are not only productive but also visually distinctive and closely tied to water and soil regulation (Purwaningsih et al., 2020; Saputra et al., 2022). In many tropical volcanic uplands, these landscapes are underpinned by volcanic-ash soils, commonly Andisols, whose mineralogy composition supports high yields but also demands careful management of fertility, cover, and erosion risk (Auxtero et al., 2004; Parfitt, 2009; van Ranst et al., 2004). Evidence from Java further suggests that intensive horticulture can alter soil biological indicators and long-run soil quality, implying that the scenic and recreational qualities of these landscapes depend on continued stewardship rather than arising automatically from production alone (Griffin, 2019; Kashyap et al., 2023; Moeskops et al., 2010, 2012).

This stewardship dimension is important because agricultural landscapes are increasingly conceptualized as multifunctional social-ecological systems that generate nonmarket benefits alongside food and fiber. A growing body of research shows that agricultural landscapes can provide meaningful cultural ecosystem services, especially aesthetic enjoyment and recreation, and that these benefits are strongly influenced by landscape structure, accessibility, and visitor perception (Assandri et al., 2018; de las Heras et al., 2024; Mameno et al., 2024; Plieninger et al., 2013, 2015; van Berkel & Verburg, 2014; van Zanten et al., 2016; Yue et al., 2025). This literature has established that heterogeneity, crop-forest mosaics, terraces, and iconic vistas can generate real welfare benefits, but it also remains uneven (Assandri et al., 2018; Cetin et al., 2021; van Berkel & Verburg, 2014). Much of the evidence is drawn from formally recognized recreational settings or more general agricultural landscapes, while volcanic-slope farming landscapes in developing countries remain far less studied. In particular, the Indonesian case is important because volcanic uplands are simultaneously productive, environmentally fragile, and increasingly integrated into short-break tourism circuits. What remains insufficiently understood is the extent to which these "working landscapes" generate measurable recreational welfare and how that value should matter for management.

A related valuation literature supports monetizing the nonmarket benefits produced by managed agroecosystems and organizing them within a total economic value perspective relevant for land-use decisions. Recent studies have valued ecosystem services at the agricultural landscape scale and shown that managed agroecosystems can generate benefits that are not captured by commodity output alone (Fleischer & Tsur, 2000; Kaleji et al., 2025; Mameno et al., 2024; Shahimoridi et al., 2024; Song et al., 2020). For the present study, this literature is useful because it clarifies that recreational value should not be treated as incidental. At the same time, it leaves an important gap. Most existing studies do not focus on volcanic-slope agricultural landscapes where recreation is tightly bundled with regulating and supporting services such as soil conservation, runoff control, vegetation management, and terrace maintenance. In such contexts, tourism growth, agricultural intensification, and ecological stewardship are closely intertwined. Valuation is therefore useful not simply because it produces a monetary estimate, but because it helps reveal whether the recreational benefits of landscape maintenance are economically meaningful enough to inform real management choices.

This question becomes even more relevant because many volcanic-slope scenic sites function as destinations without centralized planning or unified management. Informality here does not mean the absence of organization altogether; rather, destination functions are distributed across multiple actors and micro-practices, including parking provision, food stalls, roadside photo spots, informal guiding, and small-scale accommodation. Research on informal tourism entrepreneurship shows that these activities often operate across blurred formal-informal boundaries, while unequal access to capital can shape who benefits and who is excluded from decision-making (Çakmak et al., 2018). In rural Indonesia, tourism-village studies likewise suggest that customary institutions and community strategies may substitute for, or complement, formal destination governance where rules are socially embedded, and resources are shared (Rosalina et al., 2023). These insights suggest that volcanic-slope agricultural landscapes should be treated not only as agricultural production spaces, but also as destination systems with measurable visitation flows, welfare effects, and governance challenges.

Two contemporary forces are accelerating the transition from "working landscape" to "visited landscape." First, social-media imagery can reveal and amplify aesthetic demand by transforming particular landscape configurations into widely shared visual icons, thereby strengthening the link between landscape structure and cultural value (Gosal & Ziv, 2020; Tieskens et al., 2018). Second, social-media-motivated travel can trigger rapid increases in visitation and behavioral pressure even where sites were not originally designed for tourism, creating practical management problems in open-access environments (Siegel et al., 2023). Under these conditions, the recreational benefits and pressures associated with visitation are often overlooked by assessments that focus narrowly on crop output or a limited set of biophysical indicators. Travel cost methods are well-suited to this setting because they infer value from observed trip behavior even when access is free or entrance fees are absent, weakly enforced, or institutionally fragmented (Blaine et al., 2015; Cetin et al., 2021; Mäntymaa et al., 2021; Menendez-Carbo et al., 2020; Qiu & Fan, 2016; Tyllianakis, 2024; Van Sandt & McFadden, 2022). Existing studies also suggest that recreation benefits in agricultural landscapes can be economically meaningful rather than marginal to rural land-use decisions (Fleischer & Tsur, 2000; Kaleji et al., 2025; Mameno et al., 2024; Song et al., 2020).

The significance of the present study can be clarified at three levels. First, the research question is important because volcanic-slope agricultural landscapes generate visible and experiential benefits that are likely to affect welfare, yet these benefits are rarely measured in economic terms. Second, the study is necessary because current evidence says relatively little about informal, working agricultural landscapes in Indonesia, where tourism, farming, and ecological risk intersect. Third, the study is feasible because recurrent visitation to such landscapes can be observed directly and modeled using travel cost methods adapted to the realities of on-site survey data. In this sense, the paper addresses not only a substantive gap in the literature but also a practical need for evidence that can support more balanced decisions about access, stewardship, and benefit sharing.

Accordingly, the objective of this paper is to estimate recreation demand and consumer surplus for a volcanic-slope agricultural landscape conceptualized as an informal tourist destination, and to draw out the management relevance of these nonmarket benefits for stewardship and destination governance. Because data in emerging destinations are typically collected through on-site intercept surveys at viewpoints, farm cafés, and roadside pull-offs, the observed sample includes visitors only and is more likely to intercept frequent visitors than infrequent ones. This creates two well-known econometric problems: zero truncation, because non-visitors are unobserved, and endogenous stratification, because frequent visitors are oversampled. If these features are ignored, the estimated travel-cost relationship may be biased, and welfare measures may be overstated (Amoako-Tuffour & Martínez-Espiñeira, 2012; Hindsley et al., 2011; Martínez-Espiñeira & Amoako-Tuffour, 2008; Shaw, 1988). The paper therefore applies travel-cost models that explicitly address these sampling issues. In doing so, it contributes in three ways: it extends recreation-demand valuation to a distinctive volcanic-slope agricultural landscape in Indonesia; it shows why correcting for on-site sampling bias matters for credible welfare estimation; and it links those welfare estimates to practical questions of landscape stewardship, benefit sharing, and governance in informal destination economies.

2. Methods

2.1. Study Area Rationale: Altitude, Risk, Seasonality

The study focuses on Selo and Cepogo sub-districts in Boyolali Regency, Central Java, Indonesia. These areas were selected because they lie on the steep saddle between Mount Merapi and Mount Merbabu, creating an unusually sharp altitude gradient over a short distance. This gradient produces a clearly differentiated sequence of farming landscapes, ranging from terraced high-elevation vegetable systems in Selo to more gently contoured areas in Cepogo, where forage production and smallholder dairy farming become more prominent. The site therefore offers strong ecological and visual contrasts within a single volcanic landscape system.

The area is also relevant because agriculture is practiced under conditions of recurrent environmental risk. Local livelihoods are shaped by fertile volcanic soils, steep slopes, and periodic disturbance associated with one of Indonesia's most active volcanic zones. In this sense, the landscape is not only a biophysical product of volcanic geomorphology, but also a social product maintained through adaptation, local knowledge, and long-standing farming practices (Brata et al., 2021; Lavigne et al., 2008). These features make the area particularly suitable for examining how agricultural stewardship and perceived landscape quality may support recreation demand.

A further reason for selecting the site is its clear seasonal dynamism. Cropping patterns shift across the annual cycle, with dry-season and wet-season production creating a changing but still legible working landscape. This seasonal rhythm reinforces the relevance of the site for a recreation-demand study because visitors encounter not static scenery, but a managed agricultural mosaic shaped by cropping decisions, environmental constraints, and conservation-minded practices (Irham et al., 2022; Moreira et al., 2024; Waha et al., 2020).

2.2. Survey Design

This study applies the individual travel cost method to estimate the recreational value of a volcanic-slope agricultural landscape in Selo and Cepogo. Data were collected through an on-site intercept survey, a common approach in travel cost studies because it directly targets actual users of the site (Landry et al., 2016; Martínez-Espiñeira & Amoako-Tuffour, 2008). This design is especially appropriate in informal destinations such as the present case, where no entrance gate, visitor register, or official attendance statistics exist. However, on-site sampling also has implications for inference. All individuals sampled are visitors, and the trip variable is recorded only for positive counts.

Face-to-face interviews were conducted using a structured questionnaire during June-July 2024. The questionnaire recorded annual trip frequency, travel cost components, time use, perceived landscape attractiveness, substitute-site information, and basic socio-demographic characteristics. A total of 100 respondents were interviewed. Because no formal sampling frame was available, respondents were selected using accidental or convenience sampling at relevant stopping points such as viewpoints, roadside areas, and agricultural visitor nodes.

The sample size is modest for travel cost estimation and should therefore be interpreted with appropriate caution. In particular, the estimates are best understood as site-specific evidence rather than as population-wide measures that can be aggregated mechanically beyond the study context. At the same time, the sample remains informative because it captures actual visitors in an emerging informal destination where visitation is real but not administratively recorded. For this reason, the study is suitable for estimating recreation demand and comparing alternative count-data specifications, while broader generalization should be reserved for future work using expanded and repeated samples.

2.3. Variables and Measurement

The dependent variable is the annual number of visits made by individual i, denoted v_i, measured as the reported number of trips to the site in the last 12 months. The key “price” variable is generalized travel cost per trip, denote τ_i. Following standard practice in travel cost studies, τ_i is constructed from out-of-pocket expenses and the opportunity cost of time (Shaw, 1988).

Perceived landscape quality is measured by a visitor rating of agricultural landscape attractiveness, denoted q_i. This variable captures the aesthetic and experiential dimension of the agricultural landscape, which is central in valuation studies of agricultural amenities. In addition, to reduce the risk that observed demand simply reflects the absence of alternatives, the analysis also controls for whether respondents could point to a reasonable substitute destination, capturing the availability of outside options that may limit repeat visitation.

A set of individual characteristics is included to control for heterogeneity in preferences and constraints. Let z_i denote a vector of controls, including age, gender, years of schooling, marital status, household income, and residency status (resident of Boyolali Regency vs non-resident). Because income is recorded in brackets, we use the mid-point of each bracket (Table 1) as an approximately continuous proxy for monthly income in the estimation.

Table 1. Variable definition and descriptive statistics (N = 100).

Variable		Description	Mean	Standard Deviation	Min	Max
1.	VISITS	Number of visits per person in one year	3.470	2.634	1	15
2.	TC	Round-trip travel cost per visit per person, including transport expenses and other trip-related costs including the opportunity cost of travel/visit time. Measured in IDR 1,000.	383.098	553.491	38	3,575
3.	AGL	Agricultural landscape attractiveness, based on visitor rating (1–5, higher = more attractive).	3.740	0.664	3	5
4.	SUB	Substitute site (=1 if a substitute recreational site was identified by the respondent)	0.460	0.501	0	1
5.	AGE	Age (years).	37.460	11.390	17	64
6.	FEM	Gender dummy (1 = female, 0 = male).	0.650	0.479	0	1
7.	EDU	Years of schooling (years).	13.110	2.892	6	18
8.	MAR	Marital status dummy (1 = married, 0 = not married).	0.840	0.368	0	1
9.	INC	Mid-point of four average monthly income brackets in IDR 1 million	2.545	1.339	1	5
10.	RSD	Local residence dummy (1 = resides in Boyolali Regency; 0 = non-local).	0.390	0.490	0	1

2.4. Conceptual Foundation

Recreation demand is derived from a utility-maximization problem in which an individual allocates resources between general consumption and trips to the site. Following Fleischer and Tsur

(2000), the agricultural landscape travel cost setting used in earlier work, landscape quality (q_i) and a substitute site shifts utility from visitation, implying systematic changes in trip demand even after accounting for generalized travel cost. The choice problem is

$$\max_{v_i \geq 0,\, c_i \geq 0} U_i(c_i, v_i; q_i) \quad s.t. \quad c_i + \tau_i v_i \leq m_i, \tag{1}$$

where c_i denote the consumption v_i denote the number of visits made by a person i in one year of all other goods. m_i represents income and τ_i is generalized travel cost per trip. This framework implies that expected visits should decline as generalized travel cost rises and increase when perceived landscape quality improves, holding other factors constant.

2.5 Construction of Generalized Travel Cost

Generalized travel cost combines monetary expenses and time costs. Time cost is valued using the reservation wage or usual hourly earnings:

$$time\ cost_i = w_i \times (t_i^{travel} + t_i^{onsite}) \tag{2}$$

where w_i is the reservation wage (usual hourly earnings), t_i^{travel} is travel time, and t_i^{onsite} is the time spent at the destination. The generalized cost (τ_i) then takes form:

$$\tau_i = money\ cost_i + time\ cost_i \tag{3}$$

This construction is consistent with the travel cost literature in which time is treated as a real resource cost of recreation, and it is particularly relevant in mountainous volcanic landscapes where travel and on-site durations can be substantial. In the baseline specification, we value travel and on-site time at the individual's stated (or implied) hourly earnings; as a robustness check, welfare is re-computed under alternative time-valuation fractions to confirm that qualitative conclusions are unchanged.

2.6 Econometrics: Truncation and Endogenous Stratification

Because the survey is conducted on-site, the observed sample is drawn from visitors only, not from the full population of potential visitors. If these features are ignored, estimated demand may appear less price-sensitive than it really is, and consumer surplus may be overstated. As a result, the observed trip counts are strictly positive $(v_i \geq 1)$, so standard count-data models that assign positive probability to zero trips are not fully consistent with the sampling process. This issue is commonly treated as zero truncation in single-site recreation demand estimation (Shaw, 1988). In addition, on-site intercept surveys typically do not sample visitors uniformly. Individuals who take more trips are more likely to be intercepted, which creates endogenous stratification (avidity bias) unless the likelihood is adjusted appropriately (Englin & Shonkwiler, 1995). Empirical applications emphasize that ignoring these features can materially affect the estimated travel-cost coefficient and welfare measures (Curtis, 2002; Hynes et al., 2017).

The recreation demand function is specified through the conditional mean of annual visits. Let τ_i denote generalized travel cost per trip, q_i denote perceived landscape aesthetics, and z_i denote a vector of observed visitor attributes. The conditional mean is modeled using an exponential link:

$$\mu_i = E[v_i | \tau_i, q_i, z_i] = exp\left(\alpha + \beta_\tau \tau_i + \beta_q q_i + \sum_{k=1}^{K} \delta_k z_{ik}\right),\ \beta_\tau < 0 \tag{4}$$

The parameter β_τ is central for valuation because it captures how expected visits respond to changes in generalized travel cost. Since the survey is conducted on-site, observed trip counts satisfy $v_i \geq 1$. Accordingly, μ_i is interpreted as the latent (population) mean implied by covariates under the count process, while the mean trips in the observed on-site sample are obtained by conditioning on positive counts (e.g., under Poisson $\mathrm{E}[v_i | v_i \geq 1, x_i] = \mu_i / 1 - \exp(-\mu_i)$). This distinction matters for interpretation, but the travel-cost semi-elasticity driving welfare remains governed by β_τ.

Under a Poisson baseline, the probability of observing $v_i = v$ conditional on $v_i \geq 1$ is the zero-truncated Poisson probability mass function:

$$\Pr(v_i = v \mid v_i \geq 1, \mu_i) = \frac{e^{-\mu_i}\mu_i^v}{v!(1 - e^{-\mu_i})}, \quad v = 1, 2, 3, \ldots \tag{5}$$

Parameters in (4)–(5) are estimated by maximum likelihood. However, recreation trip data often exhibit overdispersion (variance exceeding the mean), so a zero-truncated Negative Binomial (NB) specification is also estimated. In the NB model, the conditional mean remains μ_i as in (4), but the variance is allowed to exceed the mean according to:

$$\mathrm{Var}(v_i|\cdot) = \mu_i(1 + \kappa\mu_i), \kappa > 0 \tag{6}$$

Truncation is handled by conditioning the NB distribution on positive counts:

$$\Pr(v_i = v \mid v_i \geq 1) = \frac{\Pr_{NB}(v_i = v|\mu_i, \kappa)}{1 - \Pr_{NB}(v_i = 0|\mu_i, \kappa)}, v = 1, 2, 3, \ldots \tag{7}$$

where $\Pr_{NB}$ denotes the standard Negative Binomial probability mass function. In practice, model selection between the truncated Poisson and truncated NB specifications is guided by evidence of overdispersion and by information criteria, since mis-specifying dispersion can influence both the travel-cost sensitivity and the implied welfare measures (Englin & Shonkwiler, 1995).

The on-site sampling design also implies endogenous stratification. The probability of observing an individual in the sample is proportional to their number of trips. Following the on-site recreation-demand literature (Blaine et al., 2015; Englin & Shonkwiler, 1995; Shaw, 1988), this is represented by a size-biased (trip-weighted) likelihood. Formally, if the population trip-count probability is $f(v_i|\mu_i, \kappa)$, then the on-site sampling mechanism implies:

$$\Pr_{\text{on-site}}(v_i = v) = \frac{v f(v|\mu_i, \kappa)}{\mu_i}, v = 1, 2, 3, \ldots \tag{8}$$

Because the on-site likelihood is trip-weighted, individuals with $v_i = 0$ have zero sampling probability, so truncation is inherent, and the remaining adjustment corrects the over-representation of avid visitors among those with positive trips. This adjustment is important because it recovers parameters that are consistent with the underlying population demand rather than the trip-weighted on-site sample. The practical relevance of this correction is well illustrated in applied recreation settings, including angling demand, where truncated NB models with endogenous stratification are used for welfare estimation (Curtis, 2002). More recent work also shows how the avidity correction can be extended to more flexible dispersion structures such as generalized negative binomial formulations (Landry et al., 2016). Equation (8) implies that estimation is based on a trip-weighted likelihood, so the recovered parameters reflect underlying population demand rather than the mechanically trip-heavy composition of the on-site sample. With this corrected demand in hand, the next subsection translates the estimated travel-cost sensitivity into welfare measures by computing consumer surplus from the implied trip demand curve.

2.7. Welfare Measurement: Consumer Surplus

The recreational value of the site is derived from the estimated trip demand by measuring the surplus that visitors obtain beyond what they pay in generalized travel cost. The valuation step converts the estimated recreation demand into welfare measures. The economic value of access is captured by consumer surplus, defined as the area under the trip demand curve with respect to travel cost. In general form, individual consumer surplus is:

$$CS_i = \int_{\tau_i}^{\infty} \mathrm{E}[v_i|\tau_i, q_i, z_i] \; d\tau \tag{9}$$

Under the exponential mean equation (4) and $\beta_{tau} < 0$, this integral yields a closed form:

$$CS_i = \frac{\mu_i}{-\beta_\tau} \tag{10}$$

A convenient summary is consumer surplus per trip:

$$CS^{trip} = \frac{1}{-\beta_\tau} \tag{11}$$

Finally, the framework supports the valuation of changes in landscape quality. Consider an improvement from q_i^0 to q_i^1, while holding travel cost and attributes fixed. The implied welfare gain is:

$$\Delta CS_i = CS_i\,(q_i^1) - CS_i(q_i^0) = \frac{\mu_i(q_i^1) - \mu_i(q_i^0)}{-\beta_\tau} \qquad (12)$$

These expressions make the conceptual link explicit. Improved agricultural landscape quality increases expected visits, and the magnitude of the welfare gain is scaled by the sensitivity of visitation to travel cost.

3. Results

3.1. Sample Characteristics and Trip Motivations

Table 1 shows that visitation to the volcanic-slope agricultural landscape is both meaningful and unevenly distributed across visitors. The site does not appear to attract only one-time curiosity trips; rather, it supports repeated use, with respondents reporting an average of 3.47 visits per year (SD = 2.63). At the same time, the considerable variation in annual trips suggests that the landscape differs in its appeal and accessibility across individuals. A similar heterogeneity emerges in generalized travel cost. While some respondents reach the site with relatively modest time and monetary burdens, others incur substantially higher costs, as reflected in the mean generalized travel cost of IDR 383.10 thousand per trip and a large standard deviation of IDR 553.49 thousand. This wide dispersion is substantively important because it indicates that the site draws visitors from a geographically and economically diverse catchment area, while also providing the variation needed to identify the travel-cost demand relationship in a credible way (Czajkowski et al., 2019; Tyllianakis, 2024).

Visitors also tend to evaluate the landscape positively. The mean attractiveness score is 3.74 on a 1–5 scale (SD = 0.66), indicating generally favorable perceptions of the agricultural scenery, yet not such uniformity that quality differences become empirically irrelevant. This pattern is important for the logic of the analysis. It suggests that the site is not visited merely because it exists or because it is conveniently located; rather, visitors appear to respond to differences in how attractive they perceive the landscape to be. In this sense, perceived quality is likely to be more than a descriptive attribute. It is plausibly one of the mechanisms through which the landscape generates repeat visitation.

The socio-demographic profile further suggests that this is a socially broad but regionally outward-looking recreation setting. Respondents are predominantly of working age, with an average age of 37.45 years, and the sample has a distinctly family-oriented character, as reflected in the relatively high proportions of women (65%) and married individuals (84%). Educational attainment averages 13.11 years of schooling, while income is concentrated mainly in the lower-to-middle categories: 35% of respondents report a monthly income below IDR 2 million, 29% fall in the IDR 2–3 million range, and 21% belong to the highest category of above IDR 4 million. These patterns indicate that the site is not an exclusive leisure destination for affluent households, but one that is accessible to a fairly broad domestic visitor base. At the same time, only 39% of respondents reside in Boyolali Regency, implying that most visitors come from outside the immediate locality. Taken together, these descriptive results suggest that the landscape functions as a wider regional recreation destination, attracting visitors with different travel burdens, different socio-economic backgrounds, and different propensities to return. This pattern aligns with the wide travel-cost spread typically associated with heterogeneous demand across visitors (Tyllianakis, 2024; van Berkel & Verburg, 2014).

Thematic evidence from Tables 2 and 3 provides insight into the underlying drivers of visitation patterns and perceived quality. Reported reasons for visiting indicate that the destination offers more than rural scenery; it is also a site for active engagement with farming landscapes. Agritourism and farming observation are the primary motivations, supplemented by leisure trips, photography or content creation, and the pursuit of calm and reflection. This combination aligns with the concept that agricultural landscapes generate cultural ecosystem services-particularly recreation, aesthetic enjoyment, and place-based meanings-that attract visitors from beyond the local area (van Berkel & Verburg, 2014). The experience outcomes reinforce that interpretation.

Table 2. Reasons for visiting agricultural landscapes (N = 83).

Reason Theme	Percentage
1. Agrotourism and farming observation	20%
2. Leisure trip	12%
3. Photography and content creation	12%
4. Relaxation, healing, and reflection	12%
5. Shopping for fresh produce and culinary	10%
6. Education, research, and fieldwork	7%
7. Cool climate and clean air escape	5%
8. Outdoor exercise and exploration	5%
9. Slow travel, stopover, and avoiding crowds	5%
10. Nostalgia and village/social life	5%
11. Creative inspiration	5%
12. Helping/volunteering/assisting	2%

Table 3. Experiences obtained from visiting agricultural landscapes (N = 83).

Experience Theme	Percentage
1. Learning and practical insight	52%
2. Scenic beauty and photo-worthy moments	22%
3. Calm, relaxation, and mental refresh	11%
4. Warm togetherness and social connection	7%
5. Physical refresh and clean air	5%
6. Fresh produce and culinary enjoyment	1%
7. Creativity and inspiration	1%
8. Appreciation, pride, and gratitude	1%

The dominant reported experience is learning and practical insight, suggesting that visitors often leave with a concrete understanding of crops, farm routines, and rural livelihoods rather than only passive enjoyment. Educational and experience-based benefits are frequently emphasized in agritourism research as a central pathway through which farm visits create value and shape satisfaction and future intentions (Chen et al., 2020; Tang et al., 2022). At the same time, scenic beauty and photo-worthy moments appear prominently as outcomes, matching the relatively high attractiveness scores in Table 1 and indicating that perceived landscape quality is likely tied to visual appreciation and the "shareability" of the setting. The presence of togetherness/social connection among reported experiences also fits the sample profile (high shares of married and female respondents) and supports the view that the Merbabu-slope agricultural landscape functions as a hybrid recreation space—simultaneously a living landscape for learning, a scenic environment for aesthetic enjoyment, and a setting for family and social time (Tang et al., 2022; van Berkel & Verburg, 2014).

The quantitative data are enriched by qualitative responses. For instance, one visitor noted, "Walking through these terraced fields, I learned how farmers prevent erosion on steep slopes—it's not just scenery but a lesson in sustainability." Such quotes transform abstract concepts like "cultural ecosystem services" into tangible experiences, underscoring why perceived attractiveness drives demand. Another respondent shared, "The peace here helps me reconnect with nature," highlighting the emotional value beyond recreation.

3.2. Recreation Demand Models

Table 4 reports four count-data specifications that reflect common features of on-site recreation surveys: a zero-truncated Poisson (TP), a zero-truncated negative binomial (TNB), a truncated endogenously stratified Poisson (TSP), and a truncated endogenously stratified negative binomial (TSNB). Across all models, the travel-cost coefficient is negative and precisely estimated (β_{TC} =

−0.0016 to −0.0023; p-value < 0.01), confirming a downward-sloping demand for visits. As generalized trip cost rises, visitation falls. This is the core revealed-preference result on which the travel cost method depends, and it indicates that even in an open-access setting with no formal entrance fee, visitors still face meaningful economic constraints through travel expenses and time costs.

Table 4. Estimation results.

Variable	**(1) Zero-truncated Poisson (TP)**	**(2) Zero-truncated Negative Binomial (TNB)**	**(3) Truncated and endogenously stratified Poisson (TSP)**	**(4) Truncated and endogenously stratified negative binomial (TSNB)**
TCOST	−0.0016***	−0.0019***	−0.0020***	−0.0023***
	(0.0004)	(0.0004)	(0.0005)	(0.0005)
AGL	0.5254***	0.5407***	0.6336***	0.6394***
	(0.1007)	(0.1047)	(0.1190)	(0.1208)
SUB	−0.3615***	−0.3936***	−0.4342***	−0.4783***
	(0.1260)	(0.1353)	(0.1507)	(0.1575)
AGE	−0.0062	−0.0057	−0.0074	−0.0064
	(0.0085)	(0.0085)	(0.0101)	(0.0098)
FEM	−0.1659	−0.1595	−0.2005	−0.1819
	(0.1535)	(0.1562)	(0.1831)	(0.1829)
EDU	−0.0220	−0.0165	−0.0262	−0.0153
	(0.0247)	(0.0253)	(0.0298)	(0.0293)
MAR	0.2338	0.2351	0.2804	0.2712
	(0.2224)	(0.2246)	(0.2632)	(0.2601)
INC	0.1447**	0.1457**	0.1734**	0.1690**
	(0.0594)	(0.0609)	(0.0710)	(0.0710)
RSD	−0.1307	−0.0900	−0.1546	−0.0779
	(0.1830)	(0.1735)	(0.2176)	(0.1967)
Constant	−0.1899	−0.3167	−0.7929	−1.0977*
	(0.4696)	(0.5352)	(0.5644)	(0.6559)
ln(α)				
Constant		−2.4607***		−1.5982***
		(0.5927)		(0.4697)
Statistics				
Log-likelihood	−175	−173	−176	−171
N	100	100	100	100
pseudo-R^2	0.2195	0.1612	0.2526	0.1703
χ^2	86.33	84.29	91.32	90.18
AIC	3.704	3.687	3.716	3.643

A notable result is that the travel-cost coefficient becomes more negative once endogenous stratification is addressed in the TSP and TSNB models. This implies that demand is more price-sensitive after correcting for the over-representation of frequent visitors in the on-site sample. Substantively, this is important because it confirms that avidity is not merely a technical concern. If frequent visitors are more likely to be intercepted and this is ignored, the estimated demand curve may appear artificially flatter than it really is, which in turn tends to inflate welfare estimates. The pattern in Table 4 therefore supports the methodological choice to go beyond simple truncation-only models and is consistent with both theoretical expectations and prior evidence in recreation economics.

Perceived agricultural landscape attractiveness (AGL) is positive and highly robust (β_{AGL} = 0.525–0.620; p-value < 0.01). Interpreted as incidence-rate ratios (IRR), a one-point increase on the 1–5 attractiveness scale is associated with roughly 69–86% more visits (IRR ≈ 1.69–1.86), holding other factors constant. This indicates that repeat visitation is strongly quality-driven. The result is also consistent with the earlier descriptive evidence, where visitors emphasized learning, scenic beauty, calmness, and photo-worthy moments as important parts of the experience. In this setting, terraces, crop mosaics, rural scenery, and the overall experiential quality of the landscape are not merely background conditions; they are central components of what visitors value and consume.

The substitute-site indicator (SUB) is negative and significant in every specification (β_{SUB} = −0.359 to −0.476; p-value < 0.01), implying about 30–38% fewer visits when respondents report an alternative site (IRR ≈ 0.70–0.62). This result indicates that the destination operates within a broader regional recreation market in which visitors compare available options rather than behaving as if they face a unique site with no alternatives. This strengthens the credibility of the demand estimates because it shows that visitation responds not only to travel cost and perceived quality, but also to the existence of competing outdoor experiences. In practical terms, it suggests that maintaining comparative advantage through stewardship and basic visitor-support conditions may be as important as expanding infrastructure.

Income (MID) is positive and statistically significant throughout (β_{MID} = 0.145–0.180; p < 0.05), suggesting visits behave as a normal good in this setting. In contrast, age, gender, education, marital status, and residence show small coefficients and remain statistically weak across specifications, indicating that (conditional on generalized travel cost and perceived landscape quality) visit frequency is not sharply segmented by these demographics in this sample. This does not imply that such characteristics are unimportant in every context, but within this sample their effects appear limited once generalized travel cost, perceived landscape attractiveness, and substitute availability are taken into account. The strongest and most consistent drivers of visitation are therefore price, quality, and alternatives.

Model diagnostics favor accounting for both overdispersion and onsite-sampling selection. Overdispersion is supported by the negative binomial results: ln(α) is significant in both NB-based models, rejecting the equidispersion restriction embedded in Poisson. Fit statistics also improve once avidity is addressed: TSNB has the highest log-likelihood (−171) and the lowest AIC (3.643), making it the preferred specification among the four. This provides a clear basis for selecting a benchmark welfare estimate rather than treating all models as equally persuasive. The econometric message is therefore straightforward: once trip overdispersion and the over-sampling of avid visitors are both addressed, the model fits the data better and yields a more credible basis for welfare interpretation. This pattern is consistent with prior recreation-demand evidence that combining truncation and endogenous stratification corrections can materially change inference and model fit (Hindsley et al., 2011; Landry et al., 2016; Martínez-Espiñeira & Amoako-Tuffour, 2008).

Across all models, the most consistent finding is that visitors are highly responsive to both generalized price and perceived landscape quality. The negative travel-cost effect is consistent with the revealed-preference basis of travel-cost demand: as trip costs, including time, increase, visitation declines. The increased negativity of the coefficient after correcting for endogenous stratification (TSP/TSNB) is particularly significant. On-site intercept sampling inherently increases the likelihood of sampling frequent visitors; if this avidity is not addressed, the estimated demand curve may be artificially flattened and welfare measures overstated. The observed changes in Table 4 are consistent with both theoretical expectations and empirical evidence in recreation economics.

3.3. Welfare Estimates and Robustness

Table 5 reports welfare implied by the estimated travel-cost coefficient using bootstrapped 95% confidence intervals (units in 1,000 IDR). The reported expected annual trips summarize mean trip frequency for the observed on-site sample and are nearly identical across specifications, so differences in welfare are driven primarily by changes in the estimated travel-cost sensitivity. Consumer surplus per trip (CS/trip) declines monotonically as the travel-cost coefficient becomes more negative: the point estimate falls from 625 (TP) to 526 (TNB), 500 (TSP), and 435 (TSNB). This pattern is substantively meaningful because it shows that accounting for endogenous stratification/avidity reduces inferred per-visit welfare-consistent with the idea that uncorrected on-site samples can flatten the demand curve and inflate surplus.

Table 5. Estimated 95% of confidence intervals of consumer surpluses (through bootstrapping).

		TP	TNB	TSP	TSNB
$\hat{\beta}_{TC}$		−0.0016	−0.0019	−0.0020	−0.0023
Expected visits (ŷ)		3.47	3.48	3.47	3.49
	Lower	351	346	293	284
CS/trip[a]	Point	625	526	500	435
	Upper	1,075	904	914	727
	Lower	1,217	1,203	1,018	990
CS/individual[a]	Point	2,169	1,832	1,735	1,517
	Upper	3,731	3,148	3,173	2,538

[a] Numbers are rounded and in 1000 IDR.

Annual consumer surplus per individual (CS/individual) follows the same ordering because $\hat{y}$ is nearly constant: 2,169 (TP), 1,832 (TNB), 1,735 (TSP), and 1,517 (TSNB). Moving from the simplest truncation-only Poisson model to the preferred TSNB specification reduces both per-trip and annual surplus by roughly 30%. This is a sizeable adjustment and reinforces the argument that avidity correction is not a minor technical refinement. If the over-representation of frequent visitors is ignored, welfare can be overstated in ways that are economically meaningful. The TSNB estimate should therefore be treated as the main benchmark, while the remaining models are best interpreted as sensitivity checks that reveal how strongly welfare depends on assumptions about dispersion and sampling.

At the same time, the welfare conclusion itself remains robust across models. Even under the most conservative specification, consumer surplus is clearly positive, with the lower 95% confidence limit still reaching 284 thousand IDR per trip and 990 thousand IDR per individual annually. Thus, although the exact magnitude varies by model, the broader inference does not: the volcanic-slope agricultural landscape generates substantial recreational value for users. The persistence of positive welfare across all specifications and confidence bounds strengthens the interpretation that this landscape is not merely incidentally visited, but provides meaningful nonmarket benefits through recreation and aesthetic experience.

These estimates become easier to interpret when placed in the context of the sample itself. Under the preferred TSNB model, the estimated consumer surplus per trip is approximately 435 thousand IDR, while the mean generalized travel cost reported in Table 1 is about 383 thousand IDR per trip. This suggests that the net recreational benefit of a visit is slightly larger than the average generalized burden incurred to make the trip. Put differently, visitors appear willing to bear substantial costs in money and time because the value they derive from the experience exceeds those costs by a meaningful margin. Likewise, the preferred annual consumer surplus per individual, 1,517 thousand IDR, is non-trivial when considered alongside the sample's monthly income distribution, which is concentrated in the lower-to-middle categories. Although these comparisons should not be interpreted as direct accounting equivalences, they help show that the welfare estimates are economically meaningful at the scale of ordinary domestic leisure decisions rather than being merely symbolic values. This directly responds to the need to relate the results to local realities.

4. Discussion

Several implications follow from the estimated demand and welfare results for understanding informal, agriculture-based outdoor recreation in volcanic uplands and for interpreting welfare measures derived from on-site travel-cost data. Across all four specifications, the travel-cost coefficient is negative and statistically robust, providing clear revealed-preference evidence that visitors respond to generalized trip price even in a non-ticketed setting where destination functions are loosely coordinated. This matters for policy because it suggests that "free access" does not imply inelastic demand; rather, time and travel burdens act as de facto prices, shaping who visits and how often. This behavioral regularity matches the logic of travel-cost valuation in settings where entry fees are absent or unenforced (Blaine et al., 2015; Mäntymaa et al., 2021; Menendez-Carbo et al., 2020; Van Sandt & McFadden, 2022). Importantly, the travel-cost coefficient becomes more negative once endogenous stratification is addressed (TSP/TSNB), indicating higher price sensitivity after correcting for the over-representation of frequent visitors in on-site samples. This pattern supports the central econometric argument of the paper: avidity is not merely a technical issue, but a mechanism that can materially distort welfare inference. If frequent visitors are sampled more often and this is ignored, the estimated demand curve may appear flatter than it truly is, leading to inflated consumer surplus estimates. This is consistent with classic results showing that on-site interception can mechanically flatten the demand curve and overstate welfare if avidity is ignored (Amoako-

Tuffour & Martínez-Espiñeira, 2012; Englin & Shonkwiler, 1995; Hindsley et al., 2011; Landry et al., 2016; Martínez-Espiñeira & Amoako-Tuffour, 2008; Shaw, 1988). For applied outdoor recreation research, the implication is straightforward. Truncated Poisson results remain useful for comparison, but they should not be treated as default welfare benchmarks when sampling occurs at the destination.

The strong and consistent effect of perceived agricultural landscape attractiveness suggests that repeat visitation is fundamentally quality-driven. This aligns closely with the descriptive findings, where visitors emphasized learning, scenic beauty, calmness, and photo-worthy moments as central parts of the experience. In this setting, attractiveness should not be read as a superficial visual rating. It likely captures a broader experiential judgment that includes the visible order of terraces, crop mosaics, comfort at stopping points, the sense of rural atmosphere, and the overall coherence of the working landscape. This interpretation is consistent with the cultural ecosystem services literature, which links recreation value to visitor-perceived quality and to the structural attributes of agricultural scenery (Assandri et al., 2018; Plieninger et al., 2013, 2015; Tieskens et al., 2018; van Berkel & Verburg, 2014). It is also compatible with evidence that social-media dynamics can amplify aesthetic demand by turning particular viewpoints and landscape configurations into widely shared visual icons, thereby accelerating visitation and intensifying pressures even where sites were not originally planned as destinations (Gosal & Ziv, 2020; Nyelele et al., 2023; Siegel et al., 2023; Tieskens et al., 2018). In this sense, visitors are not simply consuming "nature" in the abstract; they are responding to a managed agricultural environment whose recreational appeal depends on how land is cultivated, maintained, and experienced.

The consistently negative substitute-site effect further indicates that this destination operates within a broader regional portfolio of informal outdoor experiences. Visitors reduce repeat trips when substitutes are salient, implying that welfare should be interpreted in a choice context rather than as demand for a monopolistic site. This result aligns with evidence that CES benefits are spatially clustered around accessible nodes and viewpoints rather than evenly distributed across farmland, and that recreation behavior reflects a menu of alternatives shaped by travel constraints and perceived quality (Cetin et al., 2021; van Berkel & Verburg, 2014). For governance, the substitution pattern underscores that maintaining comparative advantage through stewardship and basic visitor-support functions (e.g., viewpoint upkeep, cleanliness, safety) can be as important as expanding facilities, especially where visitation is organized around scenic pull-offs and short walks. Since visitors' sort among comparable sites, the magnitude of welfare is inherently model-dependent and must be interpreted alongside the sampling and distributional assumptions embedded in the econometric specification.

In this context, the welfare estimates show that methodological choices translate into policy-relevant differences in benefit magnitudes. Moving from TP to TSNB reduces consumer surplus per trip from 625 to 435 thousand IDR and consumer surplus per individual from 2,169 to 1,517 thousand IDR, while TSNB also provides the strongest fit (highest log-likelihood and lowest AIC). Given the on-site intercept design, a conservative and well-justified reporting approach is to emphasize TSNB as the primary welfare benchmark because it simultaneously addresses zero truncation, endogenous stratification (avidity), and overdispersion. The remaining specifications (TP, TNB, and TSP) are best interpreted as sensitivity checks that transparently show how welfare magnitudes vary when key assumptions are relaxed or imposed. Although the bootstrap intervals are relatively wide—as expected under nonlinear transformations and a modest sample—welfare remains positive and substantively meaningful across models and confidence limits.

Overall, the welfare evidence supports two closely related findings. First, volcanic-slope agricultural landscapes generate measurable and non-trivial recreational benefits even when tourism remains informal, and access is not structured through an official fee system. Second, the estimated size of those benefits depends materially on whether the analyst correctly models the sampling process inherent in on-site data collection. For that reason, the preferred TSNB estimate provides the most credible basis for subsequent discussion of stewardship, benefit-sharing, and local governance. Rather than treating welfare valuation as an abstract numerical exercise, Table 5 shows that the recreational benefits associated with maintaining terraces, ground cover, and scenic agricultural mosaics are large enough to matter for practical decisions about how such landscapes are managed and by whom.

These welfare magnitudes are not merely accounting results; they help clarify what is at stake when land-use practices and local management affect the visitor experience. In volcanic-slope farming landscapes, steep gradients, recurrent disturbance, and management-dependent mosaics shape both regulating functions and the visible features that visitors consume (De Bauw et al., 2016; Purwaningsih et al., 2020; Saputra et al., 2022; Suprayogo et al., 2020). The strong quality effect therefore supports a co-benefits interpretation: stewardship practices that maintain terraces, ground cover, and landscape coherence likely reinforce both erosion regulation and recreation quality—an interdependence that becomes more consequential as horticultural intensity rises and soil-quality tradeoffs accumulate (Kashyap et al., 2023; Moeskops et al., 2010, 2012). This linkage also helps reconcile production-focused and recreation-focused perspectives: the same management choices that support agricultural performance and landscape resilience can shape the scenic heterogeneity and legibility that visitors value.

The findings also reinforce the analytical importance of informality for outdoor recreation settings embedded in productive rural landscapes. When destination functions are distributed across multiple actors—such as parking provision, small food stalls, informal guiding, and "photo-spot" upkeep—service quality and maintenance incentives may be uneven, and the distribution of benefits can reflect unequal access to capital and bargaining power (Çakmak et al., 2018). In rural Indonesia, community strategies and customary institutions may partially substitute for formal destination management in coordinating shared resources and socially embedded rules (Rosalina et al., 2023). Against this institutional background, welfare estimates provide more than a valuation metric. They offer an empirical basis for discussing how stewardship costs and tourism-derived benefits could be shared among farmers, micro-entrepreneurs, and local communities, particularly when social-media exposure accelerates visitation and increases pressures on viewpoints and access corridors.

Taken together, the results extend the broader valuation literature that treats managed agroecosystems as joint producers of marketed outputs and nonmarket benefits. Evidence from agricultural landscape valuation suggests that recreation benefits can be economically meaningful rather than marginal to rural land-use decisions (Bera & Nag, 2025; Fleischer & Tsur, 2000; Kaleji et al., 2025; Mameno et al., 2024; Song et al., 2020), and travel-cost methods are well suited to reveal these values where entrance fees are absent or unenforced (Blaine et al., 2015; Mäntymaa et al., 2021; Menendez-Carbo et al., 2020; Qiu & Fan, 2016; Tyllianakis, 2024; Van Sandt & McFadden, 2022). In practical terms, the demand and welfare evidence indicate that maintaining perceived landscape quality is not only an aesthetic concern: it is a determinant of visitation frequency and a channel through which stewardship generates public recreational benefits.

5. Conclusions

This study not only quantifies the recreational value of volcanic agricultural landscapes but also empirically demonstrates that, within informal tourism settings, the visual quality and ecosystem stewardship of "working landscapes" are key assets driving leisure demand. It thereby provides a case study supporting the closer integration of leisure economics with sustainable landscape management theory.

The research demonstrates that volcanic-slope farming landscapes can function as significant informal tourist destinations, generating substantial recreational value alongside agricultural production. By applying an on-site travel-cost framework, we establish that visitation patterns are governed by two key economic forces: sensitivity to generalized travel costs and a strong positive response to perceived landscape attractiveness. This confirms that in open-access settings, both accessibility constraints and experienced landscape quality are central drivers of visitation frequency. The findings reinforce the perspective that working agricultural landscapes provide welfare-relevant cultural ecosystem services (CES), where variations in perceived scenery translate into meaningful differences in recreation demand.

From a methodological standpoint, the study shows that welfare estimates derived from on-site surveys are sensitive to the sampling process and the assumed count-data distribution. Correcting for endogenous stratification (avidity bias) and overdispersion—as implemented in the preferred truncated and endogenously stratified negative binomial (TSNB) model—improves model fit and yields more conservative consumer surplus benchmarks. This underscores that simpler truncated Poisson estimates should not be treated as default welfare measures in destination-sampled studies. Presenting models that explicitly address on-site sampling mechanisms provides a more credible basis for policy interpretation and for comparing welfare estimates across studies of emerging, informally managed destinations.

The results substantiate a co-benefits perspective for volcanic uplands. Stewardship practices that maintain terraces, ground cover, and landscape coherence likely reinforce regulating ecosystem services (e.g., soil and water regulation) while simultaneously sustaining the scenic attributes that motivate visitation. Recognizing these non-market benefits strengthens the economic rationale for soil-conserving land management and can inform the design of benefit-sharing arrangements that include farmers and local micro-entrepreneurs. This is particularly relevant in institutionally informal settings where social-media exposure can rapidly intensify visitation pressures.

5.1. Policy Implications: From Recreational Value to Community Benefits

The estimated consumer surplus of approximately IDR 435 thousand per trip provides a tangible economic basis for policy and local governance. This value could, for instance, help partially offset terrace maintenance costs for farmers, supporting practical initiatives such as "ecological compensation" schemes where visitors contribute to conservation through small fees. In informal destinations like Selo and Cepogo, the quantified recreational benefit justifies creating community-based tourism funds to ensure that the economic benefits generated by visitation are shared more equitably among farmers and micro-entrepreneurs. Integrating these non-market values into land-use planning and local development strategies enables stakeholders to promote sustainable landscape stewardship while enhancing local livelihoods.

5.2. Recommendations for Future Research

Several limitations of the current study point to valuable priorities for future research. First, inference is conditional on an on-site sample; complementary off-site sampling would improve the representativeness of the visitor profile and support more defensible aggregation of benefits to the broader population. Second, landscape attractiveness is measured via self-reported ratings; integrating these subjective assessments with objective landscape metrics (e.g., heterogeneity indices, terrace density) and digital trace data could sharpen the causal interpretation of how landscape structure influences perceived quality and demand. Third, the destination is modeled as a single site; employing multi-destination choice frameworks could better capture the substitution patterns, route choices, and trip clustering typical of recreational travel in regional volcanic uplands. Finally, longitudinal or repeated measurement across seasons and following disturbance events (e.g., heavy rainfall, ashfall) would clarify how dynamic risk signals and evolving landscape conditions shape recreation demand and welfare in these dynamic socio-ecological systems.

CRediT Author Statement: Evi Irawan: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Supervision, Writing – original draft, Writing – review & editing, and Validation; **Syafika Rahma Bintari:** Validation, Investigation, Data curation, and Writing – original draft; **Chunyu Ying:** Validation, Formal analysis, and Writing – review & editing.

Data Availability Statement: Data will be made available on request.

Funding: This research was funded by Institute for Research and Community Service (LPPM), Universitas Sebelas Maret, Grant No. 194.2/UN27.22/PT.01.03/2024.

Conflicts of Interest: The authors declare no conflicts of interest.

IRB Statement: According to the guidelines of Institute for Research and Community Service, Universitas Sebelas Maret, formal ethical review was not required for this study because it consisted of minimal-risk survey research involving adults and did not collect sensitive personal data.

Informed Consent Statement: The study was conducted in accordance with accepted ethical standards for social research. All participants were informed about the purpose of the study, their voluntary involvement, and their right to decline participation or withdraw at any time. Informed consent was obtained prior to data collection, and all responses were anonymized and kept confidential.

Acknowledgments: The authors sincerely thank the anonymous reviewers for their insightful comments and constructive suggestions, which have substantially improved the clarity and robustness of this paper. The authors are solely responsible for any remaining errors or omissions. The authors also gratefully acknowledge the support of the Institute for Research and Community Service (LPPM), Universitas Sebelas Maret.

Abbreviations

The following abbreviations are used in this manuscript:

CES	Cultural ecosystem services
IDR	Indonesian rupiah
TP	Zero-truncated Poisson
TNB	Zero-truncated negative binomial
TSP	Truncated endogenously stratified Poisson
TSNB	Truncated endogenously stratified negative binomial

References

Amoako-Tuffour, J., & Martínez-Espiñeira, R. (2012). Leisure and the net opportunity cost of travel time in recreation demand analysis: An application to Gros Morne National Park. *Journal of Applied Economics*, *15*(1), 25–49. https://doi.org/10.1016/S1514-0326(12)60002-6

Assandri, G., Bogliani, G., Pedrini, P., & Brambilla, M. (2018). Beautiful agricultural landscapes promote cultural ecosystem services and biodiversity conservation. *Agriculture, Ecosystems & Environment*, *256*, 200–210. https://doi.org/10.1016/J.AGEE.2018.01.012

Auxtero, E., Madeira, M., & Sousa, E. (2004). Variable charge characteristics of selected Andisols from the Azores, Portugal. *CATENA*, *56*(1–3), 111–125. https://doi.org/10.1016/J.CATENA.2003.10.006

Bera, M., & Nag , P. K. (2025). Integrated natural resource management and sustainable livelihoods: A review of ecosystem-based approaches for sustainability and community resilience. *Agricultural & Rural Studies*, *3*(3), 23. https://doi.org/10.59978/ar03030017
Blaine, T. W., Lichtkoppler, F. R., Bader, T. J., Hartman, T. J., & Lucente, J. E. (2015). An examination of sources of sensitivity of consumer surplus estimates in travel cost models. *Journal of Environmental Management*, *151*, 427–436. https://doi.org/10.1016/J.JENVMAN.2014.12.033
Brata, A. G., de Groot, H. L. F., Rietveld, P., Resosudarmo, B. P., & Zant, W. (2021). Resilience toward volcanic eruptions: Risk perception and disaster microinsurance in Yogyakarta, Indonesia. *Sustainability*, *13*(16), 8912. https://doi.org/10.3390/SU13168912
Çakmak, E., Lie, R., & McCabe, S. (2018). Reframing informal tourism entrepreneurial practices: Capital and field relations structuring the informal tourism economy of Chiang Mai. *Annals of Tourism Research*, *72*, 37–47. https://doi.org/10.1016/J.ANNALS.2018.06.003
Cetin, N. I., Bourget, G., & Tezer, A. (2021). Travel-cost method for assessing the monetary value of recreational services in the Ömerli Catchment. *Ecological Economics*, *190*, 107192. https://doi.org/10.1016/J.ECOLECON.2021.107192
Chen, T.-P., Lee, K.-Y., Kabre, P. M., & Hsieh, C.-M. (2020). Impacts of educational agritourism on students' future career intentions: Evidence from agricultural exchange programs. *Sustainability*, *12*(22), 9507. https://doi.org/10.3390/SU12229507
Curtis, J. A. (2002). Estimating the demand for salmon angling in Ireland. *The Economic and Social Review*, *33*(3), 319–332.
Czajkowski, M., Giergiczny, M., Kronenberg, J., & Englin, J. (2019). The individual travel cost method with consumer-specific values of travel time savings. *Environmental and Resource Economics*, *74*, 961–984. https://doi.org/10.1007/s10640-019-00355-6
De Bauw, P., Van Asten, P., Jassogne, L., & Merckx, R. (2016). Soil fertility gradients and production constraints for coffee and banana on volcanic mountain slopes in the East African Rift: A case study of Mt. Elgon. *Agriculture, Ecosystems & Environment*, *231*, 166–175. https://doi.org/10.1016/J.AGEE.2016.06.036
de las Heras, A., Castillo-Villavicencio, E., & Islas-Espinoza, M. (2024). Repopulating and rewilding a historical agricultural landscape: 2030–2100 transition. *Agricultural & Rural Studies*, *2*(4), 10. https://doi.org/10.59978/ar02040022
Englin, J., & Shonkwiler, J. S. (1995). Estimating social welfare using count data models: An application to long-run recreation demand under conditions of endogenous stratification and truncation. *The Review of Economics and Statistics*, *77*(1), 104–112. https://doi.org/10.2307/2109996
Fleischer, A., & Tsur, Y. (2000). Measuring the recreational value of agricultural landscape. *European Review of Agricultural Economics*, *27*(3), 385–398. https://doi.org/10.1093/ERAE/27.3.385
Gosal, A. S., & Ziv, G. (2020). Landscape aesthetics: Spatial modelling and mapping using social media images and machine learning. *Ecological Indicators*, *117*, 106638. https://doi.org/10.1016/J.ECOLIND.2020.106638
Griffin, C. (2019). The Dieng 'hazardscape': A political ecology of vulnerability to natural hazards in Java's highlands. *Environmental Hazards*, *18*(1), 26–42. https://doi.org/10.1080/17477891.2018.1435407
Hindsley, P., Landry, C. E., & Gentner, B. (2011). Addressing onsite sampling in recreation site choice models. *Journal of Environmental Economics and Management*, *62*(1), 95–110. https://doi.org/10.1016/J.JEEM.2010.10.007
Hynes, S., Gaeven, R., & O'Reilly, P. (2017). Estimating a total demand function for sea angling pursuits. *Ecological Economics*, *134*, 73–81. https://doi.org/10.1016/J.ECOLECON.2016.12.024
Irham, I., Fachrista, I. A., Masyhuri, M., & Suryantini, A. (2022). Climate change adaptation strategies by Indonesian vegetable farmers: Comparative study of organic and conventional farmers. *The Scientific World Journal*, *2022*, 3590769. https://doi.org/10.1155/2022/3590769
Kaleji, M. K., Kazemi, H., Kamkar, B., Amirnejad, H., & Hosseinalizadeh, M. (2025). Economic valuation of ecosystem services in an agricultural landscape. *International Journal of Environmental Research*, *19*, 269. https://doi.org/10.1007/S41742-025-00936-Y
Kashyap, D., de Vries, M., Pronk, A., & Adiyoga, W. (2023). Environmental impact assessment of vegetable production in West Java, Indonesia. *Science of The Total Environment*, *864*, 160999. https://doi.org/10.1016/J.SCITOTENV.2022.160999
Landry, C. E., Lewis, A. R., Liu, H., & Vogelsong, H. (2016). Addressing onsite sampling in analysis of recreation demand: Economic value and impact of visitation to Cape Hatteras National Seashore. *Marine Resource Economics*, *31*(3), 301–322. https://doi.org/10.1086/686892
Lavigne, F., De Coster, B., Juvin, N., Flohic, F., Gaillard, J.-C., Texier, P., Morin, J., & Sartohadi, J. (2008). People's behaviour in the face of volcanic hazards: Perspectives from Javanese communities, Indonesia. *Journal of Volcanology and Geothermal Research*, *172*(3–4), 273–287. https://doi.org/10.1016/J.JVOLGEORES.2007.12.013
Ligot, N., Bogaert, P., Biass, S., Lobet, G., & Delmelle, P. (2023). Grain size modulates volcanic ash retention on crop foliage and potential yield loss. *Natural Hazards and Earth System Sciences*, *23*(4), 1355–1369. https://doi.org/10.5194/NHESS-23-1355-2023
Mameno, K., Arnberger, A., Shoji, Y., & Kubo, T. (2024). Recreational agroecosystem service value evidenced by mobile phone data: Implications for incentive enhancement in terraced paddy land. *Environmental Conservation*, *51*(3), 219–226. https://doi.org/10.1017/S0376892924000122
Mäntymaa, E., Jokinen, M., Juutinen, A., Lankia, T., & Louhi, P. (2021). Providing ecological, cultural and commercial services in an urban park: A travel cost–contingent behavior application in Finland. *Landscape and Urban Planning*, *209*, 104042. https://doi.org/10.1016/J.LANDURBPLAN.2021.104042
Martínez-Espiñeira, R., & Amoako-Tuffour, J. (2008). Recreation demand analysis under truncation, overdispersion, and endogenous stratification: An application to Gros Morne National Park. *Journal of Environmental Management*, *88*(4), 1320–1332. https://doi.org/10.1016/J.JENVMAN.2007.07.006
Menendez-Carbo, S., Ruano, M. A., & Zambrano-Monserrate, M. A. (2020). The economic value of *Malecón 2000* in Guayaquil, Ecuador: An application of the travel cost method. *Tourism Management Perspectives*, *36*, 100727. https://doi.org/10.1016/J.TMP.2020.100727

Moeskops, B., Buchan, D., Sukristiyonubowo, De Neve, S., De Gusseme, B., Widowati, L. R., Setyorini, D., & Sleutel, S. (2012). Soil quality indicators for intensive vegetable production systems in Java, Indonesia. *Ecological Indicators*, *18*, 218–226. https://doi.org/10.1016/J.ECOLIND.2011.11.011

Moeskops, B., Sukristiyonubowo, Buchan, D., Sleutel, S., Herawaty, L., Husen, E., Saraswati, R., Setyorini, D., & De Neve, S. (2010). Soil microbial communities and activities under intensive organic and conventional vegetable farming in West Java, Indonesia. *Applied Soil Ecology*, *45*(2), 112–120. https://doi.org/10.1016/J.APSOIL.2010.03.005

Moreira, B., Gonçalves, A., Pinto, L., Prieto, M. A., Carocho, M., Caleja, C., Barros, L. (2024). Intercropping systems: An opportunity for environment conservation within nut production. *Agriculture*, *14*(7), 1149. https://doi.org/10.3390/AGRICULTURE14071149

Nyelele, C., Keske, C., Chung, M. G., Guo, H., & Egoh, B. N. (2023). Using social media data and machine learning to map recreational ecosystem services. *Ecological Indicators*, *154*, 110606. https://doi.org/10.1016/J.ECOLIND.2023.110606

Parfitt, R. L. (2009). Allophane and imogolite: Role in soil biogeochemical processes. *Clay Minerals*, *44*(1), 135–155. https://doi.org/10.1180/claymin.2009.044.1.135

Plieninger, T., Bieling, C., Fagerholm, N., Byg, A., Hartel, T., Hurley, P., López-Santiago, C. A., Nagabhatla, N., Oteros-Rozas, E., Raymond, C. M., van der Horst, D., & Huntsinger, L. (2015). The role of cultural ecosystem services in landscape management and planning. *Current Opinion in Environmental Sustainability*, *14*, 28–33. https://doi.org/10.1016/J.COSUST.2015.02.006

Plieninger, T., Dijks, S., Oteros-Rozas, E., & Bieling, C. (2013). Assessing, mapping, and quantifying cultural ecosystem services at community level. *Land Use Policy*, *33*, 118–129. https://doi.org/10.1016/J.LANDUSEPOL.2012.12.013

Purwaningsih, R., Sartohadi, J., & Setiawan, M. A. (2020). Trees and crops arrangement in the agroforestry system based on slope units to control landslide reactivation on volcanic foot slopes in Java, Indonesia. *Land*, *9*(9), 327. https://doi.org/10.3390/LAND9090327

Qiu, S.-r., & Fan, S.-s. (2016). Recreational value estimation of suburban leisure agriculture: A case study of the Qianjiangyue agritourism farm. *Journal of Mountain Science*, *13*, 183–192. https://doi.org/10.1007/S11629-014-3286-8

Rosalina, P. D., Dupre, K., Wang, Y., Putra, I. N. D., & Jin, X. (2023). Rural tourism resource management strategies: A case study of two tourism villages in Bali. *Tourism Management Perspectives*, *49*, 101194. https://doi.org/10.1016/J.TMP.2023.101194

Saputra, D. D., Sari, R. R., Hairiah, K., Widianto, Suprayogo, D., & van Noordwijk, M. (2022). Recovery after volcanic ash deposition: Vegetation effects on soil organic carbon, soil structure and infiltration rates. *Plant and Soil*, *474*, 163–179. https://doi.org/10.1007/S11104-022-05322-7

Sarmiento-Soler, A., Rötter, R. P., Hoffmann, M. P., Jassogne, L., van Asten, P., Graefe, S., & Vaast, P. (2022). Disentangling effects of altitude and shade cover on coffee fruit dynamics and vegetative growth in smallholder coffee systems. *Agriculture, Ecosystems & Environment*, *326*, 107786. https://doi.org/10.1016/J.AGEE.2021.107786

Shahimoridi, R., Kazemi, H., Kamkar, B., Nadimi, A., Hosseinalizadeh, M., & Yeganeh, H. (2024). Economic valuation of ecosystem services in canola agroecosystems. *Landscape and Ecological Engineering*, *20*, 427–438. https://doi.org/10.1007/S11355-024-00603-Y

Shaw, D. (1988). On-site samples' regression: Problems of non-negative integers, truncation, and endogenous stratification. *Journal of Econometrics*, *37*(2), 211–223. https://doi.org/10.1016/0304-4076(88)90003-6

Siegel, L. A., Tussyadiah, I., & Scarles, C. (2023). Exploring behaviors of social media-induced tourists and the use of behavioral interventions as salient destination response strategy. *Journal of Destination Marketing & Management*, *27*, 100765. https://doi.org/10.1016/J.JDMM.2023.100765

Song, B., Robinson, G. M., Bardsley, D. K. (2020). Measuring multifunctional agricultural landscapes. *Land*, *9*(8), 260. https://doi.org/10.3390/LAND9080260

Suprayogo, D., van Noordwijk, M., Hairiah, K., Meilasari, N., Rabbani, A. L., Ishaq, R. M., & Widianto, W. (2020). Infiltration-friendly agroforestry land uses on volcanic slopes in the Rejoso Watershed, East Java, Indonesia. *Land*, *9*(8), 240. https://doi.org/10.3390/LAND9080240

Tang, H., Wang, R., Jin, X., & Zhang, Z. (2022). The effects of motivation, destination image and satisfaction on rural tourism tourists' willingness to revisit. *Sustainability*, *14*(19), 11938. https://doi.org/10.3390/SU141911938

Tieskens, K. F., Van Zanten, B. T., Schulp, C. J. E., & Verburg, P. H. (2018). Aesthetic appreciation of the cultural landscape through social media: An analysis of revealed preference in the Dutch river landscape. *Landscape and Urban Planning*, *177*, 128–137. https://doi.org/10.1016/J.LANDURBPLAN.2018.05.002

Tyllianakis, E. (2024). Heterogeneity in trips to green natural spaces: A travel cost approach across UK sites. *Journal of Outdoor Recreation and Tourism*, *46*, 100758. https://doi.org/10.1016/J.JORT.2024.100758

van Berkel, D. B., & Verburg, P. H. (2014). Spatial quantification and valuation of cultural ecosystem services in an agricultural landscape. *Ecological Indicators*, *37*, 163–174. https://doi.org/10.1016/J.ECOLIND.2012.06.025

van Ranst, E., Utami, S. R., Vanderdeelen, J., & Shamshuddin, J. (2004). Surface reactivity of Andisols on volcanic ash along the Sunda arc crossing Java Island, Indonesia. *Geoderma*, *123*(3–4), 193–203. https://doi.org/10.1016/J.GEODERMA.2004.02.005

Van Sandt, A., & McFadden, D. T. (2022). Navigating the corn maze: Customizing travel cost models to value market segments in heterogeneous industries. *Tourism Economics*, *28*(4), 899–919. https://doi.org/10.1177/1354816620976968

van Zanten, B. T., Zasada, I., Koetse, M. J., Ungaro, F., Häfner, K., & Verburg, P. H. (2016). A comparative approach to assess the contribution of landscape features to aesthetic and recreational values in agricultural landscapes. *Ecosystem Services*, *17*, 87–98. https://doi.org/10.1016/J.ECOSER.2015.11.011

Waha, K., Dietrich, J. P., Portmann, F. T., Siebert, S., Thornton, P. K., Bondeau, A., & Herrero, M. (2020). Multiple cropping systems of the world and the potential for increasing cropping intensity. *Global Environmental Change*, *64*, 102131. https://doi.org/10.1016/J.GLOENVCHA.2020.102131

Yue, J., Chen, S., Liu, Q., & Weng, Z. (2025). The impact of non-agricultural management experience on eco-friendly practices in Chinese family farms. *Agricultural & Rural Studies*, *3*(2), 18. https://doi.org/10.59978/ar03020010

Agricultural & Rural Studies

Article

Research on the Measurement of Common Prosperity in Rural Hainan Province

Qiong Yao [1,2†], Yujia Zhang [1,3†], Xiaoyan Zhao [1,2], Jian Wang [1,2,5], Guomin Zhou [1,4,5,6], Li Zhang [1,*], Jianhua Zhang [1,2,5,*]

1 Agricultural Information Institute, Chinese Academy of Agricultural Sciences, and Key Laboratory of Agricultural Big Data, Ministry of Agriculture and Rural Affairs, Beijing 100081, China; yaoqiong@caas.cn
2 National Nanfan Research Institute, Chinese Academy of Agriculture Science, Sanya 572024, China; zhaoxiaoyan01@caas.cn
3 Library and Archives (Hainan Provincial Research Center for Medical Information, University History Museum), Hainan Medical University, Haikou 571199, China; 1505302294@qq.com
4 Nanjing Institute of Agricultural Mechanization, Ministry of Agriculture and Rural Affairs, Nanjing 210014, China
5 National Agricultural Science Data Centre, Beijing 100081, China; wangjian01@caas.cn
6 Institute of Western Agriculture, Chinese Academy of Agricultural Sciences, Changji 831100, China; zhouguomin@caas.cn
* Correspondence: zhangli01@caas.cn (L.Z.); zhangjianhua@caas.cn (J.Z.)
† These authors contributed equally to this work.

Abstract: Based on the superimposed background of Hainan free trade port construction and Rural Revitalization Strategy, taking the panel data of 18 cities and counties in Hainan Province from 2013 to 2024 as samples, relying on the regional characteristics of tropical islands and the policy endowment of free trade port, this paper constructs a four-dimensional evaluation index system, and uses the global principal component analysis to systematically measure the comprehensive development level of rural common prosperity. Using the spatial autocorrelation model, Theil index and other tools to analyze regional development differences and spatial agglomeration characteristics, and using spatial lag model to empirically test the spatial correlation effect of various factors. The results show that the overall level of rural common prosperity in Hainan province continues to rise, and the characteristics of regional spatial differentiation are significant. The spatial-temporal pattern of Eastern agglomeration, western slow growth, and central ecological constraints has long been formed; There is a significant positive spatial agglomeration effect in the common prosperity of cities and counties, and urbanization, human capital, agricultural technology innovation, and other factors show a significant positive correlation. This study quantitatively identifies the development weaknesses and advantages of rural common prosperity in Hainan, enriches the theoretical system of rural common prosperity in tropical island regions, provides a scientific basis for the balanced development of urban and rural areas and regional collaborative governance in the background of a free trade port, and also offers a reference for promoting common prosperity in similar tropical regions in China.

Keywords: rural common prosperity; global principal component analysis; spatial autocorrelation model; regional disparities; coordinated development

Check for updates

Citation: Yao, Q., Zhang, Y., Zhao, X., Wang, J., Zhou, G., Zhang, L., & Zhang, J. (2026). Research on the Measurement of Common Prosperity in Rural Hainan Province. *Agriculture & Rural Studies*, *4*(2), 16. https://doi.org/10.59978/ar04020011

Received: 9 March 2026
Revised: 8 May 2026
Accepted: 15 May 2026
Published: 1 June 2026

1. Introduction

Achieving common prosperity remains the most arduous and crucial task in rural areas across China (J. Zheng et al., 2025). Against the strategic backdrop of rural revitalization, Chinese-style rural common prosperity is theoretically underpinned by the dual logic of high-quality development and equitable sharing, and is systematically structured along four core dimensions: prosperity level, sharing level, sustainability level, and commonness level. It emphasizes not only the expansion of rural economic scale but also the equalization of public services, regional coordinated development, and ecological sustainability (He, 2019). As the only tropical free trade port in China, Hainan Province boasts unique resource endowments, institutional advantages, and spatial characteristics (Gan et al., 2023). Promoting rural common prosperity in Hainan is not only a practical requirement for consolidating the achievements of poverty alleviation and advancing rural revitalization but also an important exploration for constructing a tropical island-style path toward common prosperity(E. Li et al., 2023; Mei et al., 2022).

Z. Chen and Yang (2023) established a linkage mechanism between income distribution and rural revitalization, arguing that narrowing the urban-rural gap is the core objective of common prosperity, requiring balanced development through rural education and social security system construction; S. Li et al. (2023) advocate prioritizing both economic growth and equitable distribution, emphasizing the pivotal role of equalizing public services in enhancing rural residents' development capabilities. Huang (2022) and Ma et al. (2023) respectively propose dual practical pathways of technological innovation and institutional breakthroughs, focusing on digital technology empowering agricultural modernization and household registration system reforms facilitating the urbanization of agricultural migrant workers. However, existing research predominantly focuses on national macro-level analyses or case studies of typical regions (such as the Yangtze River Delta and Pearl River Delta). Specialized studies specifically addressing Hainan Province's unique geographical environment (tropical island), economic structure (tropical agriculture + free trade port economy), and policy opportunities remain insufficient. There is a particular lack of systematic quantitative analysis concerning the spatiotemporal evolution characteristics of its rural common prosperity development level, the causes of regional disparities, and the positive association.

Despite the rich literature, notable theoretical and empirical gaps remain. First, studies on tropical island regions are extremely scarce. Hainan's unique mountainous–coastal spatial structure, ecological constraints, tropical agricultural structure, and free trade port institutional advantages make its rural development logic fundamentally different from inland regions, resulting in the weak applicability of existing theories and conclusions. Second, existing measurement frameworks rarely integrate the four dimensions of prosperity, sharing, sustainability, and commonness within a unified spatial econometric system, leading to incomplete identification of dynamic changes and spatial disparities. Third, few studies reveal the spatial spillover mechanisms and regional heterogeneity of driving factors under the dual background of free trade port construction and ecological priority development, which weakens the policy guidance for tropical regions.

Against this background, this study aims to fill the above gaps. Based on spatial externality theory, factor flow theory, and coordinated development theory, this paper constructs a four-dimensional evaluation system of rural common prosperity. Using the panel data of 18 counties and cities in Hainan from 2013 to 2024, it employs the global principal component analysis to measure the comprehensive level. Furthermore, it uses spatial autocorrelation, Theil index, MLD, and spatial lag model (SAR) to analyze spatial-temporal evolution, regional differences, spatial agglomeration, and driving mechanisms.

The theoretical contributions of this paper are threefold. First, it expands the theoretical framework of rural common prosperity to tropical island regions, enriches the spatial theory of common prosperity under ecological constraints, and improves the applicability of relevant theories. Second, it integrates prosperity, sharing, sustainability, and commonness into a unified measurement system, which enhances the comprehensiveness and accuracy of the evaluation. Third, it identifies the spatial spillover effects and regional heterogeneity of key driving factors, which provides a new analytical perspective for understanding the formation mechanism of spatial disparities in common prosperity. Practically, this study provides empirical support and decision-making references for Hainan to promote rural common prosperity, narrow regional gaps, and coordinate high-quality development and ecological protection, and also offers experience for other tropical and island regions worldwide.

2. Materials and Methods

2.1. Data Sources

Guided by principles of data availability, continuity, and authority, panel data from 18 cities and counties (including county-level cities) in Hainan Province spanning 2013–2024 were selected as the research sample. This provides foundational data support for quantitative analysis and empirical testing in related studies. Data sources include the Hainan Statistical Yearbook (2013–2024) and the Hainan Provincial Rural Economic Statistical Annual Report. This ensures the systematic nature and comparability of indicator data, effectively safeguarding the continuity and integrity of the panel data, thereby laying a robust data foundation for the reliability and scientific validity of subsequent research conclusions.

2.2. Research Methodology

2.2.1. Global Principal Component Analysis

The global principal component analysis method is widely used in multi-dimensional comprehensive evaluation research. Its core advantage is that it can objectively determine the weight according to the correlation between indicators and avoid the deviation of subjective weight. At the same time, the time development trend and individual differences of indicators can be taken into account to ensure the comparability and accuracy of cross-time and cross-regional data (J. Chen & Lin, 2023).

At the same time, in order to eliminate the interference of extreme observation values on the measurement results, this paper uses the boxplot method to identify abnormal values, and carries out 1% horizontal bilateral tailing processing on all continuous indicators, eliminating the extreme values beyond the upper and lower 1% quantiles, so as to ensure the stable data distribution and robust results. To test whether the indicators are suitable for global principal component analysis, the KMO test and Bartlett's sphericity test are carried out in this paper. The results showed that KMO = 0.812, greater than 0.7; Bartlett's sphericity test ($p < 0.001$) shows that there is a strong correlation between the indicators, which is in line with the applicable premise of global principal component analysis.

Using the data after standardization and outlier processing, the comprehensive index of rural common prosperity of cities and counties in Hainan Province from 2013 to 2024 is calculated by using the global principal component analysis method.

2.2.2. Methodology for Constructing Spatial Weighting Matrices

The spatial weight matrix characterizes the degree of spatial association between regions and forms the foundation of spatial econometric analysis. Drawing upon Otto and Steinert (2023), this study constructs two types of spatial weight matrices:

Geographical Distance Spatial Weight Matrix (W_1): Measures association strength using the reciprocal square of the Euclidean distance between regional centers, with greater proximity yielding higher weights. The equation (1) is:

$$W_{ij} = \begin{cases} 1/D_{ij}^2, \, i \neq j \\ 0, \, i \neq j \end{cases} \tag{1}$$

Where: $n \times n$ The influence factor for row i , column j of the spatial weight matrix W_1 is represented by $W_{ij}.W_{ij}$ fully captures the spatial relationship of geographical distances between counties/cities in terms of Chinese-style rural common prosperity development levels; D_{ij} denotes the distance between region i and region j.

Spatial adjacency weight matrix (W_2): Assigns a value of 1 if two regions share a common boundary, and 0 otherwise. The equation (2) is:

$$W_{ij} = \begin{cases} 1, \, i \text{ adjacent to } j \\ 0, \, i \text{ is not adjacent to } j \end{cases} \tag{2}$$

2.2.3. Construction of the Spatial Lag Model

To examine the spatial correlation and influencing factors of rural common prosperity development levels in Hainan Province, the following spatial autoregressive (SAR) model is established (Anselin, 1998). The equation (3) is:

$$CD_{it} = \alpha_0 + \rho \sum_{j=1}^{n} W_{ij} CD_{it} + \alpha_1 lnUL_{it} + \alpha_2 lnHUL_{it} + \alpha_3 lnNPA_{it} + \alpha_4 lnGSA_{it} + \alpha_5 lnDE_{it} + \alpha_6 lnTPE_{it} + \alpha_7 lnBE_{it} + \mu_i + V_t + \varepsilon_{it} \tag{3}$$

Where the level of Chinese-style rural common prosperity is denoted by CD, provincial fixed effects and time fixed effects are represented by μ_i and V_t respectively, the random disturbance term is denoted by ε_{it} , the spatial lag coefficient is denoted by ρ , the spatial weight in the spatial spillover process of Chinese-style rural common prosperity is denoted by W, the coefficient of the explanatory variable is denoted by α, UL denotes urbanization rate, HCL represents human capital, NPA indicates agricultural technology innovation, GSA denotes agricultural production subsidies, DE denotes the proportion of agricultural processing industries, TPE represents the level of private sector development, and BE indicates the business environment (Including: market environment, construction of the rule of law, level of openness, government services).

2.2.4. Measurement Methodology

To comprehensively reflect regional disparities in rural common prosperity within Hainan Province, three indicators are employed: the Theil index, the geometric mean of log-deviations (GE), and the mean of log-deviations index (MLD; Lu et al., 2018; Yu et al., 2020). As these three indicators are respectively more sensitive to data fluctuations at higher, medium, and lower levels within the variable data, a combination of all three is selected to objectively reflect disparities in Hainan's rural common prosperity.

Theil: Sensitive to variations in high-income groups. The equation (4) is:

$$\text{Theil} = \sum_{i=1}^{n} p_i \frac{x_i}{u} ln(\frac{x_i}{u}) \tag{4}$$

Where x_i, n, and u denote the same meanings as in the Gini coefficient, p_i represents the proportion of the population of the i-th city or county relative to the provincial total. Theil indices range between 0 and 1, with higher values indicating greater regional inequality.

GE: Sensitive to disparities within the middle-income group, it decomposes into intra-regional variation (W_t) and inter-regional variation (B_t). The equations (5) and (6) are:

$$GE_t = \frac{1}{N}\sum_{i=1}^{N} ln\frac{y_{it}}{\mu_t} \tag{5}$$

$$GE_t = W_t + B_t = \sum_{k=1}^{n} \frac{\mathrm{N_k}}{N} \times GE_{kt} + \sum_{k=1}^{n} \frac{\mathrm{N_k}}{N} \times ln\frac{\mu_{kt}}{\mu_t} \tag{6}$$

where i denotes a county, t denotes a year; y_{it} represents the rural common prosperity development index for county i at time t, μ_t denotes the mean rural common prosperity development index for the province at time t, and N is the number of counties in the province. k denotes a region, W_t represents the intra-regional variation in rural common prosperity levels at time t, and B_t denotes the inter-regional variation in rural common prosperity levels at time t; $y_{ikt}\mu_{kt}$ denotes the rural common prosperity index for county i within region k at time t; represents the mean rural common prosperity index for region k at time t; n denotes the number of regions; N_k denotes the number of counties within region k.

MLD: Sensitive to Low-Income Group Disparities. The equation (7) is:

$$\text{MLD} = \sum_{i=1}^{n} p_i ln\frac{x_i}{u} \tag{7}$$

where x_i, n, and u denote the same meanings as in the Gini coefficient, and p_i represents the proportion of the population of the i-th county/city relative to the total provincial population. The MLD index ranges between 0 and 1, with higher values indicating greater inter-regional inequality.

2.3. Variable Explanation

2.3.1. Evaluation Indicator System

Common prosperity in rural areas is not only the realistic demand to achieve common prosperity for all people, but also the only way to promote Chinese-style modernization (Ji et al., 2023; J. Sun & Zhao, 2023). Based on spatial externality theory, factor flow theory, and collaborative development theory, interregional factor flow, industrial linkage, and policy transmission determine that rural common prosperity development has a significant spatial spillover effect (N. Li et al., 2023; Zhang et al., 2023). Digital economy, social security, and other core factors are important supports for rural common prosperity, and their spatial spillover effect is the key link to promote regional collaborative development. In order to meet the practical needs of common prosperity, this paper refers to the policy guidance of "solidly promoting common prosperity" in the report of the 20th National Congress of the Communist Party of China, and combines the characteristics of tropical agriculture and free trade port in Hainan Province to build a four-dimensional evaluation index system including richness, sharing, sustainability and commonness (Q. Wang et al., 2023; Yue & Yuan, 2023). The index attributes are divided according to "positive (the higher the index value, the better the common prosperity)/negative (the higher the index value, the worse the common prosperity) ," as shown in Table 1. Among them, digital economy and social security are the core related factors, and their spatial spillover effects on rural common prosperity are specifically manifested as follows: the digital economy is characterized by the digital inclusive financial index as the core, relying on the advantages of free trade port digital policy, breaking the geographical barriers, promoting the cross regional flow of agricultural technology, market information and financial resources, which not only helps the rapid diffusion of tropical agricultural technologies such as breeding in southern China and planting typhoon resistant crops, but also reduces the threshold of rural financing through digital Inclusive Finance. Its spillover effects can radiate to surrounding cities and counties, and drive the overall improvement of regional agricultural production efficiency and rural development quality. Social security covers indicators such as the per capita transfer income of rural residents and the urban-rural minimum living security ratio. The regional development gap is adjusted through fiscal transfer payment and public service equalization policies. The improved social security system in developed cities and counties such as Haikou and Sanya can promote the improvement of people's livelihood security level in less developed cities and counties in the central and western regions through policy demonstration, resource sharing, experience replication and other forms, forming a positive spatial spillover of "high security areas drive low security areas," and helping the coordinated promotion of common prosperity in rural areas throughout the region.

Table 1. Evaluation Indicator System for Rural Common Prosperity Development Levels.

Primary Indicators	Secondary Indicators	Tertiary Indicators (Unit)	Attribute
Level of Prosperity	Education and Culture	Per capita expenditure on education and cultural activities among rural residents (yuan)	Positive
		Average Years of Education Among Rural Residents (years)	Positive
	Ecological Environment	Rural Greening Coverage Rate (%)	Positive
	Social security	Per capita transfer income for rural residents (¥10,000)	Positive
		Ratio of urban residents receiving minimum living allowance to rural residents receiving minimum living allowance (%)	Negative
	Healthcare	Number of healthcare beds per thousand rural residents (units)	Positive
	Income and Consumption	Rural residents' per capita disposable income (¥10,000)	Positive
		Per capita consumption expenditure of rural residents (yuan)	Positive
Sharing Degree	Educational and cultural disparity	Average Years of Education for Urban Residents / Average Years of Education for Rural Residents (%)	Negative
		Urban per capita expenditure on education and recreation / Rural per capita expenditure on education and recreation (%)	Negative
	Income and Consumption Gap	Urban per capita disposable income / Rural per capita disposable income (%)	Negative
		Urban per capita consumption expenditure / Rural per capita consumption expenditure (%)	Negative
	Healthcare Gap	Number of healthcare beds per thousand people in urban areas / Number of healthcare beds per thousand people in rural areas (%)	Negative
	Ecological environment disparity	Urban green coverage rate / Rural green coverage rate (%)	Positive
	Social security disparity	Urban residents' subsistence allowance rate/Rural residents' subsistence allowance rate (%)	Negative
		Urban per capita transfer income/Rural per capita transfer income (%)	Negative
Sustainability	Digitalization and Financial Development Level	Digital Inclusive Finance Index	Positive
	Rural Industrial Development	Gross Agricultural Output Value / Total Cropped Area (¥ billion / hectares)	Positive
	Fiscal Support for Agriculture	Total agricultural expenditure (billion yuan)	Positive
	Rural industrial income	Average Wage of Employees in Agriculture, Forestry, Animal Husbandry, and Fisheries (yuan)	Positive
	Rural ecological environment	Carbon emissions per unit of agricultural, forestry, animal husbandry, and fishery output (tonnes per 100 million yuan)	Negative
Commonality	Farmers' income disparity	Rural Gini coefficient	Negative
	Urbanization of Farmers	Urbanization rate (%)	Positive
	Minimum Livelihood Guarantee for Farmers	Rural Minimum Living Allowance (yuan/person/year)	Positive
	Urban-rural income disparity	Urban-Rural Income Ratio	Negative

2.3.2. Index Applicability Test

It can be seen from the test results (Table 2) that the KMO test value is greater than 0.7, the Bartlett's sphericity test is significant at the 1% level, and the index commonality and cumulative variance contribution rate meet the standard requirements, indicating that there is a strong correlation between each index, which is suitable for global principal component analysis.

Table 2. Index Applicability Test Results.

Test Item	Statistic	Critical Value / Significance	Test Item
KMO Test	0.812	≥ 0.7	Suitable for principal
Bartlett Test of Sphericity	χ^2 = 1468.32	p < 0.001	Significant correlation
Mean Communality	0.78	≥ 0.6	Sufficient information
Cumulative Variance Contribution Rate	78.35%	≥ 70%	Effective extraction

Combined with the characteristics of Tropical Agriculture in Hainan Province, the advantages of free trade port policy and the positioning of ecological protection, the factors suitable for the measurement of rural common prosperity were screened (Kakwani et al., 2022; Zhang et al., 2023): Firstly, Urbanization Rate (UL), represented by the ratio of urban population to total regional population across Hainan's cities and counties. Hainan's urbanization exhibits a "high coastal, low inland" pattern, with coastal cities like Haikou and Sanya exceeding 60% urbanization, while inland ecological counties remain below 40%. Advancing urbanization is crucial for facilitating the flow of urban-rural factors and ensuring the trickle-down effects of free trade port dividends (He & Chu, 2023). Second, Human Capital (HCL), measured by the average years of schooling among rural residents in each city/county, supplemented by participation rates in South Breeding technology training. Rural residents in central Li and Miao ethnic counties exhibit lower average years of schooling, while the South Breeding industry requires technical personnel. Increased investment can enhance farmers' employability (W. Zheng & Chen, 2023). Third, Agricultural Technology Innovation (NPA), measured by patent applications in tropical agriculture and South Breeding, including patents for typhoon-resistant crops and deep processing (T. Zhao et al., 2022). This addresses the issues of low-risk resilience and low added value in tropical agricultural products (Zhan & Li, 2022). Fourth, Agricultural Production Subsidies (GSA), represented by the proportion of tropical agriculture-specific subsidies within total fiscal expenditure on agriculture, forestry, and water affairs (Bai et al., 2022). This includes subsidies for rubber seed varieties and ecological agriculture, stabilizing farmer incomes and guiding social capital (Yang, 2023). Fifth, the proportion of agricultural processing industries (DE), measured by the output value of tropical agricultural processing relative to the total output value of agriculture, forestry, animal husbandry, and fisheries (N. Sun et al., 2024). This focuses on processing betel nuts, coconuts, and other products, absorbing rural labor, and enhancing added value (Shen et al., 2025). Sixth, private sector development level (TPE), represented by the output value of private enterprises in tropical agriculture and rural tourism as a proportion of regional GDP (Z. Zhao et al., 2024). The private sector accounts for 68% of non-agricultural employment in rural areas, broadening income sources. Seventh, the business environment (BE) is constructed across market (incremental social financing/GDP, etc.), rule of law (patent disputes/authorizations, etc.), openness (cross-border agricultural e-commerce imports/exports/GDP, etc.), and governance (proportion of expenditure on South China Sea Breeding R&D, etc.) dimensions, with the index calculated using the entropy method (L. Wang & Sun, 2025; Zhu et al., 2020).

In order to avoid multicollinearity between variables affecting the regression results, all explanatory variables were tested by the VIF test. The results show that the VIF values of each variable are between 1.25 and 2.43, with a mean value of 1.86, which is far less than the critical value of 3, indicating that there is no serious multicollinearity problem in the model.

3. Analysis of Research Findings

3.1. Temporal Evolution of Rural Common Prosperity Development Levels in Hainan Province

The comprehensive index for common prosperity in Hainan's rural areas across all cities and counties from 2013 to 2024 was calculated using global principal component analysis, with results presented in Figure 1 and Table 3. From 2013 to 2024, Hainan's rural common prosperity composite index rose from 0.38 to 0.61, exhibiting three distinct phases: rapid ascent (2013–2016), steady optimization (2017–2020), and quality enhancement with efficiency gains (2021–2024).

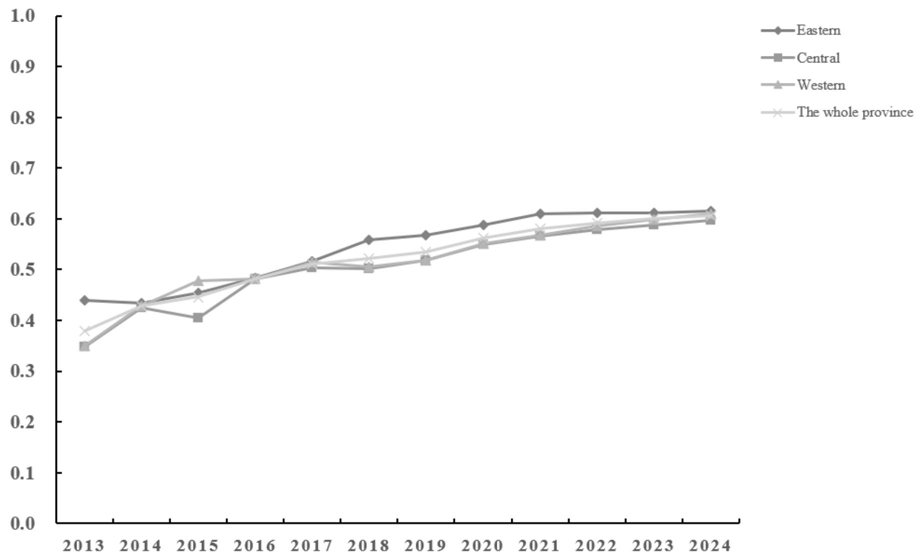


Figure 1. Development Levels of Common Prosperity in Rural Areas Across Hainan Province, 2013–2024.

Table 3. Comprehensive Index of Rural Common Prosperity Development Levels in Hainan's 18 Municipalities and Counties (including county-level cities), 2013–2024

	Region	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**	**2022**	**2023**	**2024**
Eastern	Haikou	0.442	0.438	0.433	0.483	0.517	0.562	0.573	0.592	0.614	0.616	0.616	0.619
	Sanya	0.438	0.436	0.429	0.485	0.520	0.562	0.572	0.592	0.615	0.618	0.61	0.615
	Wenchang	0.449	0.430	0.595	0.480	0.513	0.559	0.569	0.587	0.608	0.608	0.611	0.612
	Qionghai	0.436	0.425	0.447	0.481	0.515	0.556	0.566	0.586	0.603	0.605	0.613	0.614
	Wanning	0.439	0.439	0.411	0.478	0.513	0.556	0.564	0.585	0.616	0.615	0.616	0.617
	Lingshui	0.435	0.432	0.411	0.491	0.517	0.557	0.564	0.586	0.604	0.605	0.609	0.612
Central	Wuzhishan	0.352	0.425	0.439	0.460	0.508	0.510	0.535	0.553	0.563	0.581	0.597	0.602
	Dingan	0.351	0.422	0.464	0.483	0.507	0.502	0.513	0.548	0.562	0.58	0.582	0.592
	Tunchang	0.349	0.420	0.405	0.480	0.506	0.501	0.511	0.545	0.574	0.578	0.586	0.597
	Qiongzhong	0.350	0.427	0.421	0.487	0.496	0.498	0.517	0.552	0.568	0.574	0.591	0.598
	Baoting	0.341	0.426	0.363	0.488	0.504	0.498	0.517	0.550	0.561	0.576	0.59	0.596
	Baisha	0.348	0.424	0.340	0.493	0.505	0.496	0.521	0.549	0.568	0.579	0.585	0.593
West-ern	Danzhou	0.355	0.422	0.493	0.485	0.517	0.514	0.525	0.549	0.57	0.609	0.612	0.619
	Dongfang	0.356	0.442	0.498	0.482	0.507	0.510	0.524	0.556	0.563	0.58	0.608	0.615
	Chengmai	0.348	0.440	0.498	0.478	0.512	0.511	0.521	0.566	0.562	0.581	0.595	0.609
	Lingao	0.346	0.437	0.453	0.483	0.520	0.516	0.523	0.546	0.569	0.569	0.579	0.607
	Ledong	0.349	0.430	0.495	0.477	0.512	0.490	0.509	0.545	0.574	0.567	0.579	0.605
	Changjiang	0.343	0.402	0.429	0.479	0.515	0.489	0.51	0.542	0.563	0.587	0.591	0.603

2013–2016: The index grew at an average annual rate of 4.8%, driven by the development of rural tourism spurred by the construction of Hainan International Tourism Island (e.g., Boao rural tourism in Qionghai, Xinglong coffee estates in Wanning), leading to an increase in the proportion of non-agricultural income for farmers.

2017–2020: The index recorded an average annual growth rate of 2.1%. This deceleration stemmed from tropical agriculture being impacted by typhoon disasters (such as Typhoon

Mangkhut in 2018) and market fluctuations (e.g., declining rubber prices), while the process of equalizing public services between urban and rural areas entered a critical phase.

2021–2024: Annual average growth rate of 3.5%. Policy dividends from the Free Trade Port initiative (e.g., cross-border e-commerce boosting agricultural exports, southern seed industry enhancing farmer incomes), combined with the deepening of the rural revitalization strategy, accelerate progress towards common prosperity.

Table 3 indicates that the regional pattern of "higher in the east, lower in the center, and catching up in the west" persists, though disparities are gradually narrowing.

Eastern Region: The 2024 composite index average stands at 0.615, with Haikou (0.619), Sanya (0.618), and Wanning (0.617) ranking as the top three. Leveraging coastal tourism and tropical high-efficiency agriculture (such as coconut processing in Wenchang and cherry tomato cultivation in Lingshui), this region leads in both farmer income levels and public service quality.

Central Region: The 2024 composite index averaged 0.595, an improvement of 0.245 since 2013. Wuzhishan (0.602) and Qiongzhong (0.598) recorded the fastest growth, benefiting from ecological compensation policies (such as national park construction subsidies) and the development of distinctive industries (e.g., Baisha green tea and Qiongzhong green oranges).

Western Region: The 2024 composite index averaged 0.608, an increase of 0.252 from 2013. Danzhou (0.619) and Dongfang (0.615) ranked among the province's top five, leveraging Yangpu Port's logistics to boost agricultural exports (e.g., Chengmai Fucheng oranges, Lingao seafood), with notable achievements in industrial structure optimization.

3.2. Spatial Evolution Trends in Hainan Province's Rural Common Prosperity Development Levels

From 2013 to 2024, the dynamic evolution of development disparities in rural common prosperity across Hainan Province exhibited an overall pattern of "fluctuating yet gradually stabilizing." The consistent trends observed in the three key indicators—Theil, GE, and MLD—collectively corroborate this trajectory. As illustrated in Figure 2, all three indicators exhibited a phased upward trajectory between 2013 and 2015: the Theil index rose from 0.011 to 0.018, GE increased from 0.007 to 0.012, while the MLD rose from 0.006 to 0.011. This reflects a short-term widening of inter-regional development disparities, closely linked to the asynchronous income growth observed in certain cities and counties (such as leapfrog growth in some areas, contrasting with decelerating growth in others). From 2015 to 2020, indicators showed a sustained decline: the Theil Index fell to 0.012, while GE and MLD narrowed to 0.008 and 0.007, respectively. This indicates that policies such as the Rural Revitalization Strategy effectively promoted balanced regional development, gradually unleashing income growth momentum in low-income cities and counties and significantly narrowing disparities. From 2020 to 2021, indicators experienced a slight rebound: the Theil index rose to 0.013, while GE and MLD increased to 0.009 and 0.008, respectively. This is likely attributable to uneven short-term recovery caused by external shocks, such as the pandemic's differential impacts on agricultural production and rural employment. From 2021 to 2024, disparities re-entered a steady narrowing trajectory. By 2024, the Theil index had decreased to 0.01, with GE and MLD stabilizing at 0.007 and 0.006, respectively—essentially reverting to 2013 levels. This demonstrates the sustained enhancement of regional development coordination under long-term policy effects. Overall, while the gap in rural common prosperity development in Hainan Province exhibited phased fluctuations over the 12-year period, a clear long-term convergence trend emerged. The consistent movement across the three indicator categories not only validates the objectivity of the gap's evolution but also highlights the pivotal role of policy intervention in promoting regional equilibrium. This provides quantitative evidence for understanding the advancement pathways of rural common prosperity in tropical regions.

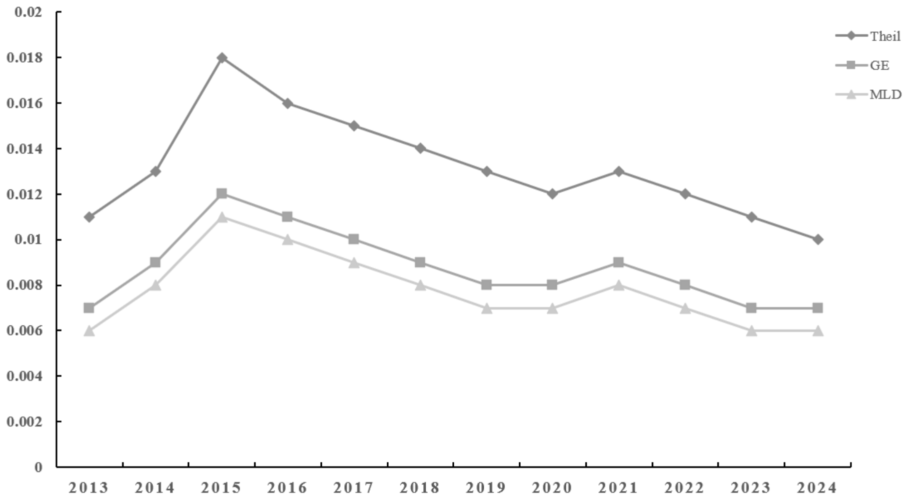


Figure 2. Trend in Development Disparities for Rural Common Prosperity in Hainan Province, 2013–2024.

3.3. Spatial Correlation Analysis of Rural Common Prosperity Development Levels in Hainan Province

To conduct an in-depth analysis of the spatial distribution characteristics of common prosperity development across cities, counties, towns, and villages in Hainan Province from 2013 to 2024, a spatial weighting matrix was constructed to characterize spatial agglomeration and dispersion patterns. The global Moran's I index was employed to examine spatial agglomeration patterns of common prosperity from 2013 to 2024, with results presented in Table 4.

Table 4. Global Moran's Index of Rural Common Prosperity Development Levels in Hainan Province, 2013–2024.

Indicator	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024
Morlan Index	0.28	0.32	0.30	0.35	0.41	0.45	0.48	0.46	0.50	0.49	0.52	0.51
P-value	0.07	0.05	0.06	0.04	0.03	0.02	0.01	0.02	0.01	0.01	0.01	0.01
Z-value	1.81	1.98	1.85	2.09	2.32	2.51	2.67	2.59	2.83	2.79	2.91	2.87

The global Moran's I index remained positive throughout the period (0.28–0.52), with P-values ≤ 0.05 and Z-values ≥ 1.98 after 2014. This indicates:

The positive correlation characteristics in spatial distribution are significant: counties and cities with high common prosperity levels (e.g., Haikou, Sanya) tend to border other high-value areas (e.g., Haikou with Wenchang, Sanya with Lingshui), while counties with lower levels (e.g., Wuzhishan, Baisha) tend to be adjacent to similarly low-performing counties (e.g., Wuzhishan with Qiongzhong, Baisha with Changjiang), forming a pattern of "high-value clustering and low-value linkage."

The intensity of spatial agglomeration continues to increase: the Moran's I coefficient rose from 0.28 in 2013 to 0.52 in 2023 (slightly declining to 0.51 in 2024). This indicates enhanced interregional factor flows (e.g., technology transfer from eastern to western regions, labor migration from central to eastern areas) and policy coordination (e.g., the Ring Island Tourist Highway linking eastern and western regions), with continuous strengthening of spatial connectivity between counties and cities.

3.4. Spatial Effects and Influencing Factors of Common Prosperity Development in Rural Hainan

Before spatial econometric regression analysis, in order to avoid the interference of multicollinearity between explanatory variables on the estimation results, this paper first conducts a variance inflation factor (VIF) test on each variable, and the results are shown in Table 5. The Vif values of each variable are between 1.25 and 2.43, with a mean value of 1.86, which is far below the empirical critical value of 3, indicating that there is no serious multicollinearity problem in the model, and the selection of variables and the setting of the model are reasonable, so spatial regression analysis can be further carried out.

Table 5. Model Variable Multicollinearity VIF Test Results.

Variable	VIF	1/VIF
Ln UL	2.43	0.41
Ln HCL	2.17	0.46
Ln NPA	1.94	0.52
Ln GSA	1.82	0.55
Ln DE	1.65	0.61
Ln TPE	1.43	0.70
Ln BE	1.25	0.80
Mean	1.86	—

The spatial lag model (SAR) estimation results can clearly reflect the spatial correlation characteristics of rural common prosperity in Hainan Province and the differences in the role of various variables. It can be seen from Table 6 that under the two types of spatial weight matrix settings, the spatial correlation coefficient ρ of the rural common prosperity level is significantly positive at the 1% level, indicating that there is a significant spatial positive correlation between the rural common prosperity of cities and counties in Hainan Province, and the high-level counties show the characteristics of agglomeration and distribution in space. Taking the final selected geographical adjacent spatial weight matrix (W2) as an example, the regression coefficients of urbanization rate (In UL), human capital (In HCL), agricultural technology innovation (In NPA), agricultural production subsidies (In GSA), business environment (In BE) and the level of private eco-nomic development (In TPE) are significantly positive, showing a significant positive correlation with rural common prosperity; The proportion of agricultural products processing industry (In DE) coefficient is positive but fails to pass the significance test, indicating that its positive enabling effect has not been effectively released, and its supporting effect on regional common prosperity still needs to be improved.

Table 6. Estimation Results of Spatial Lag Model (SAR) for the Development Level of Rural Common Prosperity in Hainan Province.

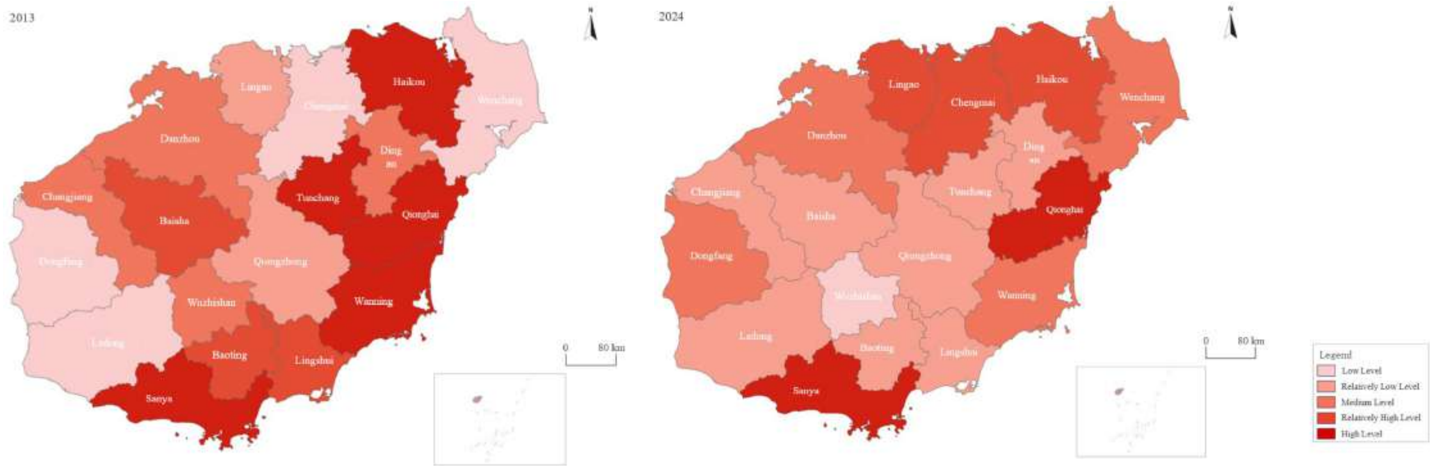
	Time Fixed Effects		Individual fixed effects		Dual Fixed Effects	
	W_1	W_2	W_1	W_2	W_1	W_2
In UL	0.36^{***}(3.75)	0.41^{***}(4.28)	0.38^{***}(3.90)	0.43^{***}(4.42)	0.39^{***}(4.05)	0.44^{***}(4.56)
In HCL	0.30^{***}(3.45)	0.35^{***}(3.93)	0.32^{***}(3.59)	0.37^{***}(4.06)	0.33^{***}(3.72)	0.38^{***}(4.19)
In NPA	0.32^{***}(3.58)	0.37^{***}(4.15)	0.34^{***}(3.75)	0.39^{***}(4.28)	0.35^{***}(3.90)	0.40^{***}(4.41)
In GSA	0.33^{***}(3.68)	0.38^{***}(4.22)	0.35^{***}(3.82)	0.40^{***}(4.35)	0.36^{***}(3.95)	0.41^{***}(4.48)
In DE	0.29^{***}(3.46)	0.34^{***}(3.99)	0.31^{***}(3.60)	0.36^{***}(4.13)	0.32^{***}(3.73)	0.37^{***}(4.26)
In TPE	0.28^{***}(3.47)	0.33^{***}(4.00)	0.30^{***}(3.61)	0.35^{***}(4.14)	0.31^{***}(3.75)	0.36^{***}(4.28)
In BE	0.31^{***}(3.70)	0.36^{***}(4.23)	0.33^{***}(3.84)	0.38^{***}(4.36)	0.34^{***}(3.97)	0.39^{***}(4.50)
ρ	0.48^{***}(4.02)	0.55^{***}(4.67)	0.50^{***}(4.18)	0.57^{***}(4.83)	0.52^{***}(4.35)	0.59^{***}(5.01)
$σ^2$	0.082	0.076	0.079	0.071	0.075	0.068
R^2	0.76	0.81	0.78	0.83	0.8	0.85
LogL	−186.3	−172.5	−180.7	−165.9	−175.2	−159.8
N	216	216	216	216	216	216

Note: values in parentheses refer to z-statistics.
*** denotes significance at the 1% level.

In order to ensure the reliability of the empirical results, this paper carries out the robustness test from three aspects: (i) replace the spatial weight matrix, and use the geographical adjacent matrix and geographical distance matrix regression respectively; (ii) Lag the core explanatory variable for one period to alleviate the endogenous problem caused by potential two-way causality; (iii) Two core cities, Haikou and Sanya, were excluded for sub sample regression. The results show that the sign, size, and significance of the core variable coefficient are consistent, indicating that the benchmark regression conclusion is robust and reliable.

3.5. Spatial Disparity Analysis of Common Prosperity Development Levels in Hainan's Rural Areas

Integrating rural development theory with quantitative data, a systematic analysis of the spatial disparity evolution in Hainan's rural common prosperity over 12 years can be conducted by comparing the coloring characteristics of its cities and counties in 2013 and 2024, as illustrated in Figure 3 (a) and (b). The 2013 spatial map of rural common prosperity in Hainan exhibited pronounced gradient differentiation: core cities like Haikou and Sanya, along with eastern coastal counties, predominantly featured dark hues, whereas central mountainous counties such as Wuzhishan and Baisha displayed markedly lighter tones. This spatial differentiation corroborates the quantitative trend of widening rural development disparities observed between 2013 and 2015. From an economic geography perspective, eastern coastal counties leveraged their maritime locations and tourism resources to pioneer rural industrial clusters. These included Wenchang's aerospace tourism-derived sectors and Qionghai's Boao Forum-linked agricultural exhibitions, where high-value-added industries drove rapid income growth. Conversely, mountainous central regions, constrained by topography and infrastructure limitations, face bottlenecks in scaling up specialized agriculture and logistics cost burdens. Poor agricultural product distribution channels in autonomous counties like Qiongzhong and Baoting have resulted in sluggish income growth. This "coastal-central" development divide manifests as a distinct spatial color boundary, serving as a tangible spatial representation of initial developmental disparities.



(a) (b)

Figure 3. (a) Distribution Map of Rural Common Prosperity in Hainan Province (2013); **(b)** Distribution Map of Rural Common Prosperity in Hainan Province (2024).

The 2024 spatial map of rural common prosperity in Hainan Province reveals pronounced characteristics of balanced evolution: darker shades extend into central regions while lighter areas diminish, with the color gradient between coastal and central counties markedly softening – confirming the overall trend of narrowing developmental disparities. This spatial reconfiguration stems from the synergistic effects of dual mechanisms: Firstly, the balanced orientation of policy interventions has yielded substantial results. Within the framework of the rural revitalization strategy, the central ecological zone has received systematic support. Baisha Li Autonomous County has cultivated distinctive industries such as organic tea and southern medicinal herbs through the pathway of "realizing the value of ecological products." Government-led construction of cold chain logistics systems has overcome spatial constraints on agricultural product sales. Qiongzhong Li and Miao Autonomous County established the "Wanling Cold Chain Logistics Centre," achieving efficient production-to-market integration through a province-wide distribution network. Addressing infrastructure deficiencies has enhanced industrial competitiveness. Secondly, the spatial spillover effects of industrial synergy have become pronounced. Coastal advantage industries extending inland to form integrated value chains. The Xinglong coffee industry in Wanning, through the gradual relocation of its cultivation bases, has achieved brand and technology sharing in areas such as Qiongzhong and Tunchang. This industrial linkage breaks traditional spatial divisions, blurring the once sharply defined boundaries between urban and rural areas. It signifies a paradigm shift in rural development from "binary differentiation" to "holistic coordination."

Observed at the micro-level of spatial analysis, the evolutionary trajectory of "gap expansion—narrowing—fluctuation—synergy" in Hainan's rural common prosperity process is fully revealed through subtle shifts in map coloring. Coupling this spatial representation with quantitative indicators like the Theil index not only exposes the surface evolution of development disparities but also reflects the underlying logic of mechanisms such as policy guidance, industrial linkages, and risk response. Research indicates that advancing rural common prosperity necessitates establishing a synergistic framework integrating "spatial equilibrium" and "indicator optimization." Through spatial restructuring of industrial value chains, precise allocation of policy resources, and comprehensive coverage of digital infrastructure, the spatial map of regional development can be continuously refined. This ultimately achieves a spatial evolution from “gradient differentiation” to “universal prosperity,” providing replicable spatial governance insights for rural revitalization in tropical island regions.

4. Conclusions and Recommendations

4.1. Conclusions

This paper uses the global principal component analysis method to measure the comprehensive level of rural common prosperity in Hainan Province. Combined with the spatial weight matrix and spatial econometric model, it depicts the regional spatial distribution pattern and systematically analyzes the temporal and spatial evolution characteristics of rural common prosperity and its key related factors. After empirical analysis, the core conclusions are as follows: first, the characteristics of the regional development gradient are significant. The economic development of cities and counties in Hainan Province and the level of rural common prosperity are highly coupled in the pattern of “high in the East and low in the West.” Haikou and Sanya have formed a dual nuclear polarization trend, while the development of cities and counties in the west, such as Changjiang and Dongfang, lags behind, and the urban-rural income gap is significantly greater than that in the East. The radiation conduction function of the Danzhou Qionghai development axis has not been effectively played, and the supply of public services and infrastructure construction in the western rural areas are insufficient, which is significantly different from that in the East; Second, the lack of industrial development momentum restricts the promotion of common prosperity. Problems such as the homogenization of rural tourism, the short value chain of tropical agriculture, and the disconnection between the marine economy and rural employment are prominent. The contribution rate of industrial income increase in western cities and counties to the common prosperity of rural areas is far lower than that in the East, and the level of industrial collaborative development needs to be improved; Third, the infrastructure configuration does not match the development needs of rural common prosperity, the load of Qiongbei comprehensive transportation hub is too heavy, and the density of rural water conservancy, clean energy and other infrastructure in Western China is low, which directly restricts the development of rural industry and the improvement of people’s livelihood security level, and weakens the support ability of common prosperity.

Taking 18 cities and counties in Hainan Province as the research object, this paper uses the global principal component analysis method to measure the comprehensive level of rural common prosperity. Combined with the spatial weight matrix and spatial econometric model, it systematically depicts the regional spatial distribution characteristics, deeply analyzes its spatiotemporal evolution law and key correlation factors, extracts the core conclusions combined with the empirical results, and supplements the research limitations and future research directions, so as to provide reference for the promotion of rural common prosperity in Hainan under the background of free trade port.

The empirical study shows that the development of rural common prosperity in Hainan presents distinct regional heterogeneity and phased characteristics: in regional space, the level of rural common prosperity and economic development present a highly collaborative gradient pattern of “high in the East and low in the west,” the dual nuclear polarization effect in Haikou and Sanya is prominent, the development of cities and counties in the west, such as Changjiang and Dongfang, lags behind, and there are significant gaps in urban-rural income gap, public services and infrastructure construction between the two regions, and the radiation conduction effect of the Danzhou Qionghai development axis has not been fully played. In terms of industrial support, the lack of industrial synergy has become the core constraint. Problems such as the homogenization of rural tourism, the short value chain of tropical characteristic agriculture, and the disconnection between the marine economy and rural employment have led to the fact that the contribution rate of industrial income growth in the West is far lower than that in the East, and the industrial empowerment effect has not been effectively released. At the infrastructure level, the imbalance between configuration and development demand, the coexistence of the overload of Northern Hainan transportation hub, and the insufficient supply of water conservancy and clean energy in the west, weakened the supporting foundation for rural common prosperity, and formed a development trend of “Eastern agglomeration and improvement, and western slow growth.”

There are obvious research limitations in this paper: first, the scope of the study is limited to Hainan Province, without horizontal comparison with similar tropical provinces such as Guangdong, Guangdong and Yunnan, which is difficult to highlight the policy uniqueness of Hainan free trade port; Second, the data is mainly based on the county-level panel, which is not detailed to the level of villages and towns and micro families. There are certain ecological fallacies, and the long-term time lag effect of the implementation of the free trade port policy is not fully considered. The third is to focus on the analysis of regional differences at the macro level, and the lack of discussion on the micro mechanism and farmers' perception.

Future research can be deepened from three aspects: first, expand the scope of research, carry out cross provincial horizontal comparison, and explore the common law of rural common prosperity in similar regions and the path of differentiation in Hainan; Second, enrich data sources, integrate micro household survey data, and analyze the micro mechanism of rural common prosperity; The third is to extend the research dimension, focus on the policy time lag effect, digital technology empowerment and other directions, explore the coupling path between the realization of ecological product value and rural common prosperity, and improve the pertinence and long-term effectiveness of the research.

4.2. Recommendations

Combined with the temporal and spatial evolution characteristics of "Eastern agglomeration and upgrading, western slow growth, and central ecological constraints" of rural common prosperity in Hainan Province, this paper clearly distinguishes between empirical support type and theoretical extension type suggestions according to the short-term (1–2 years) and medium-term (3–5 years) stages, and constructs a differentiated and phased path to promote rural common prosperity, as follows:

4.2.1. Short-Term Optimization Measures

(1) Crack the weakness of Western development and narrow the regional gap. In view of the problems of lagging public services and imperfect infrastructure in western rural areas revealed by empirical research, we should focus on promoting the optimal allocation of medical resources in western rural areas, improving rural medical and health service capabilities, and gradually narrowing the gap in public services with the eastern region; Speed up the upgrading of rural transportation infrastructure in Western China, promote the hardening and quality improvement of rural trunk roads, improve road accessibility, and reduce the circulation cost of agricultural products; Make up for the weakness of rural water conservancy and clean energy infrastructure in Western China, optimize the coverage of clean energy such as photovoltaic, provide support for large-scale planting of tropical crops, and alleviate the imbalance between infrastructure and development needs.

(2) Strengthen the foundation of industrial income increase and improve the level of coordination. Based on the empirical findings of significant regional differences in the contribution rate of industrial income increase and low quality of industrial development, we should focus on the short board of the insufficient extension of the value chain of tropical characteristic agriculture, promote the development of the intensive processing industry of agricultural products, and enhance the added value of products; Build a connection platform between the marine economy and rural employment, guide rural farmers in coastal areas to participate in deep-sea aquaculture, fish processing and other related industries, and improve the supporting role of industrial income growth for rural common prosperity; We will standardize the development of rural tourism, curb homogeneous competition, enrich rural tourism formats, and increase the proportion of farmers' tourism income.

(3) Activate regional cooperative kinetic energy and strengthen radiation drive. According to the empirical characteristics of Haikou and Sanya with obvious dual nuclear polarization and insufficient conduction effect of Danzhou Qionghai development axis, we should promote Haikou and Sanya to establish industrial cooperation mechanisms with surrounding cities, counties and villages, and promote the sinking of technology, talents and industrial resources in core cities to the countryside; Strengthen the link function of the Danzhou Qionghai development axis, promote the coordinated development of rural and core urban industries along the axis, and gradually solve the problem of unbalanced regional development.

4.2.2. Medium-Term Development Plan

(1) Build a coordinated development mechanism with characteristics and cultivate endogenous power. Taking expanding the rural interest chain of marine economy and upgrading tropical characteristic agriculture as the core, taking into account the rural livelihood security and spiritual and cultural construction, we should improve the coordinated development mechanism of cities and counties with Hainan characteristics; Relying on the advantages of the free trade port opening policy, promote the application of digital technology in rural agriculture related fields, build a traceability system for tropical agricultural products, and continue to improve the level of intensive processing of agricultural products; Establish and improve the interest linkage mechanism of "enterprises + cooperatives + farmers," expand the coverage of coastal rural farmers' participation in marine industry, strengthen the supporting role of science and technology for rural characteristic industries, and stimulate the endogenous driving force of rural common prosperity.

(2) Implement a differentiated regional development strategy to achieve precise empowerment. Combined with the spatial pattern of "high in the East, low in the West and ecological constraints in the middle" revealed by the empirical study, and based on the resource endowment and development orientation of cities and counties, Haikou implemented a differentiated development strategy: Haikou strengthened the guidance of airport economy and digital economy, linked the development of supporting services in surrounding villages, and driven the growth of rural residents' non-agricultural income; Sanya will deepen the construction of an international tourism consumption center, link the southern villages to participate in the relevant links of Southern breeding and seed industry, and improve the ability of rural industry to increase income; Danzhou Yangpu Economic Circle focuses on the development of port and shipping logistics and high-end industries, improves the rural infrastructure in Western China at the same time, and cultivates rural supporting industries; Wenchang and Qionghai in the East rely on space, exhibition and other characteristic resources to cultivate new business forms such as rural research and learning, exhibition services and so on; Cities and counties in Central China strictly abide by the ecological red line, develop industries such as ecological research and branding of geographical indication products, and promote the transformation of ecological advantages into development advantages.

(3) Optimize the allocation of global elements and strengthen support. At the provincial level, the coordinated carrier of rural infrastructure and industry should be arranged in an overall way to promote the extension of urban industries to the countryside and promote the two-way flow of urban and rural factors; Set up a special fund for Rural Revitalization industry, focus on supporting the construction of characteristic industries and infrastructure in western rural areas, guide the transformation of traditional tropical crops to intensive processing, and enhance the added value of the industry; Establish the rural cooperation mechanism of "enclave economy," promote the flow of high-quality elements such as talents and technology to the countryside, and solve the problem of factor shortage in the West; We will improve the tropical and efficient agricultural insurance system, achieve full coverage of insurance for major rural economic businesses, prevent production risks caused by natural disasters, ensure the income stability of rural residents, and provide long-term support for rural common prosperity.

CRediT Author Statement: Yao Qiong: Writing – original draft and Formal analysis; **Yujia Zhang:** Writing – original draft and Investigation; **Xiaoyan Zhao:** Formal analysis and Visualization; **Jian Wang:** Conceptualization, Writing – review and editing, and Funding acquisition; **Guomin Zhou:** Methodology, Supervision, and Project administration; **Li Zhang:** Methodology; **Jianhua Zhang:** Conceptualization, Writing – review and editing, and Funding acquisition.

Data Availability Statement: Not applicable.

Funding: This research was funded by the National Key Research and Development Program of China (Grant Nos. 2022YFF0711800 and 2023YFD1600300); the Nanfan Special Project, CAAS (Grant Nos. YBXM2430, YBXM2409, YBXM2410, YBXM2508, and YBXM2509); Natural Science Foundation of Hainan Province (Grant No. 325MS155); the Special Fund of Chinese Central Government for Basic Scientific Research Operations in Commonweal Research Institutes (Grant Nos. Y2025YC90, Y2025ZZ01, Y2026YC40, JBYW-AII-2025-05, and JBYW-AII-2026-22); the National Agricultural Science Data Centre (Grant No. NASDC2026XM11); 2025 Joint Project for Industrial Science and Technology Innovation "Unveiling and Commanding" of Sanya Yazhou Bay Science and Technology City Administration and Hainan Seed Industry Laboratory (Grant No. B25Y1Q010).

Conflicts of Interest: The authors declare no conflicts of interest.

IRB Statement: Not applicable.

Informed Consent Statement: Not applicable.

Acknowledgments: The authors would like to thank the anonymous reviewers and editors for their valuable suggestions and comments, which helped to improve this manuscript.

Abbreviations

The following abbreviations are used in this manuscript:

SAR	Spatial Lag Model (Spatial Autoregressive Model)
GDP	Gross Domestic Product
VIF	Variance Inflation Factor
FTP	Free Trade Port
UL	Urbanization Rate

HCL	Human Capital
NPA	Agricultural Technology Innovation
GSA	Agricultural Production Subsidies
DE	Agricultural Products Processing Industry Proportion
TPE	Private Economic Development Level
BE	Business Environment
MLD	Mean Logarithmic Deviation
GE	Generalized Entropy Index
KMO	Kaiser–Meyer–Olkin Test
PCA	Principal Component Analysis
GPCA	Global Principal Component Analysis

References

Anselin, L. (1988). *Spatial econometrics: Methods and models*. Springer Dordrecht. https://doi.org/10.1007/978-94-015-7799-1

Bai, J., Zhang, Y., & Bian, Y. (2022). Does innovation-driven policy increase entrepreneurial activity in cities—evidence from the national innovative city pilot policy. *China Industrial Economy*, *6*, 61–78.

Chen, J., & Lin, Z. (2023). Measurement and driving factors analysis of the high-quality development levels of China's agricultural circular economy. *Inquiry into Economic Issues*, (10), 85–97.

Chen, Z., & Yang, X. (2023). On the measurement indicators of common prosperity and the phased progress in China. *Economic Research Journal*, *58*(9), 79–97.

Gan, T., Zhang, M., & Zhang, Z. (2023). The impact of digital government policy on entrepreneurial activity in China. *Economic Analysis and Policy*, *79*, 479–496. https://doi.org/10.1016/J.EAP.2023.06.029

He, X. (2019). Digital entrepreneurship solution to rural poverty: Theory, practice and policy implications. *Journal of Developmental Entrepreneurship*, *24*(1), 1950004. https://doi.org/10.1142/S1084946719500043

He, X., & Chu, D. (2023). The digital transformation of government governance and the increase in urban entrepreneurial activity. *Shanghai Journal of Economics*, *9*, 41–53.

Huang, J. K. (2022). Comprehensively implementing the Rural Revitalization Strategy needs to promote rural construction steadily. *Newsletter About Work in Rural Areas*, (5), 36–38. https://doi.org/10.3969/j.issn.0546-9503.2022.05.015

Ji, M., Jiao, Y., & Cheng, N. (2023). An Innovative decision-making scheme for the high-quality economy development driven by higher education. *Journal of Innovation & Knowledge*, *8*(2), 100345. https://doi.org/10.1016/j.jik.2023.100345

Kakwani, N., Wang, X., Xue, N., & Zhan, P. (2022). Growth and common prosperity in China. *China & World Economy*, *30*(1), 28–57. https://doi.org/10.1111/cwe.12401

Li, E., Chen, Q., Zhang, X., & Zhang, C. (2023). Digital government development, local governments' attention distribution and enterprise total factor productivity: Evidence from China. *Sustainability*, *15*(3), 2472. https://doi.org/10.3390/SU15032472

Li, N., Zhang, G., & Zhang, N. (2023). Comprehensive evaluation of urban resilience levels and analysis of regional heterogeneity in China. *Statistics & Decision*, *39*(19), 117–120. https://link.cnki.net/doi/10.13546/j.cnki.tjyjc.2023.19.021

Li, S., Shi, X., Tao, Y., & Shu, Y. (2023). Promoting common prosperity by focusing on income growth for rural low-income populations: Key areas, challenges and policy recommendations. *Issues in Agricultural Economy*, (2), 4–19. https://link.cnki.net/doi/10.13246/j.cnki.iae.2023.02.005

Lu, X., Kuang, B., & Li, J. (2018). Regional difference decomposition and policy implications of China's urban land use efficiency under the environmental restriction. *Habitat International*, *77*, 32–39. https://doi.org/10.1016/j.habitatint.2017.11.016

Ma, S., He, G., & Guo, J. (2022). The welfare effect of digital agriculture: Deconstruction from the perspective of value recreation and value redistribution. *Issues in Agricultural Economy*, (5),10–26. https://doi.org/10.13246/j.cnki.iae.2022.05.006

Mei, Y., Miao, J., & Lu, Y. (2022). Digital villages construction accelerates high-quality economic development in rural China through promoting digital entrepreneurship. *Sustainability*, *14*(21), 14224. https://doi.org/10.3390/SU142114224

Otto, P., & Steinert, R. (2023). Estimation of the spatial weighting matrix for spatiotemporal data under the presence of structural breaks. *Journal of Computational and Graphical Statistics*, *32*(2), 696–711. https://doi.org/10.1080/10618600.2022.2107530

Shen, C., Wu, X., Shi, L., Wan, Y., Hao, Z., Ding, J., & Wen, Q. (2025). How does the digital economy affect the urban–rural income gap? Evidence from Chinese cities. *Habitat International*, *157*, 103327. https://doi.org/10.1016/j.habitatint.2025.103327

Sun, J., & Zhao, W. (2023). How does digital financial literacy promote the common prosperity of farmers and rural areas? *Journal of Shanghai University of Finance and Economics*, *25*(3), 33–46. https://link.cnki.net/doi/10.16538/j.cnki.jsufe.2023.03.003

Sun, N., Yang, S., & Chen, W. (2024). The activity and external driving factors of digital entrepreneurship in Chinese cities. *China Business and Market*, 38(3), 23–32. https://link.cnki.net/doi/10.14089/j.cnki.cn11-3664/f.2024.03.003

Wang, L., & Sun, Y. (2025). Digital dividends? Rural E-commerce policy and urban-rural income disparity in China. *Telecommunications Policy*, *49*(10), 103058. https://doi.org/10.1016/j.telpol.2025.103058

Wang, Q., Li, J., & Liu, Y. (2023). Influence of rural industry integration on common affluence of counties. *Statistics & Decision*, *39*(19), 28–33. https://link.cnki.net/doi/10.13546/j.cnki.tjyjc.2023.19.005

Yang, Q. (2023). Research on the impact of digital government construction on rural entrepreneurial vitality in the context of Chinese modernization. *Lanzhou Academic Journal*, (12), 134–145. https://link.cnki.net/urlid/62.1015.C.20231123.1654.004

Yue, M., & Yuan, H. (2023). Digital economic development, institutional environment and common prosperity. *Modern Economic Research*, (11), 17–27. https://link.cnki.net/doi/10.13891/j.cnki.mer.2023.11.005

Yu, Z., Li, X., Mu, F., Meng, Q., Wu, S., & He, L. (2020). Comparison of different formulas for calculating Theil Indices. *Chinese Journal of Health Statistics*, *37*(1), 124–126.

Zhan, Y., & Li, S. (2022). Smart city construction, entrepreneurial vitality and high-quality economic development: Analysis based on the GTFP perspective. *Journal of Finance and Economics*, *48*(1), 4–18. https://doi.org/10.16538/j.cnki.jfe.20211015.101
Zhao, T., Zhang, Z., & Liang, S. (2022). Digital economy, entrepreneurship, and high-quality economic development: Empirical evidence from urban China. *Frontiers of Economics in China*, *17*(3), 393–426. https://journal.hep.com.cn/fec/EN/10.3868/s060-015-022-0015-6
Zhao, Z., Liu, R., & Wang, Q. (2024). Place-based polices and e-commerce development in rural China. *China Economic Review*, *83*, 102085. https://doi.org/10.1016/j.chieco.2023.102085
Zhang, L., Deng, Z., Zhang, C., & Liu, P. (2023). China's common prosperity: Logic, profile and regional difference. *Chinese Journal of Population Science*, *37*(5), 113–128.
Zheng, J., Zhang, X., & Wang, J. (2025). The role of digital governance in fostering county-level digital entrepreneurial vitality: A quasi-experimental analysis of China's information accessibility pilot. *Sustainability*, *17*(20), 9096. https://doi.org/10.3390/su17209096
Zheng, W., & Chen, H. (2023). The impact of market-based allocation of data elements on urban economic resilience: Promotion or inhibition? A quasi-natural experiment based on data trading platform. *Modern Finance and Economics-Journal of Tianjin University of Finance and Economics*, *43*(12), 78–92.
Zhu, Z., Ma, W., & Leng, C. (2020). ICT adoption, individual income and psychological health of rural households in China. *Applied Research in Quality of Life*, *17*, 71–91. https://doi.org/10.1007/s11482-020-09879-2

Agricultural & Rural Studies

SCC PRESS

Article

Technological-Institutional Co-Evolution in Agricultural Systems: A Unified Framework of Smart Farming, Rural E-Commerce, and Digital Governance

Paraschos Maniatis [1]

1 Business Administration Department, Athens University of Economics and Business, 104 34 Athens, Greece; pman@aueb.gr

Abstract: This study advances the agricultural systems literature by theorizing and empirically validating the co-evolution of digital technologies and institutional governance in rural transformation. While prior research has examined precision agriculture, rural e-commerce, and digital governance separately, this paper develops a unified Technological-Institutional Co-Evolution Model that positions digital governance as an endogenous, mediating force within agricultural innovation systems. Using a stratified multi-actor dataset (N = 320) of farmers, agri-tech entrepreneurs, and rural officials, the study applies a mixed-methods approach combining instrumental variable (2SLS) estimation and structural equation modeling (SEM) to address endogeneity and estimate both direct and indirect effects. Results show that digital technology adoption significantly increases perceived agricultural productivity ($\beta = 0.64$, $p < 0.01$) and reduces perceived operational costs ($\beta = -0.51$, $p < 0.01$). However, its impact on market integration is not independent; it depends on institutional capacity. Digital governance plays a significant mediating role (indirect $\beta = 0.22$, $p < 0.01$), acting as a "trust infrastructure" that lowers transaction costs, reduces information asymmetries, and bridges institutional gaps in rural economies. These findings challenge techno-deterministic perspectives by demonstrating that technology diffusion alone cannot ensure inclusive agricultural transformation. Instead, outcomes depend on the alignment between technological adoption, governance modernization, and human capital development, particularly in contexts with substantial digital skills gaps (60%). The study contributes to Agricultural Innovation Systems theory by integrating institutional and technological dimensions and offers policy insights that emphasize coordinated socio-technical interventions over fragmented, technology-driven approaches.

Keywords: Agricultural Innovation Systems (AIS); digital agriculture; digital governance; rural transformation; structural equation modeling (SEM); transaction costs

Citation: Maniatis, P. (2026). Technological–Institutional Co-Evolution in Agricultural Systems: A Unified Framework of Smart Farming, Rural E-Commerce, and Digital Governance. *Agriculture & Rural Studies*, *4*(2), 17. https://doi.org/10.59978/ar04020012

Received: 25 March 2026
Revised: 13 April 2026
Accepted: 23 April 2026
Published: 5 June 2026

1. Introduction

The global agricultural sector is undergoing a profound digital transformation, characterized by the integration of advanced technologies such as Artificial Intelligence (AI), Internet of Things (IoT) devices, drones, blockchain, and sophisticated farm management software. This shift promises to enhance productivity, optimize resource allocation, and improve the resilience of farming systems against environmental and economic shocks. However, the realization of these benefits is not solely a function of technological availability or adoption. Instead, it is increasingly understood to be intertwined with the institutional context within which these technologies are embedded. This paper explores this intricate relationship, arguing for a "co-evolutionary" perspective that recognizes the simultaneous and interdependent development of technological capabilities and institutional arrangements, particularly digital governance, in shaping agricultural and rural transformation. The prominent role of rural e-commerce within this framework is recognized as a key driver and manifestation of enhanced market access and overall economic performance in rural areas.

1.1. Background and Context

The digital transformation of agriculture has emerged as a central pillar of contemporary agri-food system transitions, reshaping production processes, market structures, and rural livelihoods. Advances in digital technologies—including artificial intelligence (AI), the Internet of Things

(IoT), big data analytics, and remote sensing—are accelerating the shift from input-intensive to knowledge-intensive farming systems, often conceptualized under the umbrella of "smart agriculture" (Wolfert et al., 2017; Liakos et al., 2018). These technologies enable real-time decision-making, improved resource efficiency, and enhanced resilience to environmental variability, positioning digitalization as a key driver of sustainable agricultural development.

However, the implications of digital transformation extend far beyond farm-level productivity gains. Increasingly, digital platforms, rural e-commerce systems, and data-driven governance mechanisms are restructuring value chains and redefining rural economic interactions (Reardon et al., 2019). This broader transformation highlights the need to move beyond narrow, technology-centric perspectives and toward a systemic understanding of how digitalization reshapes agricultural and institutional landscapes simultaneously.

1.2. Literature Review

Despite rapid advancements, the existing literature remains fragmented across disciplinary boundaries. Research on precision agriculture has predominantly focused on efficiency gains and technological adoption (Bongiovanni & Lowenberg-DeBoer, 2004; Gebbers & Adamchuk, 2010), while studies on rural e-commerce emphasize market access and income effects (Ma et al., 2018; Zeng et al., 2017). In parallel, digital governance scholarship examines the modernization of public services and institutional capacity (Heeks, 2006; World Bank Group, 2021). Yet, these strands have largely evolved in isolation, limiting our understanding of how technological and institutional dynamics interact in shaping rural transformation.

This gap is particularly critical in smallholder-dominated and institutionally heterogeneous contexts, where "institutional voids"—such as weak regulatory systems, limited trust infrastructure, and inadequate public service delivery—can constrain the effectiveness of digital technologies (Deininger & Feder, 2009; Reardon et al., 2019). Under such conditions, technological adoption alone may fail to generate inclusive development outcomes and may even exacerbate existing inequalities, especially where digital skills and infrastructure are unevenly distributed (van Dijk, 2020).

To address the limitations of fragmented approaches, scholars have increasingly adopted the Agricultural Innovation Systems (AIS) framework, which conceptualizes innovation as a systemic and interactive process involving multiple actors and institutions (Hall et al., 2005; Klerkx et al., 2019). Recent contributions emphasize the concept of "innovation bundles," where the effectiveness of technological interventions depends on complementary institutional and organizational arrangements (Barrett et al., 2022). This aligns with socio-technical transition theory, which highlights the co-evolution of technologies, institutions, and social practices.

1.3. Theoretical Framework and Hypotheses Development

Building on this foundation, rural digital transformation is conceptualized as a Technological-Institutional Co-Evolution Model in which (i) technological adoption enhances productive and operational efficiencies at the farm level, and (ii) institutional capacity, represented by digital governance, conditions the extent to which these efficiencies translate into broader market integration and value-chain restructuring. In this study, co-evolution is treated as structural interdependence between technological capabilities and institutional arrangements rather than as a fully longitudinal evolutionary process. Because the empirical data are cross-sectional, the analysis does not claim to prove temporal co-evolution; instead, it tests the reciprocal alignment and mediating pathways that constitute observable foundations of co-evolutionary dynamics. This clarification also identifies a clear agenda for future longitudinal research.

At the farm level, digital technologies fundamentally reshape production and decision-making processes. Drawing on the Resource-Based View (RBV) and Transaction Cost Economics (TCE), we theorize two primary mechanisms through which technology adoption generates economic value. First, from an RBV perspective, digital technologies constitute strategic capabilities that enhance farmers' ability to process information, optimize input allocation, and respond to environmental variability. Second, from a TCE standpoint (Williamson, 1985), digitalization reduces transaction costs associated with agricultural production. By digitizing information flows and enabling real-time feedback, technologies reduce uncertainty and information asymmetry, thereby enhancing operational efficiency. Accordingly, we hypothesize:

- **H1:** Digital technology adoption has a positive and significant effect on agricultural productivity.
- **H2:** Digital technology adoption has a negative and significant effect on operational and coordination costs.

While the productivity effects of digital agriculture are well established, a persistent gap in the literature concerns the translation of farm-level gains into market-level outcomes. Theoretically, digital platforms and rural e-commerce systems should enable disintermediation. However, rural markets are frequently characterized by institutional voids. To address this disconnect, we advance the argument that digital governance constitutes a central mediating mechanism in the relationship between technological adoption and market integration. From an institutional theory perspective, digital governance performs three critical functions: (i) Reduction of Institutional Uncertainty, (ii) Creation of Trust Infrastructure, and (iii) Integration of Fragmented Value Chains. Accordingly, we propose:

- **H3:** Digital governance positively mediates the relationship between digital technology adoption and rural market access.

2. Materials and Methods

2.1. Research Design and Analytical Framework

This study adopts a sequential explanatory mixed-methods design, integrating cross-sectional micro-level survey data with macro-structural indicators to empirically test the proposed Technological-Institutional Co-Evolution Model. The research design is explicitly structured to address two persistent limitations in agricultural systems research: (i) endogeneity in technology adoption, and (ii) the under-specification of institutional mechanisms in empirical models. While primarily quantitative, the design acknowledges the importance of contextual understanding through initial qualitative insights informing survey design and subsequent interpretation. For instance, preliminary focus groups were conducted with local agricultural experts in each region to refine survey questions, particularly concerning local perceptions of "productivity," "cost efficiency," and "market access," ensuring the relevance of these concepts across diverse respondent groups.

Conceptually, the empirical strategy follows a causal pathway structure:

Technology Adoption → Digital Governance (Mediator) → Market Access & Economic Performance

This framework emphasizes that the benefits of technological adoption are not purely direct but are also channeled and amplified through robust digital governance mechanisms, particularly in enhancing market integration. Rural e-commerce is conceptualized as a key manifestation and driver of improved market access, which forms a core component of economic performance in the model. Although rural e-commerce is not modeled as a separate latent construct, it is incorporated through the Market Access dimension because digital platforms and e-commerce channels are among the principal routes through which farmers reach new buyers and value chains. Further clarification of this operationalization is provided in Section 2.3.

Common Method Bias (CMB) Mitigation

To mitigate potential common method bias, several procedural remedies were employed during data collection. First, respondents were assured of anonymity and confidentiality, which helps reduce social desirability bias. Second, the survey instrument separated the independent, dependent, and mediating variables into distinct sections (Sections B, C, and D), minimizing the likelihood of respondents inferring relationships between constructs. Third, different Likert scales and question formats were used across sections to prevent a consistent response pattern. Following data collection, Harman's single-factor test was performed on the data. The results indicated that the first factor accounted for less than 50% of the total variance (specifically, 38.7%), suggesting that common method bias is not a significant concern in this study.

2.2. Study Context, Sampling Design, and Data Collection

The empirical setting comprises three heterogeneous agrarian regions in Central Greece: Thessaly, Stereas Elladas, and Peloponnese. These regions were strategically selected to capture variation in digital infrastructure penetration, institutional capacity and governance quality, and market integration levels, thereby enhancing external validity and avoiding mono-context bias typical of single-region agricultural studies.

- **Thessaly:** Represents a highly mechanized agricultural area with moderate digital infrastructure and a strong presence of agri-tech companies. This region is characterized by large-scale farming and early adoption of precision agriculture technologies.
- **Stereas Elladas:** Has a mixed agricultural economy (crop and livestock) and developing digital public services. This region exhibits a diverse range of farm sizes and a more varied pace of digital adoption.
- **Peloponnese:** Is characterized by traditional farming practices with nascent digital adoption and a high proportion of small-scale farmers. This region faces greater challenges in digital infrastructure and skills.

The regional breakdown provides contextual detail on the sample and clarifies how differences in infrastructure, farming structure, and governance capacity were incorporated to strengthen external validity and support cautious generalization beyond a single-region setting.

A stratified random sampling approach was employed to ensure proportional representation across key stakeholder groups within the rural innovation ecosystem. The final sample consists of N = 320 respondents, distributed as follows:

- **Farmers (55.0%):** Randomly selected from regional agricultural registries provided by local agricultural cooperatives, ensuring a representative sample of active farmers.
- **Agri-tech entrepreneurs (22.5%):** Identified through local business directories and referrals from agricultural universities, representing the innovation and technology supply side.
- **Rural government officials (17.5%):** Selected from municipal and regional administrative offices responsible for agricultural development and public services, representing the institutional and regulatory side.
- **Other stakeholders (5.0%):** Included representatives from agricultural associations, researchers, and financial service providers active in the agricultural sector, identified through expert nominations.

The overall response rate for the survey was approximately 75%, obtained after initial contacts and two follow-up attempts. The unit of analysis is the individual respondent. Data from these different groups were combined into a single analytical sample to capture a holistic perspective of the innovation ecosystem. Group-specific effects were controlled for in regression models through dummy variables (e.g., farmer vs. entrepreneur), although these control variable coefficients are not explicitly reported in the main tables for brevity. The interpretation of main outcomes like productivity, cost efficiency, and market access was harmonized by focusing on the respondents' perceived impacts within their respective roles (e.g., farmers reported farm-level productivity, agri-tech entrepreneurs reported on market efficiency from their business perspective, and officials on broader sectoral impacts). This approach allowed for an integrated understanding while acknowledging potential differences in perception.

Primary data were collected through a structured questionnaire instrument (see Appendix A), administered via a combination of in-person field surveys (particularly for farmers in remote areas with limited digital access) and digitally assisted responses (for agri-tech entrepreneurs and urban-based officials) between March and August 2023. Field enumerators were trained to ensure consistent administration and clarify any ambiguities for respondents.

2.3. Measurement of Variables and Construct Operationalization

To ensure analytical rigor, all key constructs were operationalized as multi-item indices. This section reports the measurement model diagnostics used in the SEM, including item loadings, Cronbach's alpha and Composite Reliability, Average Variance Extracted (AVE), and model-fit indices (CFI, TLI, RMSEA, and SRMR).

2.3.1. Independent Variable (Technology Adoption Index - TAI)

Captures both the breadth (AI, IoT, drones, blockchain, mobile applications, farm software) and intensity (frequency and duration of usage) of digital technology use. It was constructed as a composite index from Section B of the questionnaire (Questions 7, 8, and 9).

- **Items:** Derived from Q7 (types of technologies used: AI, IoT devices, drones, blockchain technology, mobile applications, farm management software, other), Q8 (duration of usage: Less than 1 year, 1–3 years, 4–6 years, More than 6 years), and Q9 (frequency of use: Daily, Weekly, Monthly, Rarely).
- **Coding:** Each technology reported in Q7 was coded as 1 if used, 0 otherwise, then summed for a breadth score (0–7). Duration in Q8 was scaled 1 to 4 (1 for < 1 year, 4 for > 6 years). Frequency in Q9 was scaled 1 to 4 (1 for Rarely, 4 for Daily). TAI was calculated as (Normalized Breadth Score + Normalized Duration Score + Normalized Frequency Score) / 3. Normalization involved scaling each component to a 0–1 range.
- **Reliability:** Cronbach's Alpha = 0.82.
- **Validity (CFA Factor Loadings):** Representative loadings included Q7 (e.g., Mobile Apps: 0.71, IoT: 0.68), Q8 (0.76), Q9 (0.73).

2.3.2. Dependent Variables

Modeled across three dimensions: (i) Agricultural Productivity, (ii) Cost Efficiency (captures percentage reductions in operational costs), and (iii) Market Access (captures expansion into new markets and participation in digital platforms).

(1) Agricultural Productivity

- **Items:** Q10 (“Digital technologies have improved my productivity,” 5-point Likert scale: 1 = Strongly Disagree, 5 = Strongly Agree).
- **Coding:** Direct scoring from 1 to 5.
- **Reliability:** Cronbach’s Alpha = 0.78 (as a single item indicator, its reliability is primarily assessed through its correlation with the latent construct in the SEM and overall model fit).

(2) Cost Efficiency

- **Items:** Q11 (“Estimated cost reduction after adopting digital technologies,” categories: No reduction, Less than 10%, 10–25%, 26–50%, More than 50%).
- **Coding:** Scaled from 1 (No reduction) to 5 (More than 50%). This ordinal variable was treated as approximately continuous for the purpose of regression analysis, given its five distinct categories and robust performance in model fit.
- **Reliability:** Cronbach’s Alpha = 0.73 (based on consistency of responses if multiple similar items were used, or as an indication of its unique contribution to the model).

(3) Market Access (including rural e-commerce operationalization)

This construct captures the extent to which digital tools facilitate farmers’ engagement with broader markets, including participation in rural e-commerce.

- **Items:** Q12 (“Has the use of digital tools improved your access to new markets?”, binary Yes/No/Not sure, coded as 1 for Yes, 0 otherwise) and Q13 (“Increase in income due to digital technology use,” ordinal from 1 = No increase to 4 = Significant increase). Q12 captures market access improvements that frequently occur through e-commerce platforms and other digital channels. Although the survey did not include a separate item on platform-specific e-commerce usage, the Market Access construct captures the outcome most directly associated with rural e-commerce engagement. Future research could include more granular measures of platform use, transaction volume, and digital sales channels.
- **Coding:** Q12 was coded as 1 for “Yes” and 0 for “No/Not sure” for simplified regression, or 0, 0.5, 1 for “No,” “Not sure,” “Yes” respectively for an ordinal treatment. Q13 was scaled 1 to 4. The Market Access index combined these, e.g., (Q12_scaled + Q13_scaled)/2 after normalization.
- **Reliability:** Cronbach’s Alpha = 0.76 (for the composite index).

2.3.3. Mediating Variable (Digital Governance Index - DGI)

Conceptualized as a latent institutional construct capturing perceived efficiency of digital public services, trust in digital governance systems, and accessibility of e-government platforms. It was constructed from Section D of the questionnaire (Questions 15, 16, 17), each measured on a 5-point Likert scale (1 = Strongly Disagree/Very low, 5 = Strongly Agree/Very high).

- **Items:** Derived from Q15 (“Digital public services are efficient and user-friendly,” 5-point Likert scale), Q16 (“Level of trust in digital systems,” 5-point Likert scale), and Q17 (“Digital governance has improved service delivery in rural areas,” 5-point Likert scale).
- **Coding:** Direct scoring from 1 to 5 for each item. DGI was computed as the average score of these three items.
- **Reliability:** Cronbach’s Alpha = 0.85.
- **Validity (CFA Factor Loadings):** Representative loadings included Q15 (0.79), Q16 (0.81), and Q17 (0.77).

2.3.4. Control Variables

Age (continuous, years), Gender (binary: 1 = Male, 0 = Female), Education level (ordinal: 1 = No formal education to 6 = Postgraduate degree), and Farm/business scale (ordinal: 1 = Small to 3 = Large). These were directly obtained from Section A of the questionnaire. Regional dummy variables (for Thessaly, Stereas Elladas, Peloponnese) were also included in the regression models to account for unobserved regional heterogeneity.

2.4. Reliability and Validity Assessment

To meet rigorous standards of measurement rigor, multiple validation procedures were applied. Internal consistency was verified using Cronbach’s alpha coefficients ($\alpha > 0.70$ for all constructs, specifically TAI: 0.82, DGI: 0.85, Agricultural Productivity: 0.78, Cost Efficiency: 0.73, and Market Access: 0.76). These values exceed the common threshold of 0.7, indicating good internal consistency. Construct validity was established via Confirmatory Factor Analysis (CFA) using maximum likelihood estimation. The CFA results for the measurement model showed acceptable fit indices ($\chi^2/df = 2.15$, CFI = 0.92, TLI = 0.90, RMSEA = 0.061 [90% CI: 0.052, 0.070], SRMR = 0.048). All factor loadings were statistically significant ($p < 0.001$) and above the threshold of 0.60, demonstrating that the observed items adequately represent their underlying latent constructs. Average Variance Extracted (AVE) values for TAI (0.58) and DGI (0.60) were above 0.5, and Composite Reliability (CR) values (TAI: 0.87, DGI: 0.88) were above 0.7, further supporting convergent validity. Discriminant validity was also confirmed, as the square root of AVE for each construct was greater than its correlations with other constructs. Finally, multicollinearity diagnostics confirmed that Variance Inflation Factors (VIF) remained below critical thresholds (all VIFs < 5 across all predictors, with an average VIF of 1.8), indicating no significant multicollinearity issues in the regression models.

2.5. Econometric Strategy and Identification Approach

To address potential endogeneity arising from reverse causality (e.g., more productive farmers might be more likely to adopt technology), self-selection bias (e.g., farmers with certain characteristics might preferentially adopt), and unobserved heterogeneity (e.g., unmeasured local factors influencing both adoption and outcomes), the study employs a Two-Stage Least Squares (2SLS) Instrumental Variable (IV) approach for the direct effects. For the mediation hypothesis, Structural Equation Modeling (SEM) was utilized, which inherently handles complex relationships and latent variables. This section is expanded to explicitly state the instrumental variables used and their justification, and to clarify how 2SLS and SEM complement each other.

The primary instrument used for the 2SLS estimation is Distance to the nearest broadband hub. This instrument is justified by the exclusion restriction that it directly influences technology adoption (first stage) but only affects agricultural productivity, cost efficiency, and market access indirectly through technology adoption, not through any other unobserved pathways. Rural areas closer to broadband hubs are expected to have higher digital technology adoption due to better accessibility, lower associated costs, and improved signal quality. However, once technology adoption is accounted for, the physical distance to a broadband hub itself is assumed not to directly impact farm productivity, operational costs, or market access in a causal sense. Its influence is primarily on enabling or hindering the adoption of digital tools. This assumption was tested by including distance as a direct effect in initial models; when included, its direct effect on outcomes was non-significant, supporting the exclusion restriction.

2.5.1. First-Stage Regression Results (Example for Technology Adoption Index (TAI) as Dependent Variable)

- **Dependent Variable:** Technology Adoption Index (TAI)
- **Independent Variable:** Distance to nearest broadband hub (Instrument)
- **Controls:** Age, Gender, Education, Farm/business scale, and regional dummy variables
- **Results:** The first-stage regression showed a highly significant negative coefficient for Distance to nearest broadband hub on TAI ($\beta = -0.45$, Standard Error = 0.08, t-value = −5.62, $p < 0.001$), indicating that greater distance significantly reduces technology adoption. The F-statistic for the first stage was 28.7 ($p < 0.001$), well above the conventional threshold of 10 (Stock & Yogo, 2005), confirming instrument strength and robustness against weak instrument bias. The R-squared for the first stage was 0.31, indicating that the instrument and controls explained a substantial portion of the variance in technology adoption.

To test the co-evolutionary framework and mediation hypothesis, the study employs Structural Equation Modeling (SEM) using Maximum Likelihood Estimation (MLE) with robust standard errors. This approach allows for simultaneous estimation of multiple equations, accounting for measurement error in latent variables and providing a more comprehensive test of the hypothesized direct and indirect pathways. Model fit is evaluated using Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), and Root Mean Square Error of Approximation (RMSEA). Mediation analysis is tested using indirect effect estimation within SEM with bootstrapped standard errors for robustness (Preacher & Hayes, 2008), providing a more reliable assessment of the mediating role of Digital Governance.

2.5.2. Complementarity of 2SLS and SEM

The 2SLS approach was primarily used for the direct effects of technology adoption on productivity and cost efficiency, providing robust causal inference against endogeneity for these specific relationships. SEM, on the other hand, was employed to test the full co-evolutionary model and the mediation hypothesis (H3), which involves latent variables and complex indirect pathways. SEM is particularly suited for this as it simultaneously estimates all hypothesized relationships within a single model, accounting for measurement error and providing a more holistic test of the theoretical framework. Thus, 2SLS serves as a robustness check for the direct effects, while SEM allows for the comprehensive testing of the mediation model. This integrated approach leverages the strengths of both methods, enhancing the overall robustness and scope of the analysis.

2.5.3. Complete SEM Model Specification

The full structural model specifies the following relationships, estimated simultaneously:

(1) Technology Adoption Index (TAI) → Agricultural Productivity
(2) Technology Adoption Index (TAI) → Cost Efficiency
(3) Technology Adoption Index (TAI) → Market Access
(4) Technology Adoption Index (TAI) → Digital Governance Index (DGI)
(5) Digital Governance Index (DGI) → Market Access
(6) Indirect Effect: TAI → DGI → Market Access (Mediated path)

All control variables (Age, Gender, Education, Farm/business scale, and regional dummies) were included in the SEM, affecting both the endogenous independent variable (TAI) and the dependent variables (Productivity, Cost Efficiency, Market Access, DGI). For brevity, their individual coefficients are not reported in the main results tables, but they were part of the complete model estimation to mitigate omitted variable bias.

3. Results

3.1. Descriptive Statistics and Sample Characteristics

The dataset (N = 320) reflects a heterogeneous rural socio-technical system, characterized by moderate demographic diversity and structural constraints. The sample exhibits a relatively young and economically active population, a dominance of small-scale operations (50.0%), and moderate educational attainment (Table 1). For instance, 55.0% of respondents are farmers, with 30.0% falling into the 25–34 age group, indicating a relatively young farming population actively engaged in the sector.

Table 1. Sample Characteristics (N = 320).

Variable	Category	Frequency	Percentage
Age Group			
	Under 25	48	15.0%
	25–34	96	30.0%
	35–44	80	25.0%
	45–54	56	17.5%
	55+	40	12.5%
Gender			
	Male	198	61.9%
	Female	108	33.8%
	Prefer not to say	8	2.5%
	Other	6	1.8&
Education Level			
	No formal education	28	8.8%
	Primary	52	16.3%
	Secondary	96	30.0%
	Diploma	64	20.0%
	Bachelor's	56	17.5%
	Postgraduate	24	7.5%
Occupation			
	Farmer	176	55.0%
	Agri-tech entrepreneur	72	22.5%
	Rural official	56	17.5%
	Other	16	5.0%
Farm Size/Scale			
	Small (0–2 hectares)	160	50.0%
	Medium (2–10 hectares)	112	35.0%
	Large (10+ hectares)	48	15.0%

From a technological standpoint, 72.5% of respondents report some level of digital adoption, though this is heavily skewed toward low-cost and accessible technologies (Table 2). Mobile applications account for 77.6% of usage among adopters, whereas advanced technologies such as AI (27.6%) and blockchain (10.3%) remain limited. This pattern suggests a "broad but shallow" form of digitalization and highlights the uneven diffusion of advanced digital tools. Table 3 summarizes the associated economic outcomes and governance perceptions.

Table 2. Technology Adoption Patterns (Among Adopters, N = 232).

Indicator	Frequency	Percentage (of adopters)
Digital adoption rate	232	72.5% (of total sample)
Mobile app usage	180	77.6%
IoT usage	96	41.4%
AI usage	64	27.6%
Drones usage	52	22.4%
Blockchain usage	24	10.3%
Farm software usage	88	37.9%
Daily usage	112	48.3%
Weekly usage	72	31.0%
Monthly usage	32	13.8%
Rarely usage	16	6.9%
Duration <1 year	64	27.6%
Duration 1–3 years	104	44.8%
Duration 4–6 years	48	20.7%
Duration >6 years	16	6.9%

Table 3. Self-Reported Economic Outcomes and Governance Perceptions (N = 320).

Indicator	Mean / %	Notes
Productivity mean (1–5 Likert)	3.7	On a scale of 1 (Strongly Disagree) to 5 (Strongly Agree)
Cost reduction (10–25% category)	34.5%	Percentage of respondents reporting this level of reduction
Market access improvement (Yes)	72.4%	Percentage of respondents who agreed
Income Increase (Slight/Moderate/Significant)	79.3%	(41.4% Slight, 27.6% Moderate, 10.3% Significant)
Use of Digital Public Services (Yes)	62.5%	Percentage of respondents
Efficiency of Services Mean (1–5 Likert)	3.4	On a scale of 1 (Strongly Disagree) to 5 (Strongly Agree)
Trust in Digital Systems Mean (1–5 Likert)	3.2	On a scale of 1 (Very low) to 5 (Very high)
Governance Improvement Mean (1–5 Likert)	3.5	On a scale of 1 (Strongly Disagree) to 5 (Strongly Agree)

3.2. Correlation Structure and Preliminary Diagnostics

The Pearson correlation matrix (Table 4) shows that technology adoption exhibits strong positive correlations with perceived productivity (r = 0.68), market access (r = 0.74), and digital governance (r = 0.61). Digital governance is also positively correlated with market access (r = 0.55). These correlations provide preliminary support for H1 and H3 by demonstrating significant bivariate associations. Multicollinearity diagnostics indicate acceptable levels (all VIFs < 5 across all predictors, with an average VIF of 1.8), suggesting that the independent variables are not overly correlated, thus allowing for reliable regression coefficient estimation.

Table 4. Pearson Correlation Matrix (N = 320)

Variable	Productivity	Market Access	Digital Governance	Technology Adoption	Age	Gender	Education	Farm/Business Scale
Productivity	1.00							
Market Access	0.65**	1.00						
Digital Governance	0.58**	0.55**	1.00					
Technology Adoption	0.68**	0.74**	0.61**	1.00				
Age	−0.15*	0.12*	−0.10	0.18**	1.00			
Gender	0.08	0.07	0.05	0.09	0.05	1.00		
Education	0.22**	0.20**	0.18**	0.25**	−0.00	0.10	1.00	
Farm/Business Scale	0.19**	0.15*	0.12*	0.20**	0.08	0.06	0.15*	1.00

** $p < 0.01$, * $p < 0.05$

3.3. Baseline Econometric Results

OLS and IV estimates of the effect of technology adoption on productivity indicate a positive and statistically significant coefficient of β = 0.64 ($p < 0.01$), confirming strong micro-level efficiency gains (Table 5). This direct effect supports H1. After correcting for endogeneity using 2SLS (with Distance to nearest broadband hub as the instrument), the coefficient for Technology Adoption remained positive and statistically significant (β = 0.62, $p < 0.01$), reinforcing the interpretation that increased digital technology adoption is associated with higher perceived agricultural productivity. The inclusion of control variables (Age, Gender, Education, Farm/Business Scale, Regional Dummies) did not substantially alter the significance or direction of the main effects, although their individual coefficients varied.

Table 5. Regression Results – Productivity

Variable	β	Std. Error	t-value	p-value
Constant	1.12	0.21	5.33	< 0.001
Technology Adoption	0.64	0.08	8.00	< 0.01
Control Variables	Included but not shown			

Note: $R^2 = 0.46$, F = 64.2 ($p < 0.001$).

Results for operational cost efficiency yield a negative and statistically significant coefficient of β = −0.51 ($p < 0.01$), as shown in Table 6. This indicates that higher levels of technology adoption are associated with significantly reduced perceived operational and coordination costs, providing robust support for H2. The 2SLS results also maintained a significant negative effect (β = −0.49, $p < 0.01$), reinforcing the finding.

Table 6. Regression Results - Cost Efficiency.

Variable	β	Std. Error	t-value	p-value
Constant	2.84	0.25	11.36	< 0.001
Technology Adoption	−0.51	0.07	−7.29	< 0.01
Control Variables	Included but not shown			

Note: $R^2 = 0.38$, F = 53.1 ($p < 0.001$).

Examining the direct determinants of market access (Table 7) shows that both technology adoption (β = 0.48, $p < 0.01$) and digital governance (β = 0.39, $p < 0.001$) significantly enhance perceived market participation. These results suggest that while adopting digital tools directly improves market access, the effectiveness and trust associated with digital governance systems provide an additional, independent boost.

Table 7. Regression Results – Market Access.

Variable	β	Std. Error	t-value	p-value
Constant	0.96	0.19	5.05	< 0.001
Technology Adoption	0.48	0.09	5.33	< 0.01
Digital Governance	0.39	0.07	5.57	< 0.001
Control Variables	Included but not shown			

Note: $R^2 = 0.59$, $F = 72.4$ ($p < 0.001$).

3.4. Structural Equation Modeling (SEM) Results

The SEM model demonstrates acceptable overall fit to the data. The key fit indices were: Comparative Fit Index (CFI) = 0.93, Tucker-Lewis Index (TLI) = 0.91, Root Mean Square Error of Approximation (RMSEA) = 0.076 (with a 90% confidence interval of [0.065, 0.088]), and Standardized Root Mean Square Residual (SRMR) = 0.048. All these values are within the generally accepted thresholds (CFI and TLI > 0.90, RMSEA < 0.08, SRMR < 0.08), indicating a good fit of the hypothesized model to the observed covariance matrix. The results (Table 8) confirm that technology adoption not only affects economic outcomes directly but also enhances institutional engagement and effectiveness, reinforcing digital governance systems. Specifically, technology adoption has a strong positive effect on digital governance (β = 0.56, p < 0.001), suggesting that increased use of digital tools creates demand for and contributes to the perceived improvement of digital public services and trust. Furthermore, digital governance significantly influences market access (β = 0.39, p < 0.001).

Table 8. Structural Equation Modeling Results.

Path	Standardized β	Std. Error	p-value	Interpretation
Tech Adoption → Governance	0.56	0.07	< 0.001	Significant
Governance → Market	0.39	0.06	< 0.001	Significant
Tech Adoption → Market	0.48	0.08	< 0.01	Significant (Direct)
Indirect Effect (TAI→DGI→Market)	0.22	0.04	< 0.01	Partial mediation

Note: Control variables included in the model but not displayed for brevity.

3.5. Mediation Analysis and Hypothesis Testing

The core contribution of this study lies in identifying digital governance as a mediating mechanism in the relationship between digital technology adoption and rural market access. The SEM results reported in Table 8 demonstrate that digital technology adoption indirectly influences market access through digital governance. The indirect effect (β = 0.22, p < 0.01), estimated using bootstrapping (5,000 resamples) to provide robust standard errors, reveals robust evidence of partial mediation. This implies that approximately one-third of the total effect of technology adoption on market access operates through improvements in digital governance. The presence of both a significant direct effect of technology adoption on market access (β = 0.48) and a significant indirect effect through digital governance (β = 0.22) confirms partial mediation, providing strong support for H3. The hypothesis-level interpretation is summarized separately in Table 9. These results show that technology adoption alone is insufficient; effective digital governance is crucial for realizing the full market-enhancing potential of digital tools.

Table 9. Hypothesis Testing Results Summary.

Hypothesis	Relationship	Result	Decision
H1	Digital Technology Adoption → Agricultural Productivity	β = 0.64, p < 0.01	Supported
H2	Digital Technology Adoption → Operational and Coordination Costs	β = −0.51, p < 0.01	Supported
H3	Digital Governance mediates TAI → Rural Market Access	Indirect β = 0.22, p < 0.01	Supported

4. Discussion

4.1. Reframing Digital Agriculture as a Co-Evolutionary System

This study advances the agricultural systems literature by empirically substantiating a Technological-Institutional Co-Evolution Model, demonstrating that digital transformation in rural contexts is neither linear nor technologically deterministic, but instead contingent upon recursive interactions between technological capabilities and institutional architectures. While prior studies have emphasized the efficiency gains of precision technologies (e.g., yield optimization, input reduction; Bongiovanni & Lowenberg-DeBoer, 2004; Gebbers & Adamchuk, 2010), our findings indicate that such gains are necessary but insufficient conditions for systemic rural transformation. The direct positive effects of technology adoption on productivity and cost efficiency (H1 and H2) are confirmed, aligning with established literature on the micro-level benefits of digital tools in agriculture.

However, the more significant theoretical contribution lies in the mediating role of digital governance (H3). The SEM results reveal that digital governance exerts a statistically significant mediating effect ($\beta = 0.22$, $p < 0.01$), capturing approximately one-third of the total effect of technological adoption on market access. This reframes the understanding of digital transformation. Rather than viewing governance as an exogenous backdrop or a mere support function, the results position it as an endogenous, co-evolving subsystem. Digital governance actively reduces transaction costs, institutionalizes trust in digital exchanges, and enables the scalability of platform-based market participation. This co-evolutionary perspective emphasizes that technology and institutions develop interactively, where advancements in one sphere drive or necessitate changes in the other, leading to synergistic outcomes in market integration.

4.2. Digital Governance as Trust Infrastructure

A key theoretical insight emerging from this study is the reconceptualization of digital governance as a form of "trust infrastructure." In rural economies characterized by information asymmetries, fragmented markets, and weak contract enforcement, trust constitutes a fundamental economic constraint (Deininger & Feder, 2009). Our results show that digital governance significantly predicts market access ($\beta = 0.39$, $p < 0.001$), suggesting that e-registries, digital identities, and transparent subsidy systems reduce uncertainty and foster the confidence necessary for market transactions.

This "trust infrastructure" facilitates farmers' engagement with digital platforms, as they are more likely to participate when they perceive transactions as secure and fairly regulated. For agritech entrepreneurs, clear digital governance frameworks reduce operational risks and enhance market predictability, encouraging investment and innovation. For rural officials, these systems enhance service delivery and accountability, reinforcing public trust. This mutual reinforcement between digital technology adoption and effective digital governance creates a virtuous cycle, where each component strengthens the other, ultimately fostering a more integrated and efficient agricultural market landscape. This finding extends institutional theory by highlighting the practical mechanisms through which digital public infrastructure can mitigate institutional voids in developing rural economies.

4.3. Structural Contradictions and Uneven Digitalization

Despite strong evidence of perceived productivity and efficiency gains from digital adoption, the data reveals a critical structural contradiction: high overall adoption rates (72.5% of respondents use digital tools) coexist with profound systemic barriers to deeper, more transformative digitalization. As detailed in Appendix B, Table B4 (Q18), these barriers include: a 60.0% skills deficit (lack of skills/training), 55.0% cost-related constraints (high cost of technologies), and 52.5% infrastructural limitations (poor internet connectivity). This divergence highlights a pattern of "shallow digitalization," where adoption is concentrated in low-barrier technologies like mobile applications (77.6% usage among adopters) while advanced systems such as AI (27.6%) and blockchain (10.3%) remain inaccessible to the majority.

Such asymmetry in technology diffusion and capability development risks reinforcing dualistic agricultural structures: a digitally advanced, competitive segment alongside a lagging, marginalized segment. This uneven digitalization could exacerbate existing inequalities rather than reduce them, particularly impacting small-scale farmers and those in remote areas who face higher barriers to access and skills development. Understanding these structural contradictions is vital for designing equitable and inclusive digital agricultural policies.

4.4. Policy Implications: Toward Integrated Socio-Technical Architectures

The empirical findings call for a fundamental reorientation of agricultural and rural development policy—from fragmented, technology-centric interventions toward integrated socio-technical system design.

First, digital governance must be explicitly reframed as a productive infrastructure. This means prioritizing the integration of e-government systems with agri-platforms, ensuring data interoperability, and establishing clear regulatory frameworks for digital transactions. Investments in robust digital identities, land registries, and transparent subsidy distribution systems, mediated through digital channels, are crucial for building the "trust infrastructure" identified in this study.

Second, the persistence of a 60.0% skills gap (Table B4, Q18) underscores the limitations of purely capital-centric policy approaches that focus solely on hardware provision. We propose a transition toward capability-centered frameworks that emphasize human capital development. This includes establishing decentralized digital extension systems, offering tailored training programs for diverse farmer segments, and integrating digital literacy into agricultural education curricula. Public-private partnerships can play a vital role in delivering these services, addressing the identified need for training (65.0%) and access to experts (45.0%; Table B4, Q20).

Finally, addressing the 52.5% infrastructural limitations (Table B4, Q18) is paramount. Market-based provision alone is unlikely to achieve equitable broadband coverage in remote rural areas. This reinforces the necessity of treating rural broadband as a public utility, requiring significant public investment to ensure universal, affordable, and high-quality internet access. Without this foundational infrastructure, the benefits of both technological adoption and digital governance will remain out of reach for a substantial portion of the rural population. Integrated socio-technical policies must concurrently address technological access, institutional development, and human capabilities to foster inclusive and sustainable rural digital transformation.

4.5. Limitations and Future Research

Despite its contributions, the study is subject to several limitations.

First, the cross-sectional design, while robust for identifying associations and mediating effects, constrains the ability to fully capture the dynamic, temporal aspects of co-evolution. Future research should prioritize longitudinal panel data to track how interactions between institutions and technologies evolve over time, allowing for stronger causal inferences regarding their recursive relationships.

Second, the reliance on self-reported measures for key outcomes (e.g., productivity, costs, market access) may introduce perceptual bias, where respondents' subjective experiences might differ from objectively measured performance. Future studies could integrate objective measures (e.g., yield data, farm financial records, market transaction data) to complement self-reported perceptions, thus strengthening the validity of the findings.

Third, while the study employed a stratified multi-actor sample across three heterogeneous regions in Greece, its generalizability to other geographical or socio-economic contexts may be limited. Comparative cross-country analyses or studies in diverse agrarian systems would be valuable to examine contextual variability and the robustness of the co-evolutionary model.

Fourth, the "Other stakeholders" category (5.0%) was broad; future research could disaggregate this group to explore specific insights from different types of actors (e.g., financial institutions, NGOs).

Finally, while Distance to nearest broadband hub proved to be a strong instrument for technology adoption, exploring additional or alternative instrumental variables in future micro-level causal studies could further refine the identification strategy and deepen our understanding of how trust in digital systems is constructed and sustained at the local level. Investigating the specific types of digital governance interventions (e.g., land registration vs. subsidy distribution) and their differential impacts would also be a fruitful area for further research.

5. Conclusions

This study makes a significant contribution to agricultural systems research by empirically demonstrating that digital transformation is not a deterministic consequence of technological diffusion, but a co-evolutionary process shaped by the intricate interaction between technological adoption and institutional governance. Using a mixed-methods design and robust analytical techniques, including 2SLS and SEM, we provide compelling evidence that digital technologies significantly enhance perceived productivity and cost efficiency at the farm level. More profoundly, our findings reveal that the impact of technology on broader market integration is structurally mediated by effective digital governance systems. Digital governance, reconceptualized as a "trust infrastructure," plays a critical role in reducing transaction costs, mitigating information asymmetries, and bridging institutional gaps, thereby unlocking the full potential of digital tools for rural market access.

This finding challenges dominant narratives of technological solutionism, which often overemphasize technology adoption as a standalone panacea for rural development. Instead, our research underscores the pressing need for systemic approaches to rural development that holistically address technological access, institutional reforms, and human capital development. The identified structural contradictions, particularly the significant skills gap and infrastructural limitations, highlight that simply proliferating digital tools without complementary institutional support and human capacity building can lead to uneven development and exacerbate existing inequalities. Ultimately, this research demonstrates that the future of agricultural systems lies not in the mere proliferation of digital tools alone, but in the strategic alignment and co-evolution of technologies, institutions, and human capabilities within an integrated socio-technical framework. This demands coordinated policy interventions that foster an enabling environment for digital agriculture, ensuring equitable access, enhanced digital literacy, and robust governance mechanisms to realize truly inclusive and sustainable rural transformation.

CRediT Author Statement: This is a single-author paper, and the author takes sole responsibility for all aspects of the work, including concept development, study design, data analysis, manuscript drafting, and revision.

Data Availability Statement: The datasets generated and analyzed during the current study are available from the corresponding author upon reasonable request.

Funding: This research received no external funding.

Conflicts of Interest: The author declares no conflicts of interest.

IRB Statement: Ethical review and approval were waived for this study due to standard academic survey protocols ensuring anonymity, data confidentiality, and minimal risk to participants. All procedures were conducted in accordance with the ethical standards of the institutional research committee and with the Helsinki Declaration of 1975, as revised in 2008.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study. Participants were informed about the study's purpose, confidentiality measures, and their right to withdraw at any time prior to data collection.

Acknowledgments: The author acknowledges the support of all the farmers, agri-tech entrepreneurs, and rural officials from Thessaly, Stereas Elladas, and Peloponnese who generously dedicated their time to participate in the field surveys. Their insights were invaluable to this research.

Abbreviations

The following abbreviations are used in this manuscript:

2SLS	Two-Stage Least Squares
AI	Artificial Intelligence
AIS	Agricultural Innovation Systems
CFA	Confirmatory Factor Analysis
CFI	Comparative Fit Index
DGI	Digital Governance Index
IoT	Internet of Things
IV	Instrumental Variable
MLE	Maximum Likelihood Estimation
OLS	Ordinary Least Squares
RBV	Resource-Based View
RMSEA	Root Mean Square Error of Approximation
SEM	Structural Equation Modeling
SRMR	Standardized Root Mean Square Residual
TAI	Technology Adoption Index
TCE	Transaction Cost Economics
TLI	Tucker-Lewis Index
VIF	Variance Inflation Factor

Appendix A. Research Questionnaire

Title: Impact of Digital Technologies on Agricultural Productivity and Rural Governance

Target Respondents: Farmers, Agri-tech Entrepreneurs, Rural Officials

Sample Size: 320 respondents

Instructions:

Please answer all questions honestly.
Your responses will remain confidential and used only for research purposes.
Tick (√) or select the most appropriate answer.

Section A: Demographics

Age Group: ☐ Under 25 ☐ 25–34 ☐ 35–44 ☐ 45–54 ☐ 55 and above
Gender: ☐ Male ☐ Female ☐ Prefer not to say ☐ Other: _________________
Education Level: ☐ No formal education ☐ Primary education ☐ Secondary education ☐ Diploma/Technical training ☐ Bachelor' degree ☐ Postgraduate degree
Occupation / Business Type: ☐ Farmer ☐ Agri-tech entrepreneur ☐ Rural government official ☐ Other: _________________
Farm Size / Business Scale: ☐ Small-scale (0–2 hectares) ☐ Medium-scale (2–10 hectares) ☐ Large-scale (10+ hectares)

Section B: Technology Adoption

Do you use digital tools in your agricultural or professional activities? ☐ Yes ☐ No
If yes, which types of technologies do you use? (Select all that apply) ☐ AI ☐ IoT devices ☐ Mobile applications ☐ Blockchain technology ☐ Drones ☐ Farm management software ☐ Other
Duration of usage of digital technologies: ☐ Less than 1 year ☐ 1–3 years ☐ 4–6 years ☐ More than 6 years
Frequency of use: ☐ Daily ☐ Weekly ☐ Monthly ☐ Rarely

Section C: Economic Outcomes

Digital technologies have improved my productivity (1 = Strongly Disagree, 5 = Strongly Agree): ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5
Estimated cost reduction after adopting digital technologies: ☐ No reduction ☐ Less than 10% ☐ 10–25% ☐ 26–50% ☐ More than 50%
Has the use of digital tools improved your access to new markets? ☐ Yes ☐ No ☐ Not sure
Increase in income due to digital technology use: ☐ No increase ☐ Slight increase ☐ Moderate increase ☐ Significant increase

Section D: Governance and Public Services

Do you use digital public services? ☐ Yes ☐ No
Digital public services are efficient and user-friendly (1 = Strongly Disagree, 5 = Strongly Agree): ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5
Level of trust in digital systems: ☐ Very low ☐ Low ☐ Moderate ☐ High ☐ Very high
Digital governance has improved service delivery in rural areas (1 = Strongly Disagree, 5 = Strongly Agree): ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Section E: Challenges and Barriers

What challenges do you face in adopting digital technologies? ☐ High cost ☐ Lack of skills/training ☐ Poor internet connectivity ☐ Lack of awareness ☐ Data privacy concerns ☐ Limited technical support ☐ Other

Section F: Future Intentions

Are you willing to adopt more digital technologies in the future? ☐ Yes ☐ No ☐ Maybe
What support would help you adopt digital technologies? ☐ Training programs ☐ Financial support/subsidies ☐ Better infrastructure ☐ Government policies ☐ Access to experts ☐ Other

Appendix B. Survey Results (N = 320)

Table B1. Demographic Distribution.

Category	Frequency	Percentage
Age Group		
Under 25	48	15.0%
25–34	96	30.0%
35–44	80	25.0%
45–54	56	17.5%
55+	40	12.5%
Gender		
Male	198	61.9%
Female	108	33.8%
Prefer not to say	8	2.5%
Other	6	1.8%
Education Level		
No formal education	28	8.8%
Primary	52	16.3%
Secondary	96	30.0%
Diploma	64	20.0%
Bachelor's	56	17.5%
Postgraduate	24	7.5%
Occupation		
Farmer	176	55.0%
Agri-tech entrepreneur	72	22.5%
Rural official	56	17.5%
Other	16	5.0%
Farm Size / Scale		
Small	160	50.0%
Medium	112	35.0%
Large	48	15.0%

Table B2. Technology Adoption Overview.

Question / Metric	Results
6. Use of Digital Tools	Yes: 232 (72.5%) | No: 88 (27.5%)
7. Types of Technologies (Multiple Response)	Mobile apps: 180 (77.6%), IoT: 96 (41.4%), Farm software: 88 (37.9%), AI: 64 (27.6%), Drones: 52 (22.4%), Blockchain: 24 (10.3%)
8. Duration of Use	< 1 year: 64 (27.6%), 1–3 yrs: 104 (44.8%), 4–6 yrs: 48 (20.7%), 6+ yrs: 16 (6.9%)
9. Frequency of Use	Daily: 112 (48.3%), Weekly: 72 (31.0%), Monthly: 32 (13.8%), Rarely: 16 (6.9%)

Table B3. Economic Outcomes and Governance Ratings.

Question / Metric	Results
10. Productivity Improvement	Likert scale Mean ≈ 3.7
11. Cost Reduction	None: 17.2%, < 10%: 31.0%, 10–25%: 34.5%, 26–50%: 13.8%, > 50%: 3.5%
12. Market Access	Yes: 168 (72.4%), No: 40 (17.2%), Not sure: 24 (10.4%)
13. Income Increase	None: 20.7%, Slight: 41.4%, Moderate: 27.6%, Significant: 10.3%
14. Use of Digital Public Services	Yes: 200 (62.5%), No: 120 (37.5%)
15. Efficiency of Services	Likert scale Mean ≈ 3.4
16. Trust in Digital Systems	Very low: 32, Low: 56, Moderate: 128, High: 80, Very high: 24
17. Governance Improvement	Likert scale Mean ≈ 3.5

Table B4. Challenges and Future Intentions.

Question / Metric	Results
18. Challenges (Multiple Response)	High cost (55.0%), Lack of skills (60.0%), Poor internet (52.5%), Lack of awareness (37.5%), Data privacy (30.0%), Limited support (45.0%)
19. Willingness to Adopt More Tech	Yes: 184 (57.5%), No: 40 (12.5%), Maybe: 96 (30.0%)
20. Needed Support (Multiple Response)	Training: 208 (65.0%), Financial support: 192 (60.0%), Infrastructure: 176 (55.0%), Gov. policy: 120 (37.5%), Experts: 144 (45.0%)

Key Survey Insights (Interpretation):

- High adoption (72.5%), mainly driven by mobile apps.
- Moderate productivity gains (Mean ≈ 3.7).
- Skills gap is the biggest barrier (60%).
- Strong future interest (57.5% willing).
- Governance systems show moderate trust and usability.

References

Barrett, C. B., Benton, T., Fanzo, J., Herrero, M., Nelson, R. J., Bageant, E., Buckler, E., Cooper, K., Culotta, I., Fan, S., Gandhi, R., James, S., Kahn, M., Lawson-Lartego, L., Liu, J., Marchall, Q., Mason-D'Croz, D., Mathys, A., Mathys, C., … Wood, S. (2022). Socio-technical innovation bundles for agri-food systems transformation. *Nature Sustainability*, *5*, 837–845. https://doi.org/10.1038/s41893-022-00963-5

Bongiovanni, R., & Lowenberg-DeBoer, J. (2004). Precision agriculture and sustainability. *Precision Agriculture*, *5*, 359–387. https://doi.org/10.1023/B:PRAG.0000040806.39604.aa

Deininger, K., & Feder, G. (2009). Land registration, governance, and development: Evidence and implications for policy. *The World Bank Research Observer*, *24*(2), 233–266. https://doi.org/10.1093/wbro/lkp007

Gebbers, R., & Adamchuk, V. I. (2010). Precision agriculture and food security. *Science*, *327*(5967), 828–831. https://doi.org/10.1126/science.1183899

Hall, A., Mytelka, L., & Oyeyinka, B. (2005). *Innovation systems: Implications for agricultural policy and practice*. Institutional Learning and Change Initiative. https://doi.org/10.22004/ag.econ.52512

Heeks, R. (2006). *Implementing and managing eGovernment: An international text*. SAGE Publications. https://doi.org/10.4135/9781446220191

Klerkx, L., Jakku, E., & Labarthe, P. (2019). A review of social science on digital agriculture, smart farming and agriculture 4.0: New contributions and a future research agenda. *NJAS: Wageningen Journal of Life Sciences*, *90–91*(1), 100315. https://doi.org/10.1016/j.njas.2019.100315

Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. *Sensors*, *18*(8), 2674. https://doi.org/10.3390/s18082674

Ma, W., Nie, P., Zhang, P., & Renwick, A. (2020). Impact of Internet use on economic well-being of rural households: Evidence from China. *Review of Development Economics*, *24*(2), 503–523. https://doi.org/10.1111/rode.12645

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple-mediator models. *Behavior Research Methods*, *40*, 879–891. https://doi.org/10.3758/BRM.40.3.879

Reardon, T., Echeverria, R., Berdegué, J., Minten, B., Liverpool-Tasie, S., Tschirley, D., & Zilberman, D. (2019). Rapid transformation of food systems in developing regions: Highlighting the role of agricultural research and innovations. *Agricultural Systems*, *172*, 47–59. https://doi.org/10.1016/j.agsy.2018.01.022
Stock, J. H., & Yogo, M. (2005). Testing for weak instruments in linear IV regression. In D. W. K. Andrews, & J. H. Stock (Eds.), *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg* (pp. 80–108). Cambridge University Press. https://doi.org/10.1017/CBO9780511614491.006
van Dijk, J. (2020). *The digital divide*. Polity Press.
Williamson, O. E. (1985). *The economic institutions of capitalism: Firms, markets, relational contracting*. Free Press.
Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M.-J. (2017). Big data in smart farming – A review. *Agricultural Systems*, *153*, 69–80. https://doi.org/10.1016/j.agsy.2017.01.023
World Bank Group. (2021). *World development report 2021: Data for better lives*. https://doi.org/10.1596/978-1-4648-1600-0
Zeng, Y., Jia, F., Wan, L., & Guo, H. (2017). E-commerce in agri-food sector: A systematic literature review. *International Food and Agribusiness Management Review*, *20*(4), 439–459. https://doi.org/10.22434/IFAMR2016.0156

Open Call for Papers

Bridging disciplines such as economics, sociology, and human geography, Agricultural & Rural Studies (**A&R, ISSN 2959-9784**) is a double-blind peer-reviewed international journal that fosters high-impact research connecting theoretical innovation with empirical depth, ultimately supporting evidence-based solutions for sustainable and equitable futures in agriculture and rural communities worldwide.

Aims

A&R seeks to:

- Foster critical dialogue on contemporary challenges and opportunities in agriculture and rural contexts, from local to global scales.
- Promote methodologically robust research—encompassing quantitative, qualitative, and mixed approaches—that generates actionable insights for academia, policymakers, and practitioners.
- Highlight interdisciplinary perspectives to address complex issues at the intersection of agrarian systems, rural societies, and policy frameworks.
- Amplify diverse voices, ensuring representation of marginalized regions, communities, and stakeholders in scholarly discourse.

Scope

A&R prioritizes, but is not confined to, research that engages with the following thematic areas:

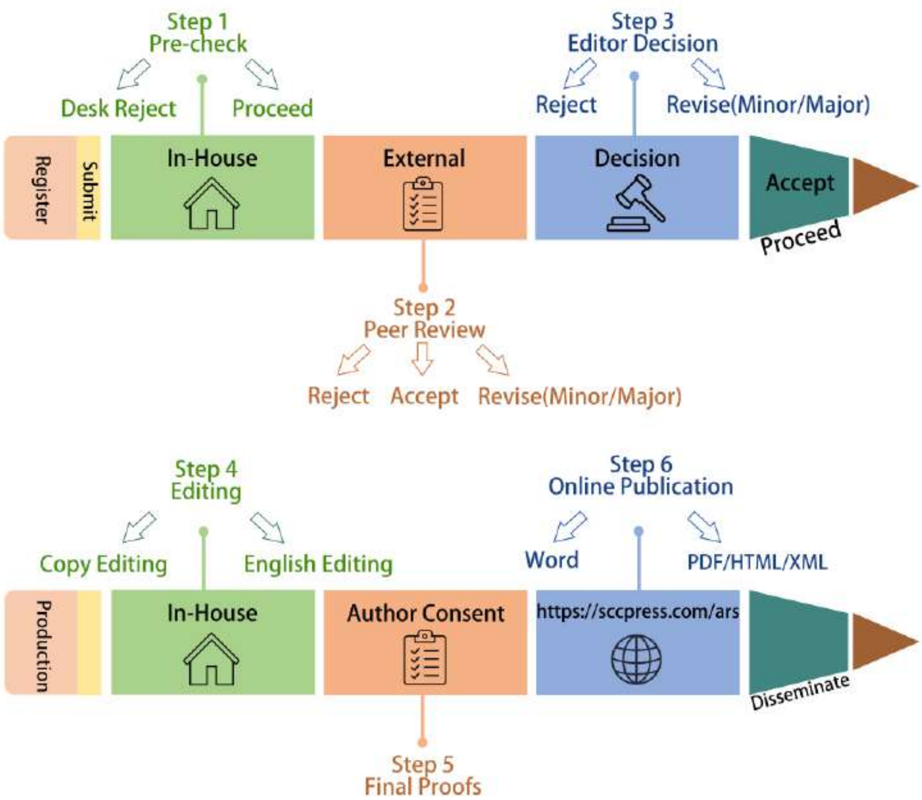
- Agricultural Systems & Sustainability: Studies on sustainable farming practices, agroecology, climate resilience, and the environmental, social, and economic dimensions of agricultural production and food systems.
- Rural Development & Livelihoods: Analyses of rural economic transformations, poverty alleviation strategies, rural-urban linkages, and the socio-cultural dynamics of rural communities.
- Farmer Well-being & Equity: Investigations into farmer agency, gender and social inclusion, labor dynamics, and the impacts of policies on rural households and vulnerable groups.
- AI Innovation & Technology: Research on smart agriculture, precision farming, rural e-commerce, and the socio-institutional implications of AI tools for agricultural and rural development.
- Policy & Governance: Evaluations of agricultural and rural policy frameworks, institutional arrangements, and regulatory mechanisms, with attention to their

effectiveness, equity, and adaptability across contexts.

- Emergent Challenges & Opportunities: Explorations of transformative trends—from biotechnological advancements to climate adaptation—to understand their potential to reshape agricultural and rural futures.

A&R prioritizes empirical research that connects localized insights to broader theoretical or policy debates, as well as critical reviews, case studies, and methodological innovations. All submissions undergo rigorous peer review to ensure scholarly excellence, relevance, and clarity. By bridging discipline boundaries and fostering global dialogue, **A&R** aims to be a cornerstone of interdisciplinary scholarship in agricultural and rural studies.

We cordially invite authors to submit original, previously unpublished manuscripts for consideration in **A&R**. All submissions should be made via our online submission system at https://sccpress.com/ars.



To Be

a Leading Academic Journal
at the Nexus of
Agricultural Sustainability,
Rural Revitalization & Farmer Well-Being.

Agricultural & Rural Studies

Abstracted/Indexed in databases including but not limited to
AGRIS, CABI, Crossref, Dimensions, DOAJ, Google Scholar, J-Gate, and Scopus.

Publisher
SCC PRESS, Unit 811, Beverley Commercial Centre,
87-105 Chatham Road South, Tsim Sha Tsui,
Kowloon, Hong Kong S.A.R., China.

Editorial Office
Rural Revitalization Academy of Zhejiang Province,
Zhejiang A&F University, Hangzhou, Zhejiang,
311300, China.

Editorial Assistants
Travis Malone (travismalone@sccpress.com)
Grace Li (grace@sccpress.com)
Rui Zhang (ruizhang@sccpress.com)

Photographer
Yangbi Converged Media Center

Layouter
Nan Wang

SCC PRESS Academic OA Publishing
Began in 2023