Czech University of Life Sciences Prague
Faculty of Economics and Management

agris on-line

Papers in Economics and Informatics

http://online.agris.cz

ISSN 1804-1930
XVII, 2025, 4

International scientific journal
Prague

Agris on-line Papers in Economics and Informatics

The international reviewed scientific journal issued by the Faculty of Economics and Management of the Czech University of Life Sciences Prague.

The journal publishes original scientific contributions from the area of economics and informatics with focus on agriculture and rural development.

Editorial office

AGRIS on-line Papers in Economics and Informatics
Department of Economics of FEM CZU Prague
Kamýcká 129, 165 00 Praha-Suchdol
Czech Republic
Phone: +420 224 382 056
E-mail: agrisonline(at)pef.czu.cz

Publisher

Faculty of Economics and Management
Czech University of Life Sciences Prague
Kamýcká 129, 165 00 Praha-Suchdol
Czech Republic
Reg. number: 60460709

ISSN 1804-1930

XVII, 2025,
30th of December 2025
Prague

Agris on-line

Papers in Economics and Informatics

ISSN 1804-1930

XVII, 2025, 4

Agris on-line Papers in Economics and Informatics

Volume XVII Number 4, 2025

Content:

Agris on-line Papers in Economics and Informatics

Volume XVII Number 4, 2025

A Highly Effective Deep Learning Tool for Identifying Plant Leaves

Adel Abdelhadi, Ouahab Kadri

Department of Computer Science, University of Batna 2, Algeria

Abstract

This work addresses pattern recognition in the agronomic domain, with a particular emphasis on identifying plant leaves using an adaptive neural network technique. We introduce a tool designed for two primary groups: botany researchers and a broader range of scientists applying it to plant identification and classification. We delve into the capabilities of Deep Learning, focusing on generalization abilities that enable accurate predictions on unseen data, which is essential for handling the variation in leaf shapes, sizes, and structures across species. The implementation details of these neural networks are described, including data preprocessing, network architecture design, training strategies, and evaluation techniques to ensure robustness and reliability in real-world applications.

Keywords

Pattern recognition, plant leaves, deep learning, classification, analysis, image processing, neural networks.

Abdelhadi, A. and Kadri, O. (2025) "A Highly Effective Deep Learning Tool for Identifying Plant Leaves", *AGRIS on-line Papers in Economics and Informatics*, Vol. 17, No. 4, pp. 3-9. ISSN 1804-1930. DOI 10.7160/aol.2025.170401.

Introduction

The problem of classifying plant leaves is closely related to the broader challenge of form recognition (FR), a fundamental step in human-machine communication. Image classification of plant leaves is complex due to several factors. First, forms exist in the physical world, and their digital transcription is often complicated by sensor limitations. Additionally, the nature and appearance of shapes can vary significantly between samples, even within the same plant family, increasing the complexity of the performance space and requiring more time for decisions and data analysis (Mufeng et al., 2023).

A key task in form recognition is accurately characterizing a shape based solely on its digital representation. This involves finding a description that distinguishes it from similar shapes and reduces misclassification risk. The description should capture features that allow comparison with neighboring shapes, ensuring only similar characteristics are grouped—a process commonly referred to as "learning" (Mufeng et al., 2023).

Another critical aspect is the decision-making process, which involves evaluating the input form and assigning it to the correct category by labeling it with a specific family or classification name. The system aims to find the best-matching family, typically by maximizing a similarity function between the input form's description and various family descriptions.

The complexity of form recognition tasks often requires processing vast amounts of data at high speed, challenging conventional methods. Traditional approaches struggle with the required processing power and intricate recognition algorithms. A promising research direction involves using artificial systems that autonomously adjust parameters, enabling generalization and adaptation to various input conditions—embodied in Deep Learning (Mufeng et al., 2023).

Deep Learning offers effective solutions for pattern recognition problems. Its learning capabilities reduce the need for human effort in research and comparison, while its adaptability enables recognizing patterns not explicitly trained on, making it powerful for complex problems like plant leaf classification.

Although Deep Learning for pattern recognition isn't new, its application to plant leaf classification is relatively novel. Previous research faced challenges leading to abandonment of the approach, but renewed interest in neural networks, especially with advancements in learning and generalization, has made them suitable for overcoming these challenges. This work ensures a clear presentation of ideas and concepts, contributing to improved plant leaf classification using Deep Learning.

Related work

Most existing plant identification work doesn't focus on automation, particularly in determining plant type by analyzing leaf properties, largely due to absent conventional classification criteria. Traditional approaches often rely on manual methods where experts use physical attributes like leaf shape, size, and fruit characteristics—effective but time-consuming, subjective, and not easily scalable. Recent comprehensive reviews, such as that by Patil and Shirdhonkar (2022), have cataloged the rapid advancement of deep learning approaches in this domain, highlighting a clear shift towards automated, data-driven solutions.

For example, Konstantinos et al. (2018) developed a reference manual of approximately 170 plant species categorized by fruit edibility and leaf shape, valuable but relying on individual expertise and manual comparison. Similarly, Arif Wani et al. (2019) published a book listing 250 tree types categorized by leaf characteristics, limited by textual descriptions and manual identification requirements.

Our work aims to bypass these manual procedures by providing an efficient, automated tool for accurate, quick plant species identification using leaf samples through modern computational methods, specifically neural networks and pattern recognition (Jackulin et al., 2022). This aligns with the broader trend of applying improved deep learning approaches for complex plant recognition tasks, as demonstrated in works on disease localization by Alqahtani et al. (2023).

Materials and methods

Deep Learning Applications for Leaf Pattern Recognition

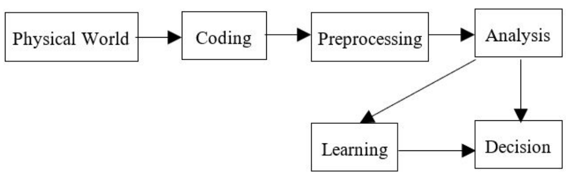
Deep learning has emerged as a powerful tool for pattern recognition, particularly in classifying and identifying plant species based on leaf characteristics using Convolutional Neural Networks (CNNs) to learn complex patterns from large leaf image datasets (Aanis et al., 2023). Its effectiveness is evident in various agricultural applications, including the automated identification of specific diseases like Northern Leaf Blight in maize from field imagery, as successfully shown by Chad et al. (2017). Recent architectural innovations, such as multi-scale feature fusion networks (Li et al., 2023), have further enhanced the ability of CNNs to capture discriminative features at various levels of abstraction, from fine textures to global shapes.

The classical approach to recognition problems involves several key steps (Guoqiang et al., 2016), shown in Figure 1:

1. **Preprocessing:** Cleaning and preparing raw data through noise reduction, image normalization, or enhancement.
2. **Feature Extraction:** Extracting relevant features that distinguish between classes/ patterns.
3. **Classification:** Using algorithms to assign input to specific categories.
4. **Post-processing:** Refining results for improved accuracy.

Figure 1 illustrates this process from input data to final classification results.

Physical World → Coding → Preprocessing → Analysis → Learning → Decision



Source: Guoqiang et al, 2016

Figure 1: Overall diagram of the Pattern Recognition System.

We explain each step in our recognition process:

- **Physical World:** Modeled as infinite-dimensional "form space" with objects described through various properties.
- **Coding:** Conversion to digital representation (representation space).
- **Pre-treatment:** Selecting relevant information by eliminating noise, standardizing data, and removing redundancy.
- **Analysis:** Computing characteristics/ parameters using recognition techniques.
- **Learning:** Refining decision-making by incorporating prior knowledge about shapes.
- **Decision:** Actual recognition using knowledge acquired during learning.

Both learning and decision phases involve neural networks—specifically a multilayer network with:

- Inputs/outputs corresponding to plant leaves and their species
- Error backpropagation learning algorithm

- Input number based on analysis phase parameters
- Output number corresponding to plant species classes (Guoqiang et al., 2016).

Characteristics of forms

Forms possess distinct features providing valuable information for pattern recognition, including perimeter, area, and other characteristics that describe shape and structure. These can be derived from both binary images (simplified black/white representations) and original images (detailed color information).

Characteristics are grouped into two categories:

1. **Boundary Features:** Describe outer contour/ perimeter
   - Perimeter: Boundary outline length
   - Compactness: Area/perimeter ratio
   - Convexity: Boundary convex/concave degree
   - Smoothness: Boundary regularity
2. **Interior Features:** Describe form interior
   - Area: Total covered space
   - Centroid: Center of mass
   - Shape descriptors: Elongation, aspect ratio, roundness (Ahmed et al., 2012).

Parameter extraction from plant leaves

Parameter extraction focuses on characteristics describing both shape and color variations, categorized as:

1. **Shape parameters:** Geometric features (perimeter, area, etc.)
2. **Color Parameters:** Color properties and variations

Single parameters may be insufficient for distinguishing between species.

Shape parameters

Essential for analyzing leaf shapes and environmental adaptations:

- **Compactness (C)**

$$C = \frac{Su}{Pe} \tag{1}$$

Where

Su = area (pixels within boundary), Pe = perimeter (pixels along boundary) demonstrates perimeter-area relationship (Figures 2-3.

Source: Authors

Figure 2: Efficiency of compactness (1) - Larger perimeter enclosing larger area.

Source: Authors

Figure 3: Efficiency of Compactness (2) - Same perimeter enclosing smaller area.

- **Elongation (Elong)**

$$Elong = \frac{Leng}{Width} \tag{2}$$

Where Leng = longest dimension (Ext1-Ext2), Width = broadest dimension (Ext3-Ext4) distinguishes broad-leaved vs. narrow-leaved types (Figures 4-5)

Source: Authors

Figure 4: Elongation of the leaf.

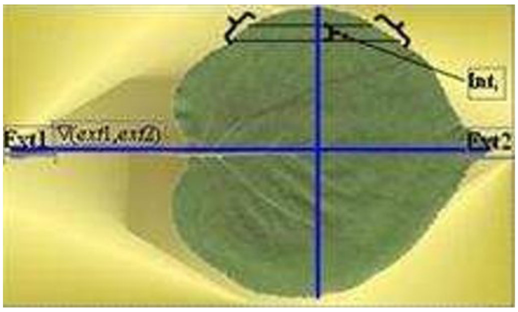
Source: Authors

Figure 5: Elongation measurement for leaves with similar shapes.

- **Level parabolic (LP)**

$$LP = \frac{Su}{Srect} \tag{3}$$

Where Srect = rectangle area around leaf using extreme points quantifies leaf curvature (Figure 6).



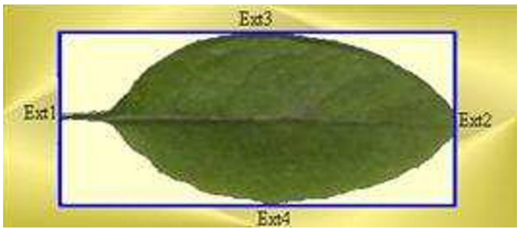
Source: Authors

Figure 6: Level parabolic shape.

- **Relativity (R)**

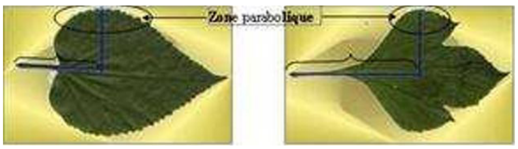
$$R = \frac{D1}{D2} \tag{4}$$

Where D1 = distance between Ext3 projection and Ext1, D2 = total leaf length distinguishes area distribution near supports (Figures 7-8).



Source: Authors

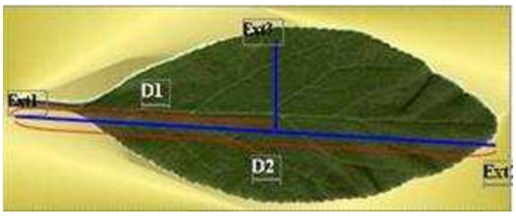
Figure 7: Relativity.



Source: Authors

Figure 8: Efficiency of LP.

- **Concentration of Intervals of Distance (CID):** Ten sub-parameters representing contour points within distance intervals from leaf end Correlates with LP shape (Figure 9).



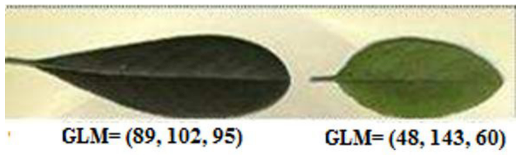
Source: Authors

Figure 9: Concentration of distance intervals.

Color parameters

Provide essential insights into plant morphology, health, and adaptation:

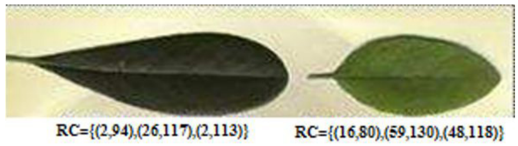
- **Gray Level Medium (GLM):** Three sub-parameters (R-CM, CM-G, CM-B) representing average color values quantifies color variations (Figure 10)



Source: Authors

Figure 10: Efficiency of GLM.

- **Range of Color (RC):** Indicates color variation range within leaf, assesses color change across surface (Figure 11).



Source: Authors

Figure 10: Efficiency of GLM.

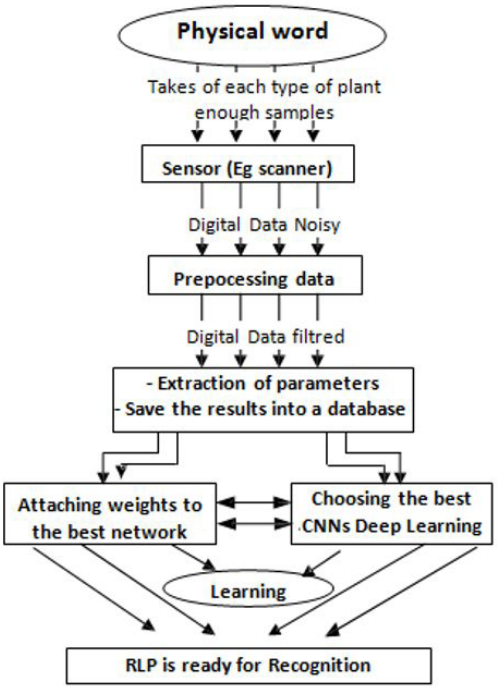
Results and discussion

PLR Tool (Plant Leaf Recognition Tool)

The PLR Tool simplifies plant species identification for researchers working in diverse environments with unfamiliar species. Unlike traditional time-consuming methods, it uses image recognition technology—users input leaf images, and the tool analyzes key features (shape, color, texture) against an extensive database for automatic identification. This automation saves time and increases accuracy, benefiting plant biologists, ecologists, and other researchers, particularly during fieldwork where quick, reliable identification is crucial. The tool also helps build and update plant databases.

General architecture of PLR

The PLR architecture combines data acquisition, image preprocessing, deep learning for feature extraction, classification, and output generation for accurate, efficient plant species recognition (Figure 12).



Source: Authors

Figure 12: General Architecture of RLP.

Sample characteristics

For optimal results, samples must meet criteria:

- Representative of plant types (collected at different growth stages)
- Exclude pathological cases with atypical characteristics
- Include leaf support (stem/petiole) for accurate identification

CNN selection for deep learning

Optimal CNN selection involves:

1. Determining iteration number
2. Validating network usage
3. Training all parameter combinations
4. Selecting network with smallest error

Application example

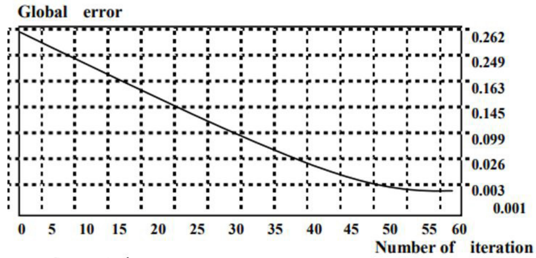
We tested the tool on 100 plant species, divided into:

- **Learning class:** 4-8 leaves/plant for CNN training
- **Test class:** 4 leaves/plant for evaluation CNN showed good performance, especially with consistent shapes. Semi-distributed coding improved convergence speed approximately threefold versus conventional coding.

Results

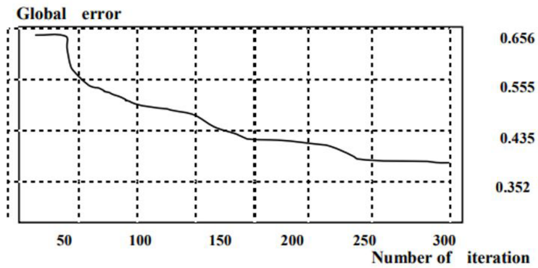
During training (100,000 iterations), CNN showed significant error reduction, especially with semi-distributed coding (error dropped from 0.7 to 10^{-8}). With semi-distributed coding, CNN correctly identified 83/100 species (85% success) versus 70% with conventional coding. This comparative result aligns with findings in the literature, where the choice of encoding and feature representation has been shown to significantly impact model performance in plant classification tasks (Sethy et al., 2024). The demonstration of improved accuracy, faster convergence, and better data handling underscores the efficacy of our approach.

Furthermore, the performance boost from semi-distributed coding suggests a pathway for future work. Integrating more sophisticated mechanisms, such as self-attention or lightweight architectures with dedicated attention modules (Zhang et al., 2022; Mbouembe and Ko, 2024), could potentially build upon this foundation to achieve even higher accuracy and efficiency.



Source: Authors

Figure 13: Graph of error (semi distributed coding).



Source: Authors

Figure 14: Graph of error (conventional coding).

The CNN achieved a strong performance using semi-distributed coding, accurately classifying 83 out of the 100 plant species, yielding a success rate of 85%. In contrast, the conventional coding approach achieved a notably lower accuracy of 70%. These results indicate that semi-distributed coding not only enhanced the model's generalization and classification capabilities but also facilitated faster convergence and more robust data processing compared to the conventional method.

Conclusion

Our objective was to create a comprehensive tool for botany researchers and scientists across disciplines. We identified mathematical parameters based on leaf shape measurements that enable plant classification into distinct categories, automated through image processing without user intervention. Using Deep Learning for recognition/classification with careful network selection, we achieved over 85% success rate with semi-distributed coding versus 70% with conventional coding. While satisfactory, this could potentially be improved by incorporating additional parameters like chemical and cellular characteristics.

Corresponding author:
Adel Abdelhadi
Department of Computer Science, Batna 2 University
53, Constantine Road, Fesdis, Batna 05078, Algeria
Phone: +213 559 364 312 / +213 657 267 599, Email: a.abdelhadi@univ-batna2.dz

References

[1] Aanis, A., Dharmendra, S. and Aly, E. (2023) "A survey on using deep learning techniques for plant disease diagnosis and recommendations for development of appropriate tools", S*mart Agricultural Technology*, Vol. 3., p. 100083. ISSN 2666-0159. DOI 10.1016/j.atech.2022.100083.

[2] Ahmed, A. A. and Gopireddy, H. R. (2021) "A Mobile-Based System for Detecting Plant Leaf Diseases Using Deep Learning", *AgriEngineering*, Vol. 3, No. 3, pp. 478-493. ISSN 2673-1906. DOI 10.3390/agriengineering3030032.

[3] Alqahtani, M., Nawaz, Y., Nazir, T., Javed, A., Jeribi, F. and Tahir, A. (2023) "An improved deep learning approach for localization and recognition of plant leaf diseases", *Expert Systems with Applications*, Vol. 230, p. 120717. ISSN 0957-4174. DOI 10.1016/j.eswa.2023.120717.

[4] Chad, D., Tyr, W., Siyuan, C., Ethan, L., Jason, Y., Michael, A., Rebecca, J. and Hod, L. (2017) "Automated Identification of Northern Leaf Blight-Infected Maize Plants from Field Imagery Using Deep Learning", *Phytopathology*, Vol. 107, No. 11. ISSN 0031-949X. DOI 10.1094/PHYTO-11-16-0417-R.

[5] Guoqiang, Z., Lina, W., Xiao, L. and Junyu, D. (2016) "An overview on data representation learning: From traditional feature learning to recent deep learning", *The Journal of Finance and Data Science*, Vol. 2, No. 4, pp. 265-278. ISSN 2405-9188. DOI 10.1016/j.jfds.2017.05.001.

[6] Jackulin, C. and Murugavalli, S. (2022) "A comprehensive review on detection of plant disease using machine learning and deep learning approaches", *Measurement: Sensors*, Vol. 24. p. 100441. ISSN 2665-9099. DOI 10.1016/j.measen.2022.100441.

[7] Konstantinos, P. F. (2018) "Deep learning models for plant disease detection and diagnosis", *Computers and Electronics in Agriculture*, Vol. 145, pp. 311-318. ISSN 0168-1699. DOI 10.1016/j.compag.2018.01.009.

[8] Lee, S. H., Chan, C. S., Mayo, S. J. and Remagnino, P. (2017) "How deep learning extracts and learns leaf features for plant classification", *Pattern Recognition*, Vol. 71, pp. 1-13. ISSN 0031-3203. https://doi.org/10.1016/j.patcog.2017.05.015.

[9] Li, W., Sun, Y., Li, J., Wang, H. and Zhang, C. (2023) "A multi-scale feature fusion convolutional neural network for high-performance plant leaf recognition", *Computers and Electronics in Agriculture*, Vol. 214, p. 108294. ISSN 0168-1699. DOI 10.1016/j.compag.2023.108294.

[10] Mufeng, T., Tommaso, S., Beren, M., Yuhang, S., Thomas, L. and Rafal, B. (2023) "Recurrent predictive coding models for associative memory employing covariance learning", *PLoS Computational Biology*, Vol. 19, No. 3, p. e1010719. ISSN 1553-7358. DOI 10.1371/journal.pcbi.1010719.

[11] Nazir, T., Iqbal, M. M., Jabbar, S., Hussain, A. and Albathan, M. (2023) "EfficientPNet—An Optimized and Efficient Deep Learning Approach for Classifying Disease of Potato Plant Leaves", *Agriculture*, Vol. 13, No. 4, p. 841. ISSN 2227-7072. DOI 10.3390/agriculture13040841.

[12] Wani, M. A., Ahmad Bhat, M. A., Saduf, F. and Asif Iqbal, K. A. (2019) "Basics of Supervised Deep Learning", In: *Advances in Deep Learning*, Chapter 2, pp. 13-29. ISBN-13 978-9811367939. DOI 10.1007/978-981-13-6794-6_2.

Volume XVII Number 4, 2025

Factors Influencing Consumer Preference Towards Horticulture Geographical Indications (GIs): A Case of Udupi Brinjal from South India

Vikram Baliga, Santhosha Shetty G, Krithika Rao, Mayuri, Ren Trevor Dias

Centre for Consultancy Training and Corporate Interface, Manipal School of Commerce and Economics, Manipal Academy of Higher Education (MAHE), Manipal, India

Abstract

Consumers' attention towards fruits and vegetables in recent times has shifted to regionally grown geographical indications (GI) due to the quality and origin of these products. This research work aims to explore the factors influencing consumer preference towards Udupi Brinjal, a horticulture GI grown in Udupi District of South India. The present study has used a mixed-method approach to gather and analyze data collected from local consumers of Udupi District. The qualitative study design involved a survey of key informants in the local region. Subsequently, data collected from consumers through a structured questionnaire were analyzed using factor analysis and regression techniques. Results from data analysis revealed that quality factors show greater importance in predicting consumer preference, followed by sensory attributes and health-related aspects. The results will help formulate marketing strategies for horticulture GIs. Agri-business marketers and State-owned agriculture promotion agencies can adopt these strategies to promote GIs and gain consumer acceptance. The results and discussions of this research work are consistent with Sustainable Development Goals (Goals 2 & 12) and contribute to sustainable agriculture.

Keywords

Consumer preference, Geographical Indications (GIs), sustainable agriculture, marketing strategies, horticulture marketing, territorial specialties, region-of-origin products.

Baliga, V., Shetty G, S., Rao, K., Mayuri and Dias, R. T. (2025) "Factors Influencing Consumer Preference Towards Horticulture Geographical Indications (GIs): A Case of Udupi Brinjal from South India", *AGRIS on-line Papers in Economics and Informatics*, Vol. 17, No. 4, pp. 11-26. ISSN 1804-1930. DOI 10.7160/aol.2025.170402.

Introduction

Geographical Indications (GIs) have attracted great literary attention in recent years. GIs are goods or products originating from a particular territory, region, or locality with a special characteristic, attributed to its geographical origin. GIs are recognized for the tremendous benefits they carry, the value they provide, and their ability to contribute to territorial development (Vinayan, 2017). From the consumers' perspective, GIs act as an assurance tool that protects against unfair competition arising out of counterfeiting and impostering from other competing varieties. In India, consumers have realized the need to consume safe and healthy products. A large portion of the Indian market constitutes young consumers who are aware of sustainability and health. Considering safety and good health, consumers have attached greater importance to products with GI tags, as they offer assurance of quality and have raised consumer expectations regarding the originality of the products (Kokthi and Kruja, 2016).

Several studies have reported that GIs provide additional revenue to the producers by influencing the consumers to pay a premium value (Cardoso et al., 2022; Biljana, 2022; Durand and Fournier, 2017; Vinayan, 2017). Additionally, studies conducted on GI products in developing countries have emphasised the need for better protection against counterfeiting (Ihsaniyati et al., 2022; Zhan et al., 2017). At the field level, many agricultural goods and services (GIs) are prone to counterfeiting and faking in India. The problem of counterfeiting and faking was referred to in another study, which revealed that consumers appreciate GI certification and brands as they provide an advantage of origin-guaranteed products (Toklu et al., 2020).

The subject of Geographical Indications is of global importance, given the growing number of traditional products grown and manufactured in the current times. India has registered 643 Geographical Indications as of February 2025 (Office of the Controller General of Patents, Designs & Trademarks, Ministry of Commerce, Government of India). The agriculture sector has been a second major contributor to the list of traditional products, with 202 products registered as GIs. Several researchers have examined the trends and consumer preferences (Kaliji et al., 2019) toward GI products. Studies have reported that GI tagging is an effective marketing tool that provides premium value to farmers and makes their products competitive in the market (Yin et al., 2024). GIs support the economic growth of rural livelihoods and contribute to regional development (Lambada-Lehnhardt et al., 2021). From the ecological perspective, studies have revealed that branded agricultural products with GIs can facilitate the reduction of agricultural carbon emissions and favourably affect consumer attitudes (Uzar et al., 2022; Zhang et al., 2024).

In India, the subject of Geographical Indications is under immense scrutiny by policymakers. The National Bank of Agriculture and Rural Development (NABARD) is an apex development bank set up by the Government of India to foster rural prosperity. NABARD has taken Udupi Brinjal as a select case to promote sustainable, equitable agriculture and support rural development. At the local level, promoting human health and providing quality products have been identified as the main pillars of responsible consumption (Sustainable Development Goals 12). Hence, increasing consumption by determining and adopting effective marketing strategies is key to the sustainability of GI products.

About Udupi Brinjal – A snapshot

Udupi Brinjal is a Geographical Indication (GI) eggplant grown in the Udupi District of South India. It is grown by 200 farm families in Mattu village with GPS coordinates of 13°15′33″N 74°44′16″E. The Geographical Indication tag was awarded to the farmers of Mattu village as a community right in 2011. The field-level operations of Udupi Brinjal are consistent with the business model owned and operated by the members of the local community, popularly known as the Community-Based Enterprises (CBE) model. The Udupi Brinjal Growers Association demonstrates the characteristics of a business enterprise that is set up to meet the common good of the local people. Additionally, Udupi Brinjal operates on local knowledge and relies on local culture, local resources, and local capacity, similar to the features of a Community-based Enterprise (Peredo and Chrisman, 2006). Therefore, Udupi Brinjal is consistent with the Community-based Enterprise (CBE) model and has been considered an effective tool to alleviate poverty at the local level and achieve sustainable development of a collective enterprise (Handy et al., 2011; D'Souza and Joshi, 2020).

Udupi Brinjal, a kharif crop, is grown between October and June. It is locally popular as "Mattu Brinjal". This green-coloured aubergine is normally spherical in shape and low in moisture content. Scientifically known as "Solanum melongena" from the Solanaceae (potato) family, this fruit is rich in vitamins and minerals (Nandi et al., 2021; Nadeeshani et al., 2021). Popularly classified as a vegetable, the brinjal is a fruit belonging to the botanical family of berry (FoodData Central Food Details Eggplant, Raw, 2018). Brinjal is a staple in Indian cuisine as well as in Bangladesh and other Mediterranean diets. It is used in preparing different types of curry and pickle. Udupi Brinjal farmers first offer the yield to the Udupi Krishna (deity) temple, and then a part of the harvest is used to recover seeds for the next cropping season. Rest is sold in the market (Peschard, 2022).

Even though the GI tag has bestowed a special status on Udupi Brinjal, the farmers have not been able to leverage its true market potential (Vinayan, 2017; Tan et al., 2024). Some of the main reasons for poor distribution and low consumer patronage for Udupi Brinjal are a) Poor market infrastructure coupled with weak market linkages in the local and neighbouring districts b) Low consumer acceptance due to lack of availability and awareness of Udupi Brinjal c) Lack of effective marketing practices that is needed to create awareness in untapped markets and boost sales d) Poor branding efforts and the inability of the farmers to understand the power of the brand which is made up of its tools like the GI logo and GI certificate. Given this scenario, the study is undertaken to examine the key factors affecting consumer preference towards Udupi Brinjal. With this background, this study proceeds to answer the following research questions:

1. What are the main factors affecting consumer preferences towards Udupi Brinjal?

2. What is the impact of such factors on consumer preference towards Udupi Brinjal?
3. What are the marketing strategies that can be formulated to increase the consumer preference for Udupi Brinjal?

This research has used a mixed-method approach to predict consumer preference towards Udupi Brinjal. In the first stage, the key informant's survey was used to identify factors affecting consumer preference for Udupi Brinjal. Subsequently, in the second stage, quantitative methods were used to analyse data collected from consumers.

The succeeding section of this work presents a review of previous studies on Geographical Indication products encompassing consumer attitudes, preferences, sustainable consumption, and consumer demographics.

Review of literature

In the context of emerging economies, limited studies have analysed health consciousness and food safety as the key determinants driving consumers' preference towards GI fruits and vegetables. It has been observed that health-conscious consumers (Wang et al., 2021) pick their purchases from local farmers who grow GIs and from organic food stores (Yu et al., 2024) and direct farm-to-store outlets. It has been found that risk aversion has a positive effect on consumer choice towards fruits and vegetables. Results reveal that higher-income groups and level of education influence the purchase of vegetables (Palaniappan and Radhakrishnan, 2020). Regarding product attributes, value-driven aspects, health-related aspects (Vijayan et al., 2019), and purchasers' lifestyles are found to influence purchase decisions towards GI fruits and vegetables. In the case of seasonal vegetables, price, quantity, and frequency of intake, liking, and intention to pay positively affect consumer preference (Herath, 2019).

Consumers in the lower-income group are driven by affordability and the availability of convenience foods. The consumption of traditional foods is influenced by both sensory attributes and cognitive aspects (Moyo et al., 2023). Consumer lifestyle attributes serve as a basis for segmenting consumers into distinct groups to analyse their behaviour toward vegetables and fruits. Studies that have used choice models (Zhu et al., 2022; Ortega et al., 2016) have reported that marketing mix variables, price, family, cost, and demographics of the consumer predicted the purchase of food products. Other studies (Vijayan et al., 2019; Rejeki et al., 2021) have reported a radical shift in consumer choice towards organic vegetables and fruits. Food preferences are largely influenced by customer loyalty (Kosiciarova et al., 2020). Studies that applied the Theory of Planned Behaviour (Ajzen, 1991) have revealed that price and quality, packaging, consumer attitudes, behavior approved by an individual or group of important people (subjective norms), and perceived behavioral control aspects influenced consumer purchase intention towards organic vegetables (Uzar and Filipovic, 2023; Choudhury et al., 2020). Also, nutrients present in Brinjal have the properties to cure several chronic diseases (Naeem and Ugur, 2019). Further, consumer preference towards the quality of fruits is influenced by the location where the fruits are grown. Consumers are ready to pay more for premium quality organic products (Wang et al., 2019; Lee et al., 2020). However, there are limited studies that test the effect of allergic reaction, deterioration in the quality of vegetables and fruits due to pest attack, and availability of chemical-free products that influence consumer choice. Our research work has identified such quality-related items to study their impact on consumer preference. A study conducted on Brinjal in India has highlighted the need for farmers to adopt sustainable management practices and improve the chances of consumer acceptance (Gautam et al., 2019). Studies have observed that the consumer is the central entity while considering and settling supply chain issues (Raut et al., 2020; Berkile et al., 2019). Due to this, marketing margin driven by consumer choices is the most important component in determining overall profitability in the supply chain.

Results have confirmed that consumers regarded high knowledge, region of origin, and education as the main indicators influencing the consumption of GIs (Uzar et al., 2022). A study by Zhan et al. (2017) has revealed that brand loyalty towards GIs is determined by consumers' attitudes toward growing areas, perceived quality, and cognition of protection. Currently, food consumption and food safety have been topics of intense scrutiny and debate in India. Regulators involved with Food Safety Standards (Food Safety and Standards Authority of India) are closely monitoring deviations, especially in the case of misleading advertisements circulated by companies, leading to unfair trade practices.

Food safety issues prevalent in the local food industry pose a serious threat to public health.

Udupi Brinjal was awarded the GI tag for its unique taste, which is a sensory attribute. Several studies have reported that sensory attributes and consumer preference are related (Caskey et al., 2021; Moyo et al., 2023). Considering the importance gained by sensory attributes in earlier studies and their significance to our study, the sensory factor has been included and analyzed in the proposed model. A study by Chen (2021) analysed the combined effect of sensory attributes and quality aspects on consumer preference and reported that the taste (Bytyci et al., 2024) of joy of quality has a higher influence compared to price-related concerns. Another study analysed the relationship between taste experiments and brain responses of coffee consumers to observe preference patterns (Artencio et al., 2021). Empirical results from earlier studies have not fully covered the combined effect of quality aspects, health factors, and sensory attributes on consumer preferences in the case of agricultural GIs. Our research work has attempted to bridge this gap by placing thrust on sensory attributes, quality, and health aspects as the three key factors influencing consumer preference towards Udupi Brinjal. GIs are unique products, and examining the combined effect of key factors influencing the consumers' preference is likely to provide interesting results. Such results will enable marketers to formulate strategies and increase the chances of consumer acceptance. Given this background, the objective of the study is to:

1. determine the factors affecting the consumers' purchase of Udupi Brinjal.
2. examine the impact of these factors on consumer preference towards Udupi Brinjal.
3. formulate and suggest marketing strategies to improve consumer acceptance of Udupi Brinjal.

The study is hypothesised as:

H1: Health factors positively affect consumer preference towards Udupi Brinjal

H2: Sensory factors positively affect consumer preference towards Udupi Brinjal

H3: Quality factors positively affect consumer preference towards Udupi Brinjal

The next section presents the research instrument used to collect that information from participants, the sampling design, and the techniques used for data analysis.

Materials and methods

This study was conducted using the mixed-method approach in three stages. In the first stage, the researchers conducted a Focused Group Discussion (FGD) of key informants (7 Nos) (Chazdon and Lott, 2010; Whittaker, 2012). The key informants included three retail vegetable and fruit merchants, two community leaders of the Udupi Brinjal Growers' Association, and two agri-business experts from the Udupi District Horticulture Board. The key informants were purposively selected for the FGD from Mattu village based on their expertise and experience in the production and sale of Udupi Brinjal. The researchers have a collective experience of over two decades in horticulture marketing. Moreover, their close connection with the Udupi Brinjal ecosystem, dating back to 2012, has played a vital role in the recruitment of key informants. It is believed that the informants purposively selected have firsthand knowledge about the community, which comprises farmers and consumers of Udupi Brinjal, gained through their roles as community leaders or influencers. Further, key informants are recruited for this study as they are familiar with the research topic. The three vegetable and fruit merchants selected have been in the wholesale and retail trade of vegetables and fruits in the Udupi district for over 30 years. Moreover, their distribution network is spread across other taluks of Udupi. Hence, merchants operating from Mattu junction (2) and Agriculture Produce Marketing Committee, Udupi (1) were recruited for this purpose. The merchants thus selected are in close contact with Mattu farmers and local consumers. They also serve as intermediaries in the agri supply chain. It is felt that the merchants can provide diverse information about consumer preferences towards Udupi Brinjal. Also, farmer community leaders (2) who were instrumental in securing the GI tag for Udupi Brinjal in 2011 were interviewed based on their seniority and experience. The leaders selected possess extensive experience in profiling consumers during the process of filing GI applications. Community leaders who were selected have frequent interactions with Mattu farmers and local consumers, as well as local agriculture development agencies such as the Horticulture Board and the Agriculture Produce Marketing Committee. Also, agriculture experts (2) who have worked in projects related to the promotion of farmers' producer organisations. were recruited for the study. The Focused Group Discussion of key informants was conducted in August 2023 in Mattu village. Research associates trained in ethnography and in-depth interviews conducted the discussion with key informants. The discussion was recorded and documented, and the inputs collected were used to carry out quantitative analysis.

Thus, key informants are felt to be the right source of useful and reliable information. The Focused Group Discussion (FGD) with key informants is believed to bring to light diverse issues regarding consumers' purchase preferences for Udupi Brinjal. Additionally, this method has the quality to identify items that are important to include as variables in a survey. To conduct the FGD, a discussion guide (Hennink et al., 2020) was used, which included opening questions, main questions, and closing questions. The opening questions solicited participation on issues like the nature of business/work, type of customers, ways of interacting with the community, farmers, and consumer profiles linked to Udupi Brinjal. The participants shared their knowledge and experience on consumers' choices and preferences, likes, dislikes, motivations, opinions, and views as a response to the main questions in the discussion. Interestingly, the key informants stated that health-related aspects like diet, nutrition, oil-free intake, attributes like taste and flavour, different types of preparations of Udupi Brinjal, and quality of the produce motivated consumers to buy Udupi Brinjal. The Focused Group Discussions have generated statements required for the questionnaire. Such statements have evolved from open discussions among participants, their diverse perspectives based on their experience, and in-depth insights. The vegetable and fruit vendors stated that consumers seek more information regarding the freshness of produce and the ease of preparation. A few consumers raise concerns about the allergic reaction that Brinjal is likely to cause occasionally, especially on the lips and tongue. Also, pest invasion on the crop has been a menace that has affected the quality of the crop. Also, the vegetable and fruit vendors referred to the growing demand for organically grown vegetables and fruits. Such consumers are health-conscious and believe in maintaining a ‘wellness-oriented' lifestyle, mainly focused on nutrition, fitness, and diet, and free from stress. The vendors also mentioned that Udupi Brinjal is facing tough competition from other local varieties. According to the merchants, Udupi Brinjal is preferred over the local competing variety mainly due to its unique taste. Halfway into the discussion, the community leaders were requested to provide their perspectives on challenges and issues they encounter in marketing and selling Udupi Brinjal. Community leaders stated that increasing the regularity of purchases has been a major issue that has hindered sales. Given this situation, community leaders expressed support from the academic community and marketers who could deep dive into exploring and understanding consumers' preferences and buying behaviour. During the discussion, the agriculture experts brought up the aspect of the size of Udupi Brinjal being a key determinant that influences its purchase. It was also suggested that the Udupi Brinjal Growers Association should scale up its branding and labelling efforts to stay relevant in the market. Thus, the views, opinions, and observations collected from the focus group discussion were transcribed, and the variables of research interest were identified. In the second stage, the transcribed information went through a text analysis of these inputs/descriptions to identify the main items and factors related to the study. After performing a text analysis of the descriptions, the items were put forth as statements in a structured questionnaire suitable for quantitative analysis. In the third and final stage, a survey questionnaire was designed to collect the data from the respondents. The first section of the questionnaire invited respondents to provide their demographic details. The next section included questions on variables in the form of statements. These statements were presented to the respondents anchored on a ‘5-point’ Likert scale as 1 = strongly disagree, 2 = Disagree, 3 = neither agree nor disagree, 4 = Agree, 5= Strongly agree. Likewise, consumer preference (CP) was also anchored using a 5-point Likert scale.

The opening section of the survey contained a consent form that had a detailed explanation of the research objectives. In the beginning, the participants were asked to tick their free consent with a yes/no box before proceeding with the other questions. The consent form also provides the right for the participants to leave the survey at any time before participating. The consent was obtained in written form. Regarding the selection of consumers for the survey, the researchers used the ‘Customer Ledger’ maintained by the Udupi Brinjal Growers Association. The customer's ledger had a list of 1036 customers who visited the warehouse of the association between October 2023 to June 2024. For screening, the researchers further filtered this list to select customers who had visited the warehouse at least twice during this period. Repeat purchases by a customer within one season were considered a testament of loyalty towards Udupi Brinjal. Out of 1036 customers, it was found that 367 buyers have purchased Udupi Brinjal at least twice directly from the association during the survey period. The respondents were contacted based on the available consumer group of loyalists. The researchers contacted all 367 buyers through the residential addresses/cell phone numbers available with the Udupi Brinjal Growers Association. However, only 307 respondents volunteered to participate in the survey. Thereafter, the researchers discarded three questionnaires that were considered unsuitable for data analysis after the data cleaning process. Hence, 304 responses were included in the final analysis.

Regarding data analysis, factor analysis was performed to validate the statistical significance of variables using factor loadings to label the factors. Further, the study has performed regression analysis with factors as independent variables and consumer preference as the dependent variable. The succeeding section of this research work will present and discuss the results of the study.

Results and discussion

Consumer responses towards Udupi Brinjal consumption trends (Tan et al., 2024) were summarized using the factor analysis technique to identify and label factors. Also, this technique has been useful to overcome the problem of multicollinearity and satisfy the assumption that all the independent variables are statistically independent. A higher KMO value (.702) and a statistically significant Bartlett's test of Sphericity (p-value 0.000) have been observed, indicating that the correlation between variables is statistically significant (refer to Table 1).

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.702
Bartlett's Test of Sphericity	Approx. Chi-Square	223.669
	df.	45
	sig.	.000

Source: Survey results

Table: 1 KMO and Bartlett's test.

Further, the rotation component matrix (refer to Table 2) with factor loads supported the labelling of factors. Factor loadings represent the reliability of the items under study and confirm the measurement fit of the model. In most cases, the study has maintained 0.7 or higher factor loads that explain sufficient variance of a construct.

Components	Factor 1	Factor 2	Factor 3
Fresh Produce	.278	.715	.095
Free from pest attack	.756	.023	.015
Ease of preparation	.253	.701	.056
Allergic reaction	.783	.047	.035
Taste	.082	.736	.349
Organic Produce	.805	.127	.301
Regular consumption	.749	.032	.004
Reasonable size	.400	.093	.442
Label and Grade	.191	.058	.707
Better than other local variety	.267	.257	.702

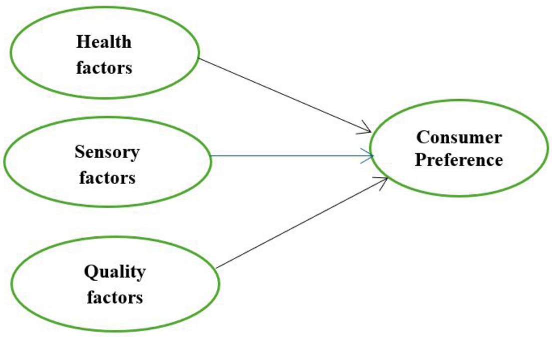
Source: Survey results, Analysis method: Principal Component Analysis & Varimax Rotation Method

Table 2: Rotation component matrix.

In this way, the study has identified three factors using a cut-off of 0.70 to extract variables under each factor. Free from pest attack (0.756), allergic reaction (0.783), organic produce (0.805), and regular consumption (0.749) represent the first factor, which is labelled as 'Health Factor'. Similarly, ease of preparation (0.701) and taste (0.736) represent the second factor labelled as the 'Sensory factor'. Finally, label and grade (0.707) other local varieties (0.702) make the third factor, 'Quality factor'. All factor loading values are > 0.70, which is suitable to perform regression analysis. The profiles of these clusters are explained as:

Consumers of the first cluster are health-conscious and execute their purchase action and consumption responsibly. They seek information on health, nutrition, and ingredients, and are concerned about the impact of the products on their overall well-being. They favour foods that are natural and organic, free from additives and chemicals. Hence, health factors and proposed to impact the nature of their consumption preferences. Sensory attributes like colour, taste, smell, and touch impact the consumer's overall preference and dining experience. For this group, sensory properties are crucial in determining how much consumers choose and enjoy the food that they consume. Regarding the quality factor, it is found that nutritional aspects, variety of vegetables and fruits available for purchase, and grading of the crop by labels determine consumers' purchases. According to Bharadwaj et al. (2024), an alarming decline in the consumption of nutritional quality of foods in India is posing health-related threats to future generations. Hence, it is felt that consumers in the future will be quality-conscious and demand vegetables and fruits, and food without compromising quality.

The proposed relationship (refer to Figure 1) is explained using the following theoretical model:



Source: Developed by authors

Figure 1: Theoretical framework of the study.

The three factors were used to obtain factor scores using Statistical Package for Social Sciences Software (SPSS 27.0) with health factor (HF), sensory factor (SF), and quality factor (QF) as independent variables to predict the dependent variable, consumer preference (CP). For data analysis, the items for Health Factors are coded as HF1- free from pest attack, HF2 - allergic reaction, HF3 - organic produce, and HF4- regular consumption. For Sensory Factors, they are coded as SF1 – ease of preparation, SF2 - taste, and for Quality factors, they are coded as QF1 - labelling and grading, and QF2 – competing variety. The dependent variable, consumer preference, is denoted as CP.

The relationship between the independent variables and the dependent variable is hypothesized as follows:

H1: Health factors positively affect consumer preference towards Udupi Brinjal

H2: Sensory factors positively affect consumer preference towards Udupi Brinjal

H3: Quality factors positively affect consumer preference towards Udupi Brinjal

The hypothesis is tested using a regression equation presented as follows:

CP = f (HF, SF, QF), where HF = Health factor, SF = Sensory factor, and QF = Quality factor.

The regression results (refer to Table 3) indicate that 32.4 per cent of the variance in the dependent variable is predicted by the model, i.e., further, a higher value (f = 17.016) shown by the ANOVA table (refer to Table 4) indicates a good model fit in predicting the dependent variable. Consumer preference level is explained by three factors, i.e., health factors, sensory factors, and quality factors. The coefficients (refer to Table 5) of these three factors (health factor 0.022, sensory factor 0.008, quality factor 0.000) are statistically significant (p values are < 0.05).

Model summary				
Model	R	R Square	Adjusted R Square	Std. Error of the estimate
1	.568	0.324	.311	1.029

Source: Survey results: Dependent variable: Consumer preference (CP), Independent variables – Health factor, Sensory factor and Quality factor

Table 3: Regression model.

Model	Sum of Squares	d.f	Mean square	F	Sig.
	50.368	3	18.123	---	---
Regression	101.632	96	1.067	17.016	.000b
Residual	153.000	99	---	---	---

Source: Survey results

Table 4: ANOVA.

Model	Coefficients					
		Unstandardized coefficients B	Std. error	Standardized coefficients Beta	t	Sig
	Constant	3.590	.104	---	34.718	.000
	Health factor score	.239	.105	.190	2.299	.022
	Sensory factor score	.279	.105	.226	2.637	.008
	Quality factor score	.615	.106	.489	5.829	.000

Source: Survey results – Dependent variable: Consumer preference

Table 5: Factor regression output.

The Unstandardized coefficients obtained (refer to Table 5) after data analysis indicate a positive relationship between health factors (0.23), sensory factors (0.279), quality factors (0.61), and consumer preference. In this regression model, a higher ‘t’ value of the factors (health factor 2.29, sensory factor 2.63, quality factor 5.82) indicates greater predictive power of the factors on the dependent variable (consumer preference). Generally, a ‘t’ value of > 2 for a factor (independent variable) reveals strong evidence to predict the dependent variable. The results also reveal that the quality factor (0.48) is the most important in explaining the variation in Consumer Preference level, as the standardized coefficient of this factor is higher than the values of the sensory factor and health factor. The results are consistent with what is witnessed in the local market as ‘quality’ as a preference indicator for agriculture. In India, the increasing popularity of plant-based diets and eco-friendly consumption is motivating farmers to go organic. Geographical Indications (GIs) assume great importance in predicting consumer preference and purchase intention. The results of this study have significant implications for marketers to design suitable strategies to improve the marketing and sales of Udupi Brinjal. The next section will compare earlier studies on consumer preferences towards agricultural products and Geographical Indications with the present study.

Discussion

The present study aims to analyze the factors influencing consumer preference towards Udupi Brinjal, a locally grown horticulture GI produce. Data analysis has been performed to indicate key empirical findings, which are presented as follows (refer to Table 6).

The results of the study reveal that health factors, sensory factors, and quality factors positively affect consumer preference towards Udupi Brinjal. The p-values (< 0.05) in H1 (.022), H2 (0.008), and H3 (.000) indicate that the findings provide support for the hypothesis proposed by the study. Further, the ‘t’ values of H1 (2.29), H2 (2.63), and H3 (5.82) indicate a positive relationship between the predictor variable and outcome variable. Thus, the study's results are consistent with the proposed hypothesis. In this study, it is observed that the quality factor has a greater influence (‘t’ value of 5.82 > 2) on consumer preference compared to the other two factors.

Quality factors and consumer preference

There are limited studies that have explored the quality dimension of consumer preference in the case of Horticulture GIs. A study has observed that grading, labelling, and packaging that represent quality assume greater importance and should be standardized (Berkile et al., 2019). Consumers’ selection of a product with a GI is prompted by a perception that the label indicates superior quality of the product (Tabanli and Akdogan, 2022). In the case of Udupi Brinjal, the term quality is denoted by Grade labels (based on size). These labels are used to categorise the produce into different grades (Grade A, B & C). It is observed that the local markets are flooded with other varieties of Brinjal, which look like Udupi Brinjal. Hence, releasing Udupi Brinjal produce without labelling may lead to impostering by the local producers and retailers of other competing varieties. Grading and labelling were introduced in 2022 to protect Udupi Brinjal from counterfeiting and impostering by local varieties. A small label sticker carrying the Udupi Brinjal logo is placed on the crop for easy identification on Grade A & Grade B. However, it is felt that packaging should complement grading if the marketing is to be effective. Results of our study indicate that grading and labelling influence consumers' decisions to buy Udupi Brinjal.

Regarding quality parameters, other studies have referred to reliability (Tabanli and Akdogan, 2022; Zhang and Du, 2023) as the key determinant

Hypothesis	Description	‘t’ values	Sig <.05	Remarks
H1	Health factors positively affect consumer preference towards Udupi Brinjal.	2.299	.022	Results are consistent with the hypothesis
H2	Sensory factors positively affect consumer preference towards Udupi Brinjal.	2.637	.008	Results are consistent with the hypothesis
H3	Quality factors positively affect consumer preference towards Udupi Brinjal.	5.829	.000	Results are consistent with the hypothesis

Source: Authors survey results

Table 6: Results of hypothesis testing.

in driving consumer attention towards GI-based local agricultural produce. Focusing on quality and maintaining consistency in quality has significant implications for how consumers prefer (Rao et al., 2024) and accept Udupi Brinjal. Our study emphasises the need for a unique packaging strategy required to position Udupi Brinjal as a 'distinct brand' in the local, domestic markets and international markets. Also, in the quality aspect, it is seen that activities like cleaning, sorting, and grading of produce are important to prevent Udupi Brinjal from going out of favour in the local market. Other studies have reported that consumers are willing to pay a higher price (Kirsten et al., 2017) for premium products with GI tags. Our study emphasizes the need to leverage the GI tag for branding purposes by positioning it as a symbol of quality to increase consumer awareness towards premium agricultural products. Another study (Ihsaniyati et al., 2022) has revealed that promoting quality among consumers may need optimizing collaboration between farmers and relevant associations. There is an information asymmetry (Cardoso et al., 2022) between producers and consumers, as producers are focused on distribution rather than customer centricity. Greater awareness of factors impacting consumer preference will serve as a tool to bridge this gap by adopting the right marketing mix. This disconnect can be resolved by organizing agricultural product exhibitions at regular intervals where producers and consumers meet and interact. Regarding Quality attributes, GIs protect consumers against faking and give a sense of security towards nutrition and information labels (Zhan et al., 2017; Albayram et al., 2014). Post-COVID-19 outbreak, organic fruits and vegetables consumption has witnessed increased consumer attention. Organically grown fruits and vegetables represent the brand of quality (Dangi et al., 2020). Concurring with these views, we believe that the producers should adopt a customer-centric approach and use natural fertilizers and manuring systems to provide organic produce. This approach will be consistent with the responsible consumption Sustainable Development Goals (SDG 12) announced by the United Nations.

Further, we also highlight that the Udupi Brinjal logo, with its local dialect, has limited literary and visual appeal to international consumers. Hence, the Branding strategy for Udupi Brinjal needs serious attention. Post-COVID-19 outbreak, organic fruits and vegetables consumption has witnessed increased consumer attention. Organically grown fruits and vegetables represent the brand of quality (Dangi et al., 2020). The findings of our study present useful input to formulate a strategy for marketers of agricultural GI products in emerging economies to improve the core quality function of labelling, grading, and packaging.

Sensory factors and consumer preference

Regarding sensory attributes, results from an earlier study have stated that sensory factors positively affect the purchase intention of GI products (Moyo et al., 2023). Additionally, other studies have reported that the consumption of traditional foods has a unique sensory experience (Van Der Hoeven et al., 2013; Caskey et al., 2020). Similarly, Dias and Mendes (2018) and Cheung et al. (2021) stated that consumers are interested in sensory-related characteristics in the case of products attributed to a specific region. Another study (Bytyci et al., 2024) revealed that sensory attributes and taste, fragrance, safety, and freshness of fruits motivate the purchase of GI fruits and vegetables. The melt-in-your-mouth texture and soft pulp of Udupi Brinjal make it the homemakers' choice for tasty dishes like Sambar (soup), gojju (salad), and palya (dry dish). Slight bitterness and rawness with subtle sweetness give Udupi Brinjal the edge over other competing local varieties. This study shows that the 'unique taste' of Udupi Brinjal influences consumer preference and loyalty.

Research work by Zhu et al. (2022) has found that brand labels, grade labels, informative labels, and descriptive labels are used as marketing tools based on sensory appeal. This study also revealed that brands' visual appeal and gustatory attributes influence consumer fruit and vegetable buying intentions. In the current study, we have introduced the ease of preparation and taste of value-added products of Udupi Brinjal. Ease of preparation is a significant factor in consumers' buying decisions, as it determines whether food is consumed or not. Consumers are more likely to prefer those vegetables whose dishes or salads are easy to prepare or can be eaten raw. In today's fast-paced life, which is more inclined to food orders from online food distributors, ease of preparation would motivate the Indian kitchen to prepare food at home. Recipes of Udupi Brinjal are easy to prepare. Snacks like Brinjal fritters (fried snack) are easy to prepare. Brinjal salads go well with lunch. Notably, marketers have a strong message to market Udupi Brinjal value-adds using the ease of preparation positioning strategy. It is also felt that consumers, especially women, who cook meals for their family, are driven to buy those items that can be prepared easily with quick hands. Further, vegetables that can be cut easily are preferred over those that are difficult and time-consuming. Udupi Brinjal is known for its unique taste in its preparations like Sambar (curry), salad, and fritters (oil fry). It is felt that marketers must consider the sensory experience of consumers as a powerful tool to increase consumer acceptance of perishable products. Visual merchandising at the store level and visual prompts on online platforms may be a good way to attract consumers to stores and online dashboards using sensory experience.

Health factors and consumer preference

Regarding consumption and health, studies have reported that modern consumer is motivated by sustainable food habits and are not price sensitive (Lizcano-Prada et al., 2024). Consumers continue to adopt new food consumption behaviour with local food being replaced by health-value products (Baliga et al., 2024; Chen and Warden, 2020; Caskey et al., 2020; Cheung et al., 2021). Consumer markets for GIs in India are different from other parts of the world, with the dominance of vegetable and fruit-based GIs in local markets. This research work has highlighted the impact of pest attacks on production, the use of fertilisers and chemicals, and allergic reactions as some of the factors that influence the buying behaviour of health-conscious Indian consumers. Udupi Brinjal consumers prefer products free from pest attack, organically grown produce, and products that are free from allergic reactions. The results of this study reveal a higher loading of organic produce (0.80) for Udupi Brinjal, confirming the consumers' preference for organically grown produce. Increased use of herbal spraying methods to combat pest attacks and focusing on organic growing methods will shift consumers' attention towards purchase.

Consumers who are particular about their health view GI as a tool to execute the correct purchase decision by avoiding the mistake of picking counterfeit and adulterated products. Hence, health marketing campaigns should encourage health-conscious consumption by promoting health-value products guided by the adoption of nutritional labels. The results of our study reveal a lower impact of health on consumer preference compared to sensory attributes and quality factors. This shows the eagerness of consumers to shift consumption preference from sensory attractions towards quality products. Of late, there has been a clear shift in food consumption patterns among Indians towards organic food, which is driven by the need to maintain good health.

Marketing and managerial implications

This study presents managerial implications for marketers. In the case of locally grown GIs (Bytyci et al., 2024), quality indicators accompany key aspects of responsibility and accountability for producers (Cardoso et al., 2022; Yin et al., 2024). Hence, it is felt that producers of Udupi Brinjal should embrace a strong quality culture. First, to promote customer centricity by using quality as a tool, producers can adopt guidelines on packaging as follows: a) follow the criteria for packaging on a) appearance, b) identification, c) instructions for use, d) information about contents, e) carry the brand name, and f) provide sales aid.

Second, producers should focus on brand building in the local market using the Udupi Brinjal logo. The authors concur with the views of the research work done by Uzundumlu and Topcu (2016) that consumer satisfaction can be maximised by augmenting actual product images as a tactic for brand building. For Udupi Brinjal, promoting quality will enable brand building by leveraging tools like labels and logos. Further, other studies have referred to branding as a powerful activity to secure premium prices (Aggarwal et al., 2014). This will ensure that the State government's objective of promoting Geographical Indications (GIs) products is achieved by breaking counterfeiting and imposter-related barriers. The study finds 'ease of Udupi Brinjal preparation' as a strong marketing tool to increase consumption. Items like dishes, pickles, and salads can be prepared very easily, thereby increasing consumption and reducing preparation time and spoilage, and wastage. When preparations are easy to prepare, consumers are more like to consume, increasing consumption and leading to a greater vegetable and fruit intake. Further, the time involved in washing, peeling, and chopping may be a significant deterrent to busy consumers. This section of consumers will welcome the 'easy to prepare' solution. Furthermore, easy-to-prepare recipes will significantly reduce spoilage and wastage, which is a major concern for consumers. Also, the easy-to-prepare solution aligns with consumer habits. Once the consumer gets adjusted to a particular recipe that is easy to prepare, it requires less effort to automate that behaviour, resulting in repeat action. In essence, ease of preparation accompanies convenience, which removes a significant hurdle for consumers. This makes consumption of vegetables and fruits more accessible and actionable for consumers. Marketers should target social media campaigns that promote ease of preparation as a 'convenience tool' to gain consumer attention. In the case of Udupi Brinjal marketing, it is often seen that 'ease of preparation' is veiled under the shadows of 'unique taste'.

The Udupi Brinjal Growers Association can design consumer education programs by combining brand awareness campaigns with health marketing campaigns to eliminate information asymmetry between producers and consumers. Also, awareness and education programs on organic farming should be extended to both farmers and consumers. The Government of India has launched various schemes to promote agriculture, like the 'Paramparagat Krishi Vikas Yojana' (PKVY) and Modern Organic Value Chain Development (MOVCDNER), to encourage organic farming using natural fertilizers. Effective promotion of such schemes by government agencies will go a long way in reducing this information asymmetry between producers and consumers. The final section presents a summary of the aim of this research work, the results achieved, and the limitations of this study. Directions for future research are also presented in the next section.

Conclusion

This research work has analyzed the effects of quality factors, sensory attributes, and health aspects on consumer preference towards Udupi Brinjal. The results indicate that the evidence of the predictive power of quality, sensory, and health-related factors is statistically significant. Given this background, the research has provided marketing and managerial observations to academicians and marketers.

The current research work has a few limitations. The limited sample size of 307 respondents made it difficult to generalise the findings to a larger consumer pool. However, it is felt that the findings of the study are valuable and informative for semi-urban and rural consumer markets in emerging economies. Also, consumers selected for the survey have self-reported their responses. A significant part of the purchase actions is determined by community-driven mechanisms of the local region. Therefore, social factors play an important role in influencing consumers' purchase decisions. Thus, this survey may have included self-desirability bias due to self-perception to a certain extent.

Regarding future research, the area of sustainable community-based enterprises is a good area for further research. Deeper insights on Branding (Aggarwal et al., 2014) place branding (Choo and Kim, 2015) of region-of-origin GI products, Packaging of GIs should attract more literary attention considering the large volume of registered agriculture GIs. Similarly, literary attention towards GI products and consumer preferences and their impact on cultural differences is limited (Wang et al., 2021). For a large country like India, responsible consumption may be driven by quality-related parameters. Additional quality-related parameters like the importance of the consumption of branded products, nutritional value, information, types of labels, and utility are unknown to many consumers. Future research can focus on brand awareness, nutritional aspects of premium products, determinants of consumer acceptance of GIs and elements of a good packaging strategy. Also, studies can determine a suitable pricing strategy for Agriculture GIs, especially in emerging economies that lack minimum support price guidelines. The outcome of the current study is consistent with the Sustainable Development Goals (SDG 2 & 12) that cover promoting Sustainable Agriculture (SDG 2) and Responsible Consumption (SDG 12).

Corresponding author:
Dr. Santhosha Shetty G
Manipal School of Commerce and Economics, Manipal Academy of Higher Education
Manipal, Karnataka 576104, India
Phone: +91-9916821343, E-mail: santhosh.g@manipal.edu

References

[1] Aggarwal, R., Singh, H. and Prashar, S. (2014) "Branding of geographical indications in India: A Paradigm to sustain its premium value", *International Journal of Law and Management,* Vol. 56, No. 6, pp. 431-442. ISSN 1754-2448. DOI 10.1108/ijlma-08-2012-0029.

[2] Ajzen, I. (1991) "The theory of planned behaviour", *Organisational Behaviour and Human Decision Processes*, Vol. 50, No. 2, pp. 179-211. ISSN 1095-9920. DOI 10.1016/0749-5978(91)90020-T.

[3] Albayram, Z., Mattas, K. and Tsakiridou, E. (2014) "Purchasing local and non-local products labelled with geographical indications (GIs)", *Operational Research*, Vol. 14, No. 2, pp. 237-251. ISSN 1109-2858. DOI 10.1007/s12351-014-0154-9.

[4] Artêncio, M. M., de Moura Engracia Giraldi, J. and de Oliveira, J. H. C. (2021) "A cup of black coffee with GI, please! Evidence of Geographical Indication influence on a coffee tasting experiment", *Physiology & Behaviour*, Vol. 245, p. 113671. ISSN 0031-9384. DOI 10.1016/j.physbeh.2021.113671.

[5] Baliga, V., Shetty, S. G., Gil, M. T., Shenoy, R. and Rao, K. (2024) "Analysing Purchase Preference Towards Geographical Indications (GIs) Using Consumer Segmentation Approach", *Agris on-line Papers in Economics and Informatics*, Vol. 16, No. 2, pp. 3-24. ISSN 1804-1930. DOI 10.7160/aol.2024.160201.

[6] Berkile, M. S., More, S. S. and Waghmare, Y. M. (2019) "Marketing cost, marketing margin and price spread of Brinjal in Latur district of Maharashtra state", *Journal of Pharmacognosy and Phytochemistry*, Vol. 8, No. 4, pp. 2326-2328. ISSN 2278-4135.

[7] Bhardwaj, R. L., Parashar, A., Parewa, H. P. and Vyas, L. (2024) "An Alarming Decline in the Nutritional Quality of Foods: The Biggest Challenge for Future Generations' Health", *Foods*, Vol. 13, No. 6, p. 877. ISSN 2304-8158. DOI 10.3390/foods13060877.

[8] Biljana Panin, B. (2022) "Market Perspectives for Serbian PDO Products in the Republic of Serbia", *DETUROPE - The Central European Journal of Tourism and Regional Development,* Vol. 14, No. 1, pp. 128-146. ISSN 1821-2506. DOI 10.32725/det.2022.007.

[9] Bytyçi, P., Kokthi, E., Hasalliu, R., Fetoshi, O., Salihu, L. and Mestani, M. (2024) "Is the local origin of a food product a nexus to better taste or is it just an information bias?", I*nternational Journal of Gastronomy and Food Science*, Vol. 35, p.100877. E-ISSN 1878-450X. DOI 10.1016/j.ijgfs.2024.100877.

[10] Cardoso, V. A., Lourenzani, A. E. B. S., Caldas, M. M., Bernardo, C. H. C. and Bernardo, R. (2022) "The benefits and barriers of geographical indications to producers: A review", *Renewable Agriculture and Food Systems*, Vol. 37, No. 6, pp. 707-719. E-ISSN 1742-1713. DOI 10.1017/S174217052200031X.

[11] Caskey, D., Chen, J. F. and Warden, C. A. (2020) "Surfacing consumer psychosensory perceptions of a nonendemic food: The case of coffee in a tea culture", *Journal of Sensory Studies*, Vol. 36, No. 1. ISSN 1745-459X. DOI 10.1111/joss.12625.

[12] Chazdon, S. A. and Lott, S. (2010) "Ready for engagement: using key informant interviews to measure community social capacity", *Community Development*, Vol. 41, No. 2, pp. 156-175. ISSN 1468-2656. DOI 10.1080/15575331003646173.

[13] Chen, J. F. and Warden, C. A. (2020) "Traditional agriculture product repositioning: goat milk in Taiwan", *International Journal of Agriculture Innovation, Technology and Globalisation*, Vol. 1, No. 3, pp 238-252. ISSN 2516-1970. DOI 10.1504/ijaitg.2020.106008.

[14] Chen, N.-H. (2021) "Geographical indication labelling of food and behavioral intentions", *British Food Journal*, Vol. 123, No.12, pp. 4093-4115. ISSN 1758-4108. DOI 10.1108/bfj-06-2020-0552.

[15] Cheung, J. T. H., Lok, J., Gietel-Basten, S. and Koh, K. (2021) "The Food Environments of Fruit and Vegetable Consumption in East and Southeast Asia: A Systematic Review", *Nutrients*, Vol. 13, No. 1, p. 148. ISSN 2072-6643. DOI 10.3390/nu13010148.

[16] Choo, S. and Kim, H. (2015) "The Brand Value of Place Names: Topics in Economic Geography", *Journal of the Economic Geographic Society of Korea*, Vol. 18, No. 4. ISSN 1226-8968. DOI 10.23841/egsk.2015.18.4.431.

[17] Choudhury, S., Shankar, B., Aleksandrowicz, L., Tak, M., Green, R., Harris, F., Scheelbeek, P. and Dangour, A. (2020) "What underlies inadequate and unequal fruit and vegetable consumption in India? An exploratory analysis", *Global Food Security*, Vol. 24, p.100332. ISSN 2211-9124. DOI 10.1016/j.gfs.2019.100332.

[18] D'souza, D. J. and Joshi, H. G. (2020) "Exploring the Consumers' Willingness of Using E-Commerce to Purchase Geographical Indication-Based Crops, a Case Study of Udupi Jasmine", *Agris on-line Papers in Economics and Informatics*. Vol. 12, No. 2, pp. 63-69. ISSN 1804-1930. DOI 10.7160/aol.2020.120206.

[19] Dangi, N., Gupta, S.K. and Narula, S. A. (2020) "Consumer buying behaviour and purchase intention of organic food: a conceptual framework", *Management of Environmental Quality: An International Journal*, Vol. 31, No. 6, pp. 1515-1530. E-ISSN 1758-6119. DOI 10.1108/meq-01-2020-0014.

[20] Dias, C. and Mendes, L. (2018) "Protected Designation of Origin (PDO), Protected Geographical Indication (PGI) and Traditional Speciality Guaranteed (TSG): A bibliometric analysis", *Food Research International*, Vol. 103, pp. 492-508. ISSN 1873-7149. DOI 10.1016/j.foodres.2017.09.059.

[21] Durand, C. and Fournier, S. (2017) "Can Geographical Indications Modernize Indonesian and Vietnamese Agriculture? Analysing the Role of National and Local Governments and Producers' Strategies", *World Development*, Vol. 98, No. 98, pp. 93-104. ISSN 1873-5991. DOI 10.1016/j.worlddev.2015.11.022.

[22] Gautam, M., Kafle, S., Regm, B., Thapa, G. and Paudel, S. (2019) "Management of Brinjal Fruit and Shoot Borer (Leucinodes orbonalis Guenee) in Nepal", *Acta Scientific Agriculture*, Vol. 3, No. 9, pp. 88-195. ISSN 2581-365X. DOI 10.31080/asag.2019.03.0632.

[23] Handy, F., Cnaan, R. A., Bhat, G. and Meijs, L. C. P. M. (2011) "Jasmine growers of coastal Karnataka: Grassroots sustainable community-based enterprise in India", *Entrepreneurship & Regional Development*, Vol. 23, No. 5-6, pp. 405-417. ISSN 0898-5626. DOI 10.1080/08985626.2011.580166.

[24] Hennink, M., Hutter, I. and Bailey, A. (2020) "*Qualitative Research Methods*", 2nd ed. London: Sage Publications. ISSN 1468-7941.

[25] Herath, U. (2019) "Consumer Behaviour and Attitudes in Purchasing Vegetables", *Agricultural Research & Technology*, Vol. 20, No. 2. ISSN 2471-6774. DOI 10.19080/ARTOAJ.2019.20.556123.

[26] Ihsaniyati, H., Setyowati, N. and Pardono (2022) "Factors Motivating the Adoption of Geographical Indication-Based Quality Standards among Robusta Coffee Farmers in Indonesia", *International Journal of Business and Society*, Vol. 23, No. 1, pp. 207-225. ISSN 1511-6670. DOI 10.33736/ijbs.4609.2022.

[27] Kaliji, S. A., Mojaverian, S. M., Amirnejad, H. and Canavari, M. (2019) "Factors Affecting Consumers' Dairy Products Preferences", *Agris on-line Papers in Economics and Informatics*, Vol. 11, No. 2, pp. 3-11. ISSN 1804-1930. DOI 10.7160/aol.2019.110201.

[28] Kirsten, J. F., Vermeulen, H., van Zyl, K., du Rand,G., du Plessis, H. and Weissnar, T. (2017) "Do South African Consumers have an Appetite for an Origin-based Certification System for Meat Products? A Synthesis of Studies on Perceptions, Preferences and Experiments", *International Journal on Food System Dynamics*, Vol. 8, No. 1, pp. 54-71. ISSN 1869-6945. DOI 10.18461/ijfsd.v8i1.815.

[29] Kokthi, E. and Kruja, D. (2016) "Consumer Expectations for Geographical Origin: Eliciting Willingness to Pay (WTP) Using the Disconfirmation of Expectation Theory (EDT)", *Journal of Food Products Marketing*, Vol. 23, No. 8, pp. 873-889. ISSN 1045-4446. DOI 10.1080/10454446.2017.1244794.

[30] Košičiarová, I., Kádeková, Z., Holotová, M., Kubicová, Ľ. and Predanocyová, K. (2020) "Consumer Preferences in the Content of Loyalty to the Yoghurt Brand", *Agris on-line Papers in Economics and Informatics*, Vol. 12, No. 1, pp. 37-48. ISSN 1804-1930. DOI 10.7160/aol.2020.120104.

[31] Lambarraa-Lehnhardt, F., Ihle, R. and Mhaouch, K. (2021) "Geographical indications for supporting rural development in the context of the Green Morocco Plan: Oasis dates", *Agricultural Economics (Zemědělská ekonomika)*, Vol. 67, No. 2, pp. 70-79. ISSN 1805-9295. DOI 10.17221/226/2020-agricecon.

[32] Lee, J. Y., Pavasopon, N., Napasintuwong, O. and Nayga, R. M. (2020) "Consumers' Valuation of Geographical Indication-Labelled Food: The Case of Hom Mali Rice in Bangkok", *Asian Economic Journal*, Vol. 34, No. 1, pp. 79-96. ISSN 1351-3958. DOI 10.1111/asej.12196.

[33] Lizcano-Prada, J., Maestre-Matos, M., Mesias, F. J., Lami, O., Giray, H., Özçiçek Dölekoğlu, C., Abdoulame Bamoi, A. G. and Martínez-Carrasco, F. (2024) "Does Consumers' Cultural Background Affect How They Perceive and Engage in Food Sustainability? A Cross-Cultural Study", *Foods*, Vol. 13, No. 2, pp. 311. ISSN 2304-8158. DOI 10.3390/foods13020311.

[34] Moyo, A., Amoah, F. and van Eyk, M. (2023) "Consumer behaviour research on traditional foods in Africa: A scoping review", *Cogent Business & Management*, Vol. 10, No. 2, ISSN 2331-1975. DOI 10.1080/23311975.2023.2213532.

[35] Nadeeshani, H., Samarasinghe, G., Wimalasiri, S., Silva, R., Hunter, D. and Madhujith, T. (2021) "Comparative analysis of the nutritional profiles of selected Solanum species grown in Sri Lanka", *Journal of Food Composition and Analysis*, Vol. 99, p.103847. ISSN 0889-1575. DOI 10.1016/j.jfca.2021.103847.

[36] Naeem, M. Y. and Ugur, S. (2019) "Nutritional Content and Health Benefits of Eggplant", *Turkish Journal of Agriculture, Food Science and Technology*", Vol. 7. No. Sp3, pp. 31-36. E-ISSN 2148-127X. DOI 10.24925/turjaf.v7isp3.31-36.3146.

[37] Nandi, L. L., Saha, P., Jaiswal, S., Lyngdoh, Y. A., Behera, T. K., Pan, R. S., Munshi, A. D., Saha, N. D., Hossain, F., Srivastava, M. and Tomar, B. S. (2021) "Bioactive compounds, antioxidant activity and element content variation in indigenous and exotic Solanum sp. and their suitability in the recommended daily diet", *Scientia Horticulturae*, Vol. 287, pp. 110232. ISSN 0304-4238. DOI 10.1016/j.scienta.2021.110232.

[38] Ortega, D. L., Hong, S. J., Wang, H. H. and Wu, L. (2016) "Emerging markets for imported beef in China: Results from a consumer choice experiment in Beijing", *Meat Science*, Vol. 121, pp. 317-323. ISSN 1873-4138. DOI 10.1016/j.meatsci.2016.06.032.

[39] Palaniappan, V. and Radhakrishnan, B. (2020) "Factors Influencing Consumers' Purchasing Behaviour on Exotic Vegetables in Coimbatore City", *Asian Journal of Agricultural Extension, Economics & Sociology*, Vol. 38, pp. 122-133. ISSN 2320-7027. DOI 10.9734/ajaees/2020/v38i1230502.

[40] Peredo, A. M. and Chrisman, J. J. (2006) "Toward a Theory of Community-Based Enterprise", *Academy of Management Review*, Vol. 31, No. 2, pp. 309-328. E-ISSN 1930-3807. DOI 10.5465/amr.2006.20208683.

[41] Peschard, K. E. (2022) "*Seed Activism*", The MIT Press eBooks. ISBN 9780262544641. DOI 10.7551/mitpress/14484.001.0001.

[42] Rao, F., Chenguang, L., Wang, L. and Gao, Z. (2024) "Chinese consumer preference for beef with Geographical indications and other attributes", *Meat Science*, Vol. 212, pp. 109475. ISSN 1873-4138. DOI 10.1016/j.meatsci.2024.109475.

[43] Raut, M., Khobarakar, V., Patil, A., Kankal, A. and Dangore, A. (2020) "Marketing of Brinjal in Akola District", *Journal of Pharmacognosy and Phytochemistry*, Vol. 9, No. 5, pp. 494-497. ISSN 2349-8234.

[44] Rejeki, S., Andriatmoko, N. D. and Toiba, H. (2021) "Factors Affecting the Intention to Purchase Organic Vegetables with Theory Planned Behaviour Approach", *Agricultural Social Economic Journal*, Vol. 21, No. 2, pp. 103-110. ISSN 1412-1425. DOI 10.21776/ub.agrise.2021.021.2.3.

[45] Tabanli, S. M. and Akdoğan, M. Ş. (2022) "Young Consumers' Attitudes Together with Geo-Labelled Food Products and the Effect of Ethnocentric Perceptions on Purchasing Intention", *Turkish Journal of Marketing*, Vol. 7, No. 3, pp. 105-124. ISSN 2458-9748. DOI 10.30685/tujom.v7i3.157.

[46] Tan, A., Hashim, S. B., Zuo, J. and Cheng, J. (2024) "Consumer responses and determinants in geographical indications agricultural product consumption: A ten-year systematic review", *F1000Research*, Vol. 13, pp. 1410. ISSN 2046-1402. DOI 10.12688/f1000research.158225.1.

[47] Toklu, I. T., Kucuk, H. O. and Toklu, A. T. (2020) "The importance of extrinsic cues in deciding to purchase meat products: A conjoint analysis on Muslim consumers", *South African Journal of Business Management*, Vol. 51, No. 1, ISSN 2078-5976. DOI 10.4102/sajbm.v51i1.1986.

[48] Užar, D., Dunđerski, D. and Pejanović, V. (2022) "Consumers' intention to buy cheeses with geographical indications: The case of Serbia", *Ekonomika poljoprivrede*, Vol. 69, No. 3, pp. 819-832. ISSN 0352-3462. DOI 10.5937/ekopolj2203819u.

[49] Užar, D. and Filipović, J. (2023) "Determinants of Consumer Purchase Intention Towards Cheeses with Geographical Indication in a Developing Country: Extending the Theory of Planned Behaviour", *Tržište*, Vol. 35, No. 2, pp. 183-204. ISSN 0353-4790. DOI 10.22598/mt/2023.35.2.183.

[50] Uzundumlu, A. S. and Topcu, Y. (2016) "Determining Turkish consumers' consumption satisfaction with Erzurum Civil cheese", *British Food Journal*, Vol. 118, No. 4, pp. 896-914. ISSN 1758-4108. DOI 10.1108/bfj-03-2015-0113.

[51] van der Hoeven, M., Osei, J., Greeff, M., Kruger, A., Faber, M. and Smuts, C. M. (2013) "Indigenous and traditional plants: South African parents' knowledge, perceptions and uses and their children's sensory acceptance", *Journal of Ethnobiology and Ethnomedicine*, Vol. 9, No. 1, p. 78. ISSN 1746-4269. DOI 10.1186/1746-4269-9-78.

[52] Vijayan, D., Ushadevi, K. N. and Krishna, R. (2019) "Determinants of Consumer Behaviour Towards Organic Vegetables", *Commerce Spectrum*, Vol. 7, No. 2, pp. 28-34. ISSN 2321-371X.

[53] Vinayan, S. (2017) "Geographical Indications in India: Issues and Challenges-An Overview", *The Journal of World Intellectual Property*, Vol. 20, No. 3-4, pp. 119-132. ISSN 1422-2213. DOI 10.1111/jwip.12076.

[54] Wang, L., Wang, J. and Huo, X. (2019) "Consumer's Willingness to Pay a Premium for Organic Fruits in China: A Double-Hurdle Analysis", *International Journal of Environmental Research and Public Health*, Vol. 16, No. 1, p. 126. ISSN 1660-4601. DOI 10.3390/ijerph16010126.

[55] Wang, T.-S., Liang, A. R.-D., Ko, C.-C. and Lin, J.-H. (2021) "The importance of region of origin and geographical labelling for tea consumers: the moderating effect of traditional tea processing method and tea prices", *Asia Pacific Journal of Marketing and Logistics*, Vol. 34, No. 6, pp. 1158-1177. ISSN 1355-5855. DOI 10.1108/apjml-02-2021-0121.

[56] Whittaker, R. (2012) "Issues in mHealth: Findings from Key Informant Interviews", *Journal of Medical Internet Research*, Vol. 14, No. 5, pp. 129. ISSN 1438-8871. DOI 10.2196/jmir.1989.

[57] Yin, X., Li, J., Wu, J., Cao, R., Xin, S. and Liu, J. (2024) "Impacts of Geographical Indications on Agricultural Growth and Farmers' Income in Rural China", *Agriculture*, Vol. 14, No. 1, p. 113. ISSN 2077-0472. DOI 10.3390/agriculture14010113.

[58] Yu, H., Jiang, Y., Sun, Y., Ding, Y. and Alita, L. (2024) "Geographical indications and organic labels: complements or substitutes? - the case of online rice consumption in China", *Applied Economics*, Vol. 57, No. 26, pp. 1-15. ISSN 1466-4283 DOI 10.1080/00036846.2024.2337814.

[59] Zhan, H., Liu, S. and Yu, J. (2017) "Research on factors influencing consumers' loyalty towards geographical indication products based on grey incidence analysis", *Grey Systems: Theory and Application*, Vol. 7, No. 3, pp. 397-407. ISSN 2043-9377. DOI 10.1108/gs-10-2016-0037.

[60] Zhang, S. and Du, B. (2023) "Tracing or not: How can the supplier of geographical indication products benefit from different traceability strategies?", *Computers & Industrial Engineering*, Vol. 184, p. 109516. ISSN 0360-8352. DOI 10.1016/j.cie.2023.109516.

[61] Zhang, S., Wen, X., Sun, Y. and Xiong, Y. (2024) "Impact of agricultural product brands and agricultural industry agglomeration on agricultural carbon emissions", *Journal of Environmental Management*, Vol. 369, p. 122238. ISSN 0301-4797. DOI 10.1016/j.jenvman.2024.122238.

[62] Zhu, Z., Shen, Q. and Gao, Z. (2022) "Consumer choices in agricultural markets with multitier collective labels and private brands", *Agribusiness*, Vol. 295, pp. 126443. ISSN 1520-6297. DOI 10.1002/agr.21747.

Volume XVII Number 4, 2025

Leveraging Deep Learning for Early Detection and Diagnosis of Wheat Diseases: Challenges and Innovations

Chemesse Ennehar Bencheriet[1], Hala Hamouchi[2], Mohamed Islem Hadri[2]

[1] LAIG Laboratory, University 08 May 1945, Guelma, Algeria

[2] Computer Science Department, University 08 May 1945, Guelma, Algeria

Abstract

This research introduces a deep learning system for the early identification and categorization of wheat illnesses, with the objective of optimizing crop health and promoting agricultural sustainability. Results in up to high classification accuracy for brown rust, yellow rust, leaf rust, and septoria. The combination of artificial intelligence (AI) with image processing methodologies such as rescaling and augmentation allows the system to accurately classify wheat crops that are well or unhealthy. The presented system is of great interest for precision agriculture, providing an affordable means to reduce the application of pesticides and encourage sustainable agricultural practices. Ongoing research involves linking this diagnostic platform with drone technology to facilitate on-demand, point-by-point disease surveillance and monitoring across large areas, further extending the platform's applicability in field applications for food securit.

Keywords

Wheat disease detection, deep learning, convolutional neural networks (CNNs).

Bencheriet, C. E., Hamouchi, H. and Hadri, M. I. (2025) "Leveraging Deep Learning for Early Detection and Diagnosis of Wheat Diseases: Challenges and Innovations", *AGRIS on-line Papers in Economics and Informatics*, Vol. 17, No. 4, pp. 27-36. ISSN 1804-1930. DOI 10.7160/aol.2025.170403.

Introduction

Wheat production is a strategic priority for food security in many agricultural locations worldwide, where it plays a vi tal role as a staple food. Disease conditions related to wheat, mainly rusts and septoria, are a major pestilence for farmers resulting in yield losses and deterioration in quality. Disease control procedures are often late in emphasizing such condi tions, resulting in a significant hit on the quantity and quality of yield. Reducing pesticide usage has become a major priority in modern agriculture due to its environmental and health impacts. Early and accurate disease detection systems can significantly minimize unnecessary chemical applications, contributing to more sustainable agricultural practices. In addition, the growing pressure of climate change increases the spread and severity of plant diseases, making efficient detection tools essential. The lag time creates tremendous losses, particularly in locations that are highly dependent on wheat as a primary food and a source of livelihood. The rapid advancements in artificial intelligence technologies have produced CNN based approaches as exciting and promising techniques in the detection and diagnosis of wheat diseases. However promising, such pieces of equipment are still far from being adopted in practical farming. This is probably one of the most crit ical questions: how to correctly introduce these tools into farming practices so that each was able to properly fit into the specific features of unsustainable disease management in heterogeneously practicing agricultural systems with varying resources and constraints?

"Reducing the pesticide applications in agriculture are of paramount importance and it is suggested an inexpensive op tical device to be implemented in the remote disease detec tion in crop fields by analyzing the canopy reflectance. Yellow rust detection in wheat was accomplished using in-field spectral images by recording in-field spectral images with a spectrograph attached on the spray boom height. To take into account reflectance and light intensity differences, a normalization technique was used, and a dataset referred "leaf spectra" was constituted. Disease detection algorithms were trained using neural network, in particular, multilayer per ceptron (MLP). Classification accuracy went up dramatically, from 95% to more than 99%. This demonstrates the promise of suchanoptical device in the early diagnosis of disease with the result of decreased pesticide applications and adoption of sustainable agricultural methods" (Moshou et al., 2004).

"Modern computer vision methods are proposed are pre sented for the detection of various wheat leaf diseases, specif ically yellow spots, yellow rust, and brown rust. Neural network architectures were applied to reach a high accuracy from 0.95 to 0.99 on specific disease categories, thereby equalling the performance of skilled phytopathologists. A new way of doing multilabel classification was proposed to enable the recognition of more than one disease from the same image. An image dataset for diseased wheat leaves manifesting different diseases was developed and utilized in the model training to encourage the creation of high precision disease detection systems. Preprocessing was applied to enhance the performance of the model: rotation, flipping, and image normalization. GoogleNet was chosen for multilabel classification because it has a very lightweight neural network architecture with great performance and effi ciency. The accuracy, precision, recall, and F1-score metrics were used in the assessment of the performance ofthis model; hence, this showed high accurate rates over individual disease classes. Its light weight makes the model deployable on mobile devices to quickly and automatically diagnose wheat crop diseases. The proposed method is scalable and possible in practical agricultural scenarios and holds great potential for application on mobile devices. The results show that the proposed models have achieved high accuracies in the diag nosis of diseases in wheat leaves, which further demonstrates the effectiveness of the models in automated agricultural disease diagnosis" (Arinichev et al., 2022).

"The objective is to develop a deep transfer learning model for disease identification in wheat crops using artificial in telligence (AI). The study utilized the WheatRust21 dataset, which was collected on the spot and included stripe, leaf, and stem rust occurrences in wheat farming. The work mainly fo cused on the use of Convolutional Neural Networks (CNN) and EfficientNet architectures for recognizing wheat rust diseases. Further, the study tried to make the model deployable on mobile devices to facilitate real-time, image-based iden tification of wheat diseases on the farm. Thus, proving that AI-based approaches were really effective in this task, a deep transfer learning model with a 99.35% accuracy level was obtained in disease identification. Disease detection perfor mance wasimproved further by using the CNN and Efficient Net architectures. Further, due to its compatibility with the use of mobile devices in real-time on-site identification, this model provides a great opportunity for easy disease manage ment in wheat crops. This research has shown the strength of the artificial intelligence-driven schemes in dealing with agri cultural diseases, especially in the step of disease identifica tion, and pointed out the possibility of applying these models on a mobile platform for real applications in the field" (Nigam et al., 2023).

"Wheat disease problem is discussed, mainly focusing on the great importance of wheat leaf diseases' impact on agricul tural yield and food safety. A new deep learning model, called RFE-CNN, has been developed based on Residual Channel Attention Blocks (RCABs), Feature Boosters (FBs) and Embedding-based Metric Learning (EML) with the LWDCD 2020 dataset and Convolutional Neural Networks (CNN) to get accurate disease detection in wheat crops. The study uses parallel CNNs, RCABs, and FBs to extract relevant features and enhance classification accuracy. The to tal classification accuracy achieved was 98.83%, whereas the highest testing accuracy achieved was 99.95%. The average accuracy score was 99.50%. RFE-CNN is superior to tradi tional CNN models in accuracy, efficiency, and adaptability and proves to be better for the task of disease identification; further research should focus on strengthening disease recog nition over the ecological regions and varieties by using hyperspectral imaging technology" (Xu et al., 2023).

"The ongoing threat to the cultivation of wheat around the world by wheat rust pathogens is discussed, underlining the economic losses running into billions of dollars annually. VG16 and Capsule Networks have been tried out with an ac curacy score of 93%. Rust fungi rely on living host cells for their growth and reproduction activities, which explains chal lenges related to controlling their spread. Continuous emer gence of new rust races is a significant hindrance to achiev ing genetic resistance in wheat varieties. Scientific studies show that race-specific genes in wheat code for NBS-LRR proteins, hence providing useful insight into the genetic ways through which resistance to rust pathogens is achieved. Man agement strategies for diseases of wheat include fungicide application, cultural practices, and optimization of planting dates to reduce the impact ofrust pathogens and other diseases on wheat yield." (Rathore and Prasad, 2022).

"Identification and control of wheat diseases, specifically yel low rust, Septoria tritici blotch, brown rust, and mildew, are critical since they cause severe damage in crop yield and quality. The challenge arises because, at some stages in their development cycles, the aforementioned diseases tend to exhibit identical appearances. To handle this challenge, deep learning approaches—particularly convolutional neu ral networks (CNNs)—are investigated in automatically detecting and classifying wheat diseases using image analysis. The images used in this research were taken in 2019 from several sites across the UK and Ireland under natural and glasshouse conditions, showing wheat leaves with variously visible symptoms of diseases and healthy leaves. The dataset used in the research had more than 19,000 images represent ing five different classes: septoria, yellow rust, brown rust, mildew, and healthy. In the present study, a CNN called CerealConv is developed and trained to automatically detect and classify wheat diseases. This CNN had 13 convolutional layers with batch normalization, max pooling, and dropout. CerealConv achieved a classification accuracy of over 97% on a test set of five well-known disease classes. Compared against manual classifications by five specialist pathologists, CerealConv outperformed the pathologists with its accuracy 2% higher than the best-performing pathologist. The convo lutional neural network also classified an abridged data set of 999 images both faster and more accurately than the pathologists. To ensure that CerealConv relies on relevant informa tion to perform its classification task, masked images were used. The large drop in classification accuracy upon blocking key image components indicates that it really depends on valid information to perform its task effectively. Overall, this study shows that deep learning methods, specially the developed CNN model—CerealConv—handle real field condition images ofwheatdiseases effectively and at least as well as expert pathologists in disease identification and classification" (Long et al., 2023).

"It emphasizes the importance of wheat in ethiopia being the second most consumed grain crop, contributing to 14% calo rie consumption. Ethiopia´s wheat production is almost ex clusively for subsistence, produced by small holder farmers. For this reason, deep learning-based classification systems to improve early identification of wheat diseases can provide an important contribution in the disease management. Wheat farmers in Ethiopian experience market constraints due to lack of access for timely information and insufficient mar ket linkage which lowers productivity and profitability. The genetic variation in wheat species is important to combat diseases; hence there are breeding programs and scope for genetics. The VGG19 model seems to be able to classify wheat diseases well, and we have seen it performing relatively well for the disease detection task. An automated system for this identification process will help in reducing the large amounts of yield losses and further back Ethiopia's agricultural sector to increase productivity and food security" (Aboneh et al., 2021).

"This includes the million-dollar losses from diseases in wheat, a major worldwide grain of 4.5 and reasons for crop spoilage Optimizing crop yield by identifying diseases and classes of the wheat disease with high accuracy—automatic classification for wheat disease detection in deep learning ar chitecture Even though the deep learning classifiers are pow erful, they might suffer from overfitting and can require large datasets computational resources. Finding: Transfer learning as a means of boosting classification performance, especially when having constraints on data and resources making use of pre-trained models. The model proposed is based on the VGG16 architecture, performs well in classifying wheat diseases, showing promise to enhance crop management practice. The newly devised deep learning model exhibited high performance, acquiring a test accuracy of 97.88% on discrim inating ten different wheat diseases which were tested Such a high accuracy is indicative of how well it can determine and classify wheat diseases, which in turn will allow for prompt management practices to be utilized enough to prevent crop yield reductions. Moreover, while the VGG16 architecture is applied to illustrate that our model can achieve high accuracy in disease classification tasks; this exhausts its application capabilities—from a practical perspective—and practically it could augment field-based strategies and processes on crop management ensuring food security" (Jadhav et al., 2021).

"Thefocus is onusing convolutional neural networks (CNNs) to identify soybean leaf diseases with the help of AlexNet and GoogleNet models. The used dataset includes soybean leaf images which were gathered from farms of Kolhapur block in Maharashtra state, India. It contains

649 images for training the AlexNet model, and 550 images for training the GoogleNet model, which were divided into four cat egories: bacterial blight, brown spot, frogeye leaf spot, and healthy leaves. Both models achieved accuracy of over 95% and above the designed targets." "Outdated recognition systems are the main limitation of farmers in understanding plants which makes them suscep tible to diseases that seriously affect agricultural productiv ity. even though farmers are knowledgeable locally, there are restrictions in terms of sharing such capabilities due to absence of platforms. Further, the constant decrease in agricultural yields was never disconnected with diseases, farming habits or weak location understanding. To address this, the information provided to the farmers is sourced through other farmers, experts and agricultural stakeholders, which informs the farmers about prevalent diseases so they can take action according to the situation. Machine Learning (ML) as a tool is always greatly appreciated which however, suffers from interfacing with static data and maltreating nuances of lo cal knowledge from experts, leading to poor generalizations across regions. In order to address this problem, an effort was made in crowdsourcing images and symptom based data for training purpose. For wheat disease recognition and clas sification, a new framework integrating Decision Trees (DT) with deep learning models is presented. The Decision Tree, following an expert validation, City University achieved a 28.5% increase in accuracy, from 51.6% to 80.1%, and an increase of 4.3% in CNN, resulting in a final accuracy of 97.2%, with most of the decision rules integrated into a decision support system for managing wheat diseases" (Niedbała et al., 2019).

"The impacte of climate change-on-wheat production scenar ios for Iran were assessed with the CERES-Wheat model, and the new wheat disease dataset (WDD2017) was built to provide a more precise diagnosis of plant diseases by the DMIL-WDDS model, achieving mean accuracies of 97.95% ...... and 95.12% via 5-fold cross-validation. The CERES-Wheat model predicts wheat yield with a clear correlation between the predicted and measured values, thereby confirming its usefulness in simulating the climate change problem in the f ield of wheat production. In addition, the DMIL-WDDS could distinguish diseases better than the traditional CNN approaches since climate change showed a negative impact on wheat yield and biomass, specifically due to temperature rise, stressing the critical importance of predicting climate change impact for sustainable agriculture in Iran" (Haider et al., 2021).

"The objective is to classify Fusarium Head Blight disease in wheat by employing deep neural networks on hyperspec tral imaging data collected in the field. The experiment took place from April 29 to May 15, 2017, under varying wind conditions and relative humidity and temperature. Ninety samples of wheat ears were segregated into 10 segments in turn, and a Deep Convolutional Neural Network (CNN) was employed for analysis. Metrics such as precision, recall, and F1 score, among others, were used to evaluate the performance of the model. The experiments maintain that deep neural networks can positively identify Fusarium Head Blight in wheat. This work proposes hybrid neural networks as a suitable approach to disease diagnosis and upgrades in agri cultural disease management" (Lu et al., 2017).

Materials and methods

System architecture

The proposed model for wheat disease detection is based on a master-slave sequential architecture proposed in (Bencheriet and Bencheriet, 2023), composed of three distinct modules (Figure 1):

Module 1: Data preprocessing

Module 2: The Master (Disease detection network)

Module 3: The Slave (Disease diagnosis network)

Remarque: All models were implemented, trained, and evaluated by the authors using Python in 2025.

Data preprocessing

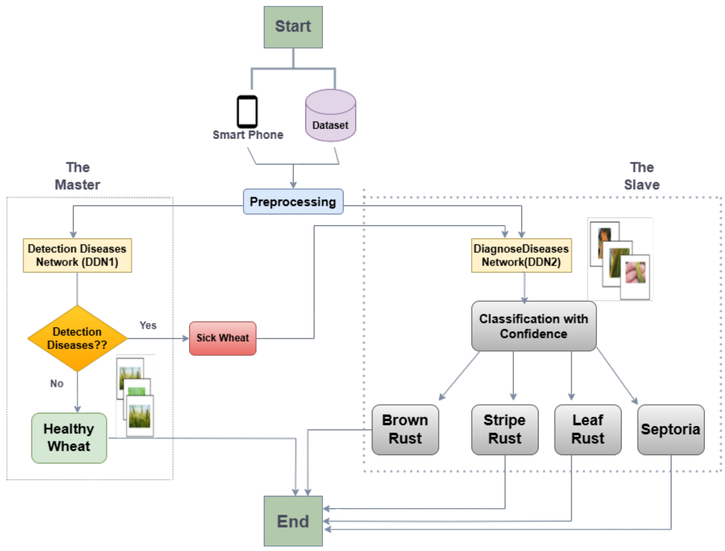
Image preprocessing starts by captur ing images of wheat ears, and then these images are resized to a uniform size of 255x255 pixels. This resizing enables the images to be standardized so that they will fit into the neural network. After the images are resized, the images are fed to the corresponding network for disease detection and diagnosis.

Image resizing

Resizing includes changes of the image sizes (255x255 pixels) to guarantee the uniformity of the input for mat. In this step, the image quality and resolution are kept in law, and the performance of the model is optimized.

Split data

The data is divided as follows: 80% for training, which is used to train the model; 10% for validation, which is used for hyper-parameter tuning and model evaluation during training; and 10% for testing, which is used to evaluate the f inal performance of the trained model (Figure 2, 3).



Source: Authors' own design (2025)

Figure 1: Architecture of our system.

Class	Train 80%	Validation 10%	Test 10%
Healthy	588	74	73
Sick	2215	277	276

Source: Authors' dataset partitioning (2025)

Figure 2: Split dataset for DDN1 model.

Class	Train	Validation	Test
Septoria	166	21	21
Leaf rust	20	8	8
Brown rust	902	113	113
Stripe rust	1116	140	140

Source: Authors' dataset partitioning (2025)

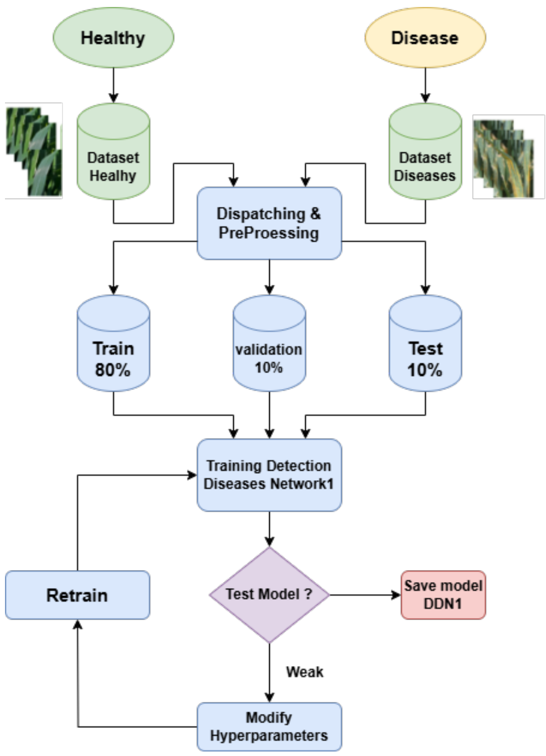
Figure 3: Split dataset for DDN2 model.

The Master Network (DDN1)

Initial classification of wheat ear by Master Network (DDN1) is carried out to diagnose wheat ear as it is healthy or not healthy. The model is based on a deep convolutional network, in which the images are processed with a succession of convolution and pooling lay ers in order to extract relevant features. After the image is preprocessed, the image is then redirected to the disease detection network for subsequent analysis (Figure 4).

Data training

Data are divided into three sets: 80/10/10% for training, validation and testing, respectively. Parameters of the training procedure are optimized and results are classified on the basis of their accuracy (high or low).

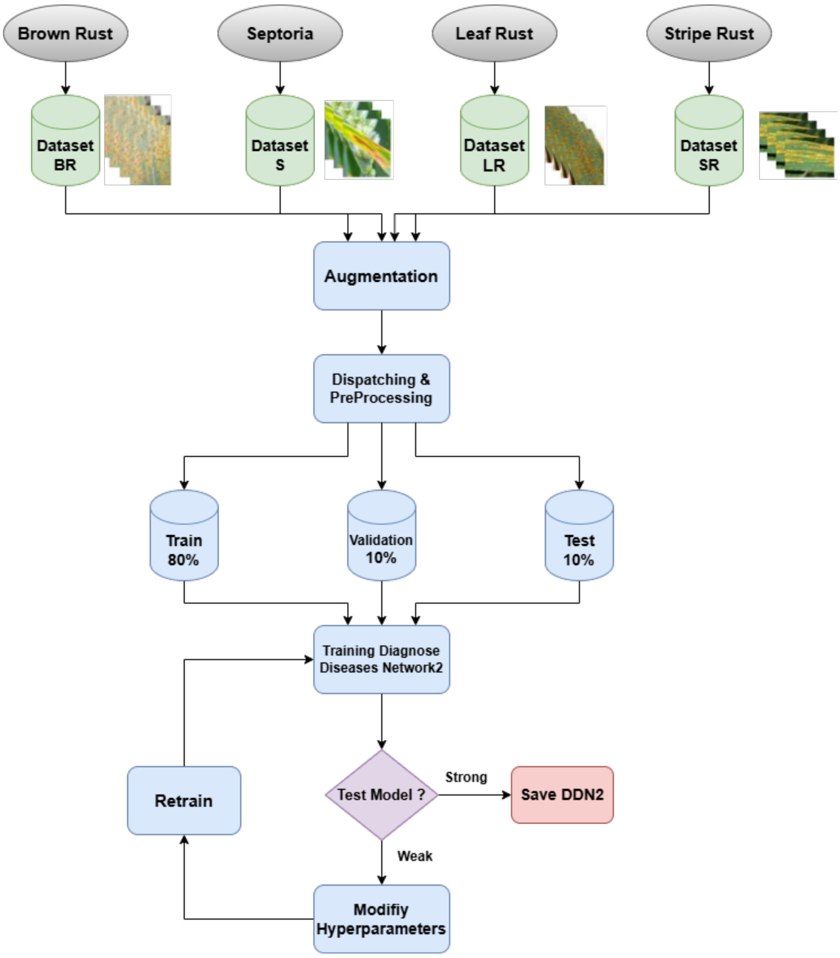


Source: Authors' own model training results (2025)

Figure 4: Train Detection Disease Network (DDN1).

The Slave Network (DDN2)

Once the disease has been de tected by the Master Network, the Slave Network takes over to classify the specific disease into four classes: yellow rust, brown rust, spot rust and healthy. This network uses addi tional layers to refine the classification and provide an accurate diagnosis (Figure 5).



Source: Authors' own model training results (2025)

Figure 5: Train Diagnose Disease Network (DDN2).

Image augmentation

Image augmentation is applied to enhance the diversity of the dataset and generalize the model. These modifications include rotation, flipping, and scaling to generate new, varied images from the same samples.

Data training

The training of this network is very similar to that of Master Network, where the data is split and trained using optimal hyper-parameters. The model is tested for its ability of classifying images into the four classes of diseases and the results are classified according to accuracy.

Model architecture

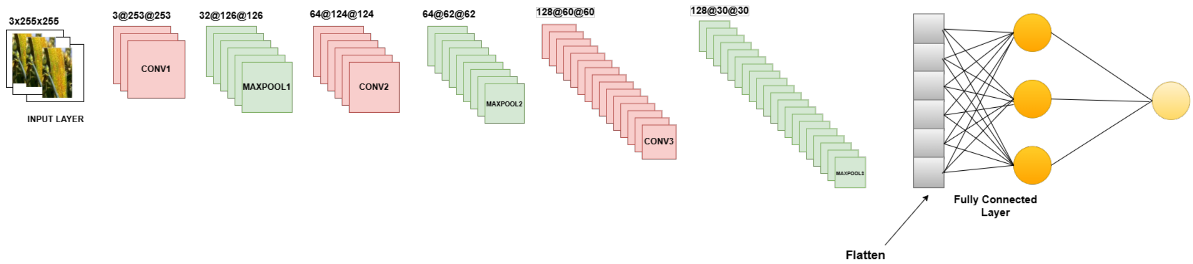
The general structure of the system is as follows:

Diseases Detection Network: General model architecture

The main network (DDN1) is responsible for checking the health of the wheat ears. It processes all the images sent by the user, resizing them to 255x255 pixels and applying filters. Next, the images pass through three layers Conv2D and Max Pooling2D, where the difference between these layers lies in the size of the kernel and the shape of the output. In addition, the images are divided into three parts to deepen the analysis, increasing the depth to 32, then 64, and finally 128. After these successive operations, the activation function used is the sigmoid function. The output classification label is either 0, indicating a healthy wheat ear, or 1, indicating a diseased wheat ear. If the final prediction is healthy, the process stops. On the other hand, if the wheat ear is classified as diseased, the secondary (slave) model is called to classify the specific disease (Figure 6).

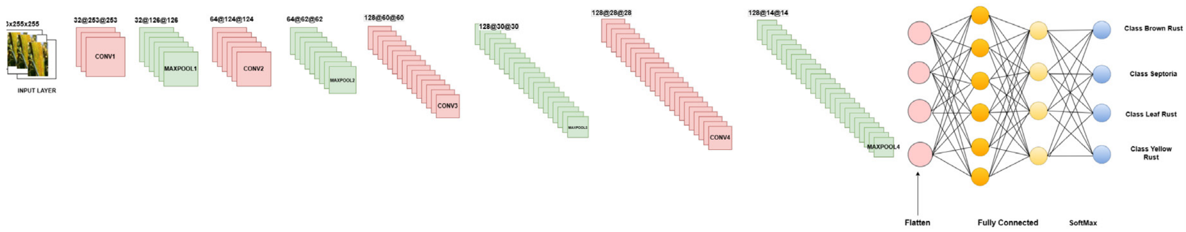
Diseases Diagnose Network: General model architecture

The DDN2 model is responsible for detecting wheat diseases and classifying them into four categories (Figure 7). It processes all images



Source: Authors' own architecture (2025)

Figure 6: Architecture of Master Network (DDN1).



Source: Authors' own architecture (2025)

Figure 7: General Architecture of the Slave Network (DNN2).

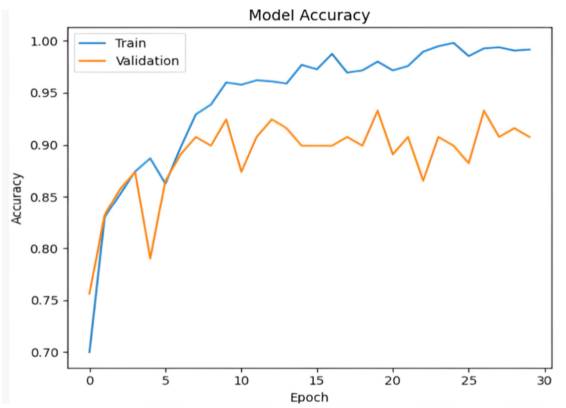
sent by the user, resizing them to 255x255 pixels and applying filters. NMS2 uses the Softmax activation function and performs operations similar to those of the first model, but with a difference in the number of classes. The same input dataset is used, and the initial structure remains un changed from the first model. The diagram below shows our second architecture, with four layers of Conv2D and Max Pooling2D. The fourth Conv2D layer is 128x28x28, followed by the MaxPooling2D layer at 128x14x14. At the end of the architecture, the feature maps are flattened to a size of 25088. Unlike the first model, where the activation function was sig moidal, the Dense layer here uses Softmax for the final clas sification of the four classes.

Results and discussion

Training of the DDN1 (Master Network)

The Detection Diseases Network (DDN1) was conducted using the following configurations (Figure 8):

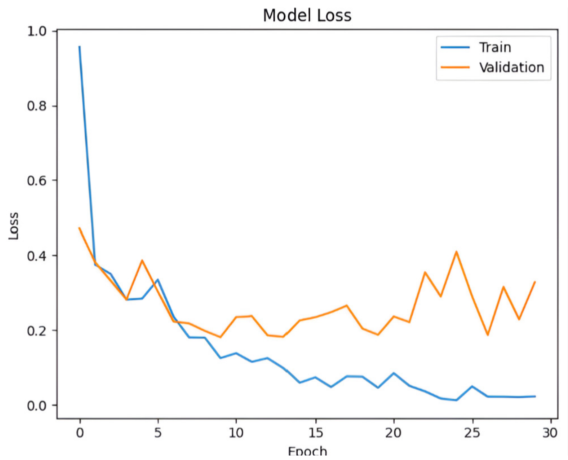
- Optimizer: Adam
- Batch size: 32
- Activation function: ReLU for all convolutional layers except the output layer (Softmax)
- Epochs: 30



Source: Authors' training results (2025)

Figure 8: Graph of the accuracy train and validation for DDN1.

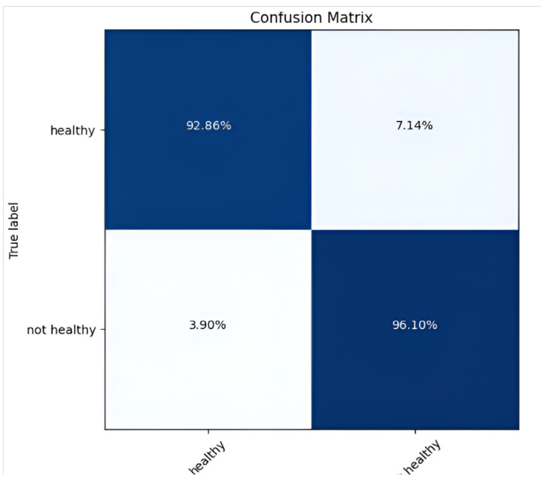
- Training Accuracy: The accuracy improved steadily with each epoch, reaching nearly 98% by the end. This indicates that the model was learning the training data very well, with minimal overfitting.
- Validation Accuracy: Initially, the validation accuracy improved, but after around 5 epochs, it started fluctuating and declined temporarily. Despite these fluctuations, the validation accuracy ultimately reached 90% (Figure 9).



Source: Authors' training results (2025)

Figure 9: Graph of the train and validation loos for the DDN1.

- Training loss : The training loss decreased steadily, starting from a high value and gradually approaching zero, suggesting that the model was fitting the training data well.
- Validation loss: The validation loss decreased at first but began fluctuating after a few epochs, remaining higher than the training loss and even slightly increasing towards the end (Figure 10).



Source: Authors' evaluation results (2025)

Figure 10: Confusion matrix for the Master Network (DDN1).

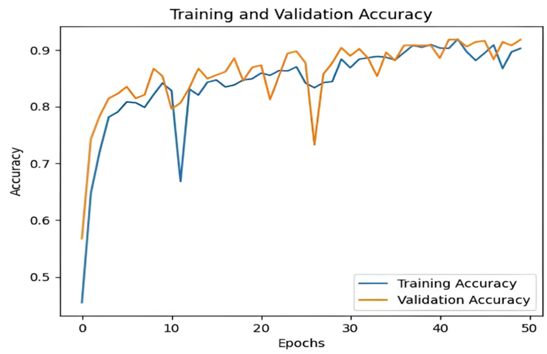
The confusion matrix for the Master Network (DDN1) reveals the following insights:

- High precision and recall: The model demonstrated excellent performance in distinguishing between healthy and unhealthy wheat, with high correct classification rates for both classes.
- False positives and false negatives: The model had relatively low rates of false positives and false negatives, indicating good precision and recall for both healthy and diseased wheat classes.
- These results suggest that the Master Network (DDN1) is highly effective in distinguishing between healthy and diseased wheat, with a very good ability to classify the wheat as either healthy or diseased.

Training of the DDN2 (Slave Network)

The Diagnose Diseases Network (DDN2) was trained using the following configurations (Figure 11)**:**

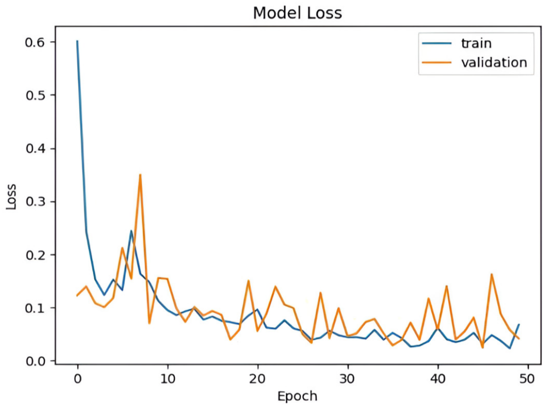
- Optimizer: Adam
- Batch size: 32
- Activation function: ReLU for all convolutional layers, with Softmax used for the output layer.
- Epochs: 50



Source: Authors' training results (2025)

Figure 11. Graph of train and validation accuracy for the Slave Network (DDN2).

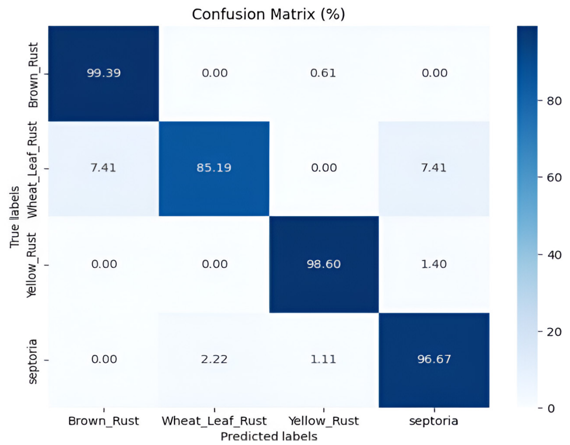
- Accuracy for training and validation starts at around 50% in the initial epoch. Gradually, these two metrics improve with each subsequent epoch, eventually reaching nearly 91% accuracy (Figure 12).



Source: Authors' training results (2025)

Figure 12: Train and validation loss for DNN2.

- The model loss for both training and validation datasets decreases significantly during the first 60% of the epochs and then stabilizes at low values. This trend indicates effective learning, strong model performance, and minimal overfitting, with a final overfitting value of approximately 0.08 (Figure 13).



Source: Authors' evaluation results (2025)

Figure 13: Confusion matrix for the slave model.

- The model shows a high accuracy for all diseases classes and, in addition, a very low misclassification rate. Brown rust, is diagnosed at the best accuracy of 99.39%, while wheat leaf rust, is diagnosed at the lowest accuracy of 85.19%. Results of misclassifications are 7.41% Wheat leaf rust samples classified as brown rust and 7.41% as septoria, probably because of the high visual similarity of wheat leaf rust, brown rust, and septoria.

To provide a more detailed analysis of the model's performance, we report additional metrics including precision, recall, and F1-score for each wheat disease class (Table 1). These metrics offer a deeper understanding of the behavior of the model beyond global accuracy, especially in terms of class-wise reliability. The F1-score, in particular, reflects the balance between false positives and false negatives, confirming the robustness of the proposed dual-CNN system across all disease categories.

Class	Precision	Recall	F1-score
Brown Rust	0.93	0.99	0.96
Wheat Leaf Rust	0.97	0.85	0.91
Yellow Rust	0.98	0.99	0.98
Septoria	0.93	0.97	0.95

Source: Authors' own calculations (2025)

Table 1: Evaluation metrics by disease class.

Conclusion

In conclusion, this study highlights the transformative potential of deep learning techniques, particularly convolutional neural networks (CNNs), in enhancing wheat disease detection and classification. Using a dual CNN system, designed to be in a master-slave fashion, the model is able, in an effective manner, to learn the diagnosis of the wheat health and to predict the diseases (yellow rust, brown rust, leaf rust and septoria). The achieved high classification accuracy shows the effectiveness of the system, and therefore the system can be considered practical in precision agriculture.

Beyond improving detection accuracy, the proposed system contributes to sustainable agriculture by optimizing pesticide usage and reducing chemical dispersion into soil and ecosystems. This environmentally friendly approach is particularly relevant considering the increasing ecological challenges posed by climate change.

This system is a significant contribution towards agricultural problem solving, sustainable food provisioning, and pesticide minimization. Through combining AI-based disease diagnosis with drone technology, the system will allow real-time surveillance, delivering quick, precise interventions to curb crop loss and guarantee food security. However, with the changing of the technology, in future investigations the model will be refined for a variety of agricultural environments and the model's robustness to uncertain cases (i.e., different wheat varieties) will be encouraged. Overall, the joint integration of smart detection technologies and drones represents the start of the "SMART Agriculture" age, greatly improving food safety, and its sustainability and crop yields.

Corresponding author:
Chemesse Ennehar Bencheriet
LAIG Laboratory, University 08 May 1945, Guelma, Algeria
Email: bencheriet.chemesseennehar@univ-guelma.dz

References

[1] Aboneh, T., Rorissa, A., Srinivasagan, R. and Gemechu, A. (2021) "Computer vision framework for wheat disease identification and classification using Jetson GPU infrastructure", *Technologies*, Vol. 9, No. 3, p. 47. E-ISSN 2227-7080. DOI 10.3390/technologies9030047.

[2] Arinichev, I. V., Polyanskikh, S. V., Arinicheva, I. V., Volkova, G. V. and Matveeva, I. P. (2022) "A Neural Network-Based Approach to Multiple Wheat Disease Recognition", *International Journal of Fuzzy Logic and Intelligent Systems*, Vol. 22, No. 1, pp. 106-115. E-ISSN 2093-744X. DOI 10.5391/IJFIS.2022.22.1.106.

[3] Bencheriet, Ch. and Bencheriet, S. (2023). “ Master-slave convolutional deep architecture for vehicle identification and type classification”, Traitement du Signal, Vol. 40, No. 2, pp. 619-627.

[4] Haider, W., Rehman, A. U. and Rehman, S. U. (2021) "A generic approach for wheat disease classification and verification using expert opinion for knowledge-based decisions", *IEEE Access*, Vol. 9, pp. 31104-31129. E-ISSN 2169-3536. DOI 10.1109/ACCESS.2021.3058582.

[5] Jadhav, S. B., Udupi, V. R. and Patil, S. B. (2021) "Identification of plant diseases using convolutional neural networks", *International Journal of Information Technology*, Vol. 13, No. 6, pp. 2461-2470. E-ISSN 2511-2112. DOI 10.1007/s41870-020-00437-5.

[6] Long, M., Hartley, M., Morris, R. J. and Brown, J. K. M. (2023) "Classification of wheat diseases using deep learning networks with field and glasshouse images", *Plant Pathology*, Vol. 72, No. 3, pp. 536-547. ISSN 1365-3059. DOI 10.1111/ppa.13684.

[7] Lu, J., Hu, J., Zhao, G., Mei, F. and Zhang, C. (2017) "An in-field automatic wheat disease diagnosis system", *Computers and Electronics in Agriculture*, Vol. 142, pp. 369-379. ISSN 0168-1699. DOI 10.1016/j.compag.2017.09.012.

[8] Moshou, D., Bravo, C., West, J., Wahlen, S., McCartney, A. and Ramon, H. (2004) "Automatic detection of yellow rust in wheat using reflectance measurements and neural networks", *Computers and Electronics in Agriculture*, Vol. 44, No. 3, pp. 173-188. ISSN 0168-1699. DOI 10.1016/j.compag.2004.04.003.

[9] Niedbała, G., Nowakowski, K., Rudowicz-Nawrocka, J., Piekutowska, M., Weres, J., Tomczak, R. J. and Álvarez Pinto, A. (2019) "Multicriteria prediction and simulation of winter wheat yield using extended qualitative and quantitative data based on artificial neural networks", *Applied Sciences*, Vol. 9, No. 14, p. 2773. ISSN 2076-3417. DOI 10.3390/app9142773.

[10] Nigam, S., Jain, R., Marwaha, S., Arora, A., Haque, M. A., Dheeraj, A. and Singh, V. K. (2023) "Deep transfer learning model for disease identification in wheat crop", *Ecological Informatics*, Vol. 75, p. 102068. E-ISSN 1878-0512, ISSN 1574-9541. DOI 10.1016/j.ecoinf.2023.102068.

[11] Rathore, N. P. S. and Prasad, L. (2022) "Hybrid deep learning model to detect uncertain diseases in wheat leaves", *Journal of Uncertain Systems*, Vol. 15, No. 3, p. 2241004. ISSN 1752-8917. DOI 10.1142/S1752890922410045.

[12] Xu, L., Cao, B., Zhao, F., Ning, S., Xu, P., Zhang, W. and Hou, X. (2023) "Wheat leaf disease identification based on deep learning algorithms", *Physiological and Molecular Plant Pathology,* Vol. 123, p. 101940. E-ISSN 1096-1178. DOI 10.1016/j.pmpp.2022.101940.

Volume XVII Number 4, 2025

Impact of Rural Out-Migration on Crop Productivity of Migrant-Sending Rural Households in Oromia Region of Ethiopia

Fassil Eshetu[1], Semeneh Bessie[2], Lamessa T. Abdisa[3], Abdulaziz Dawud[4], Fekadu Abdissa[5]

[1] Department of Economics, Arba Minch University, Ethiopia

[2] Communication and Partnership Division, Ethiopian Economics Association, Addis Ababa, Ethiopia

[3] Research and Policy Analysis Division, Ethiopian Economics Association, Addis Ababa, Ethiopia

[4] Oromia Planning and Development Commission, Addis Ababa, Ethiopia

[5] Policy Study and Research Development, Oromia Planning and Development Commission, Addis Ababa, Ethiopia

Abstract

This study quantified the impact of rural out-migration on crop productivity using the multinomial endogenous switching model as an analytical model in the Oromia region of Ethiopia. Cross-sectional data were gathered from a random sample of 384 rural households. The descriptive analysis revealed that the rate of rural-rural migration in Ethiopia decreased from 55.8 to 24.6% while the rate of rural-urban migration increased from 28.7 to 33.8 % between 1984 and 2021. The proportion of migrants in the total urban population increased from 17.2 to 49.2% in the Oromia region between 1999 and 2021. The regression results found that land size, use of irrigation, tropical livestock unit, dependency ratio, and education level of household head decrease the likelihood of participating in migration, whereas family size, number of plots, being female-headed households, and age of household head increase the probability of participating in migration. The participation in rural-urban and international migration increases the productivity of wheat producers by 341.28 and 707.21kilograms, respectively. Similarly, the participation in rural-urban and international migration increases the productivity of teff producers by 502.05 and 257.04 kilograms, respectively. This finding also supports the credit and risk hypotheses of the new economics labour migration theory. Enhancing access to finance or credit markets, agricultural land, and enhanced technology for youth in migrant-sending rural communities can leverage the gains from rural out-migration. Provision of pre-migration training, rural non-farm employment, awareness creation, promotion of safe migration, and better rural public services would capitalize the net benefit from out-migration.

Keywords

Migration, new economics labor migration, productivity, switching model, Oromia.

Eshetu, F., Bessie, S., Abdisa, L. T., Dawud, A. and Abdisa, F. (2025) "Impact of Rural Out-Migration on Crop Productivity of Migrant-Sending Rural Households in Oromia Region of Ethiopia", *AGRIS on-line Papers in Economics and Informatics*, Vol. 17, No. 4, pp. 37-56. ISSN 1804-1930. DOI 10.7160/aol.2025.170404.

Introduction

Migration is part and parcel of economic development and structural transformation in developing countries (World Bank, 2023). In structural transformation, people move from working only in rural agricultural sectors into urban non-agricultural sectors such as manufacturing and services. When migration flows from the low-productivity rural agricultural sector to more productive urban industrial sectors, it can benefit both migrants and the rural households they leave behind (ILO, 2022). Recently, migration patterns have shifted more prominently from lagging rural agricultural areas to leading urban centers and from developing to developed countries (FAO, 2021).

While the number of international migrants worldwide increased from 173 million in 2000 to 281 million in 2023, the total amount of global remittances rose from USD 128 billion to USD 860 billion over the same period (IOM, 2024; World Bank, 2024). Moreover, the share of international remittances flowing to developing countries grew from 57% in 2000 to approximately

81% in 2023 (World Bank, 2024). However, rural out-migration remains a development challenge in agrarian economies for at least three reasons: the composition, rate, and direction of migration. First, most rural out-migrants are young, better educated, more informed, and unmarried, resulting in a potential loss of human capital in rural agriculture. Second, the current rate of rural out-migration surpasses historical migration rates during the industrial revolution in now-developed countries (de Haas, 2011). Third, migration tends to move from rural areas, where job creation potential is relatively higher and underemployment is more common, to urban areas where job creation is limited and unemployment rates are higher (ILO, 2022).

Rural out-migration affects agricultural output and welfare in communities of origin by transferring labor and remittances between sending and receiving regions (Stark and Bloom, 1985). Two primary ways that migration impacts these economies are through remittances, which can increase farm investment and household welfare by reducing credit constraints, and labor loss, which can lower agricultural productivity by displacing young, educated, and skilled workers (Lucas, 1987; Lagakos et al., 2020). The opportunity cost of lost labor and the way remittances are used determine the overall effect. Remittances can help compensate for labor shortages if they are used productively; if not, migration may cause rural production to fall over time (Taylor and Wyatt, 1999; Zahonogo, 2011; United Nations, 2016).

Ethiopia is the 2nd most populous country in Africa and the 12th in the world (World Bank, 2021). Migration patterns in Ethiopia have varied across political regimes (Adugna, 2021). During the imperial era (1941–1974), both rural-urban and international migrations were limited. But, under the military regime (1974–1991), international migration increased while rural-urban migration was restricted by policy (Witten, 2007). Following the demise of the socialist government in 1991, both forms of migration have increased significantly. Between 2000 and 2023, the number of international Ethiopian migrants increased from approximately 611,000 to over 2.5 million, reflecting a growing diaspora (IOM, 2024). During the same period, remittance inflows to Ethiopia rose sharply from USD 53 million in 2000 to USD 539 million in 2023 (World Bank, 2024). Rural-urban migration also grew, with its share rising from 21.6% to 32.2%, while rural-rural migration declined from 35.5% to 23.4% between 1999 and 2021 (ESS, 2021). Oromia is the leading source of international migrants, whereas Amhara leads in rural-urban migration. Regarding destinations, 30.7% of Ethiopian migrants go to Saudi Arabia, followed by South Africa (12.4%) and the UAE (8.9%) (ESS, 2021). Ethiopians primarily migrate through three corridors: the eastern route to the Middle East, the northern route through Sudan to Europe, and the southern route to South Africa, where Ethiopia accounts for two-thirds of Horn of Africa migrants (Horwood, 2009; Massey et al., 1998; Abire and Sagar, 2016).

Despite the rising rate of rural out-migration in Ethiopia, particularly in the Oromia region, empirical studies examining its impact on crop productivity in migrant-sending areas remain limited. While some studies (Odozi et al., 2020; Mesfin et al., 2021) report a positive association between migration and crop productivity, others (Khanal et al., 2015; Imran et al., 2016; Adaku, 2019) find a negative association. However, few of these studies apply the New Economics of Labor Migration (NELM) framework or focus on specific staple crops in the Ethiopian context. Hence, this research addresses that gap by investigating the trends, determinants of rural out-migration, and quantifying its impact on the productivity of wheat and teff, two of Ethiopia's most important crops. Wheat is a key staple crop that is targeted by national self-sufficiency programs and agricultural modernization efforts, while teff is a culturally significant, indigenous crop with widespread consumption and high market value, particularly in Oromia. Understanding how migration affects the productivity of these crops is critical for evidence-based rural development and agricultural policy. The remainder of this paper is structured as follows. The second section contains a review of the relevant literature, the third section discusses the research methods, the fourth section presents the findings and a discussion, and the last section contains a conclusion of the study.

Literature review

Theoretical review

From the earliest individual-based explanations to more contemporary household-level tactics, migration ideas have evolved. Examples of classical models that have impacted the understanding of the causes of rural-to-urban migration include the push-pull framework (Lee, 1966), the dual-sector model (Lewis, 1954), the gravity model

(Ravenstein, 1885), and the human capital model (Harris and Todaro, 1970). These theories generally emphasize economic opportunities, labor productivity differences, and wage gaps as reasons influencing migration. However, many of their assumptions, such as the unlimited industrial absorption of rural labor, are less applicable in developing nations where low industrialization and high urban unemployment prevail.

While Lewis's (1954) dual-sector theory contended that excess rural labor moves to more productive urban industries, Ravenstein's (1885) gravity model emphasized the flow from rural agricultural areas to urban centers with higher economic opportunities. Lee's (1966) push-pull theory divided the reasons for migration into four categories: personal causes, intervening hurdles, pull factors (better jobs, education), and push factors (limited land, bad services). Harris and Todaro (1970) explained the coexistence of rural-urban migration and urban unemployment by framing migration as a function of predicted wage differentials. Although these viewpoints offer crucial background information, they are not very good at capturing the intricate socioeconomic tactics that influence migration choices in rural economies.

By changing the unit of study from persons to households, the new economics of labor migration (NELM) provides a more comprehensive perspective (Stark and Bloom, 1985; Taylor, 1999). It makes the case that migration is a risk management tactic to get around market inefficiencies, especially in rural credit and insurance markets, rather than just a reaction to pay disparities. Families send migrants in order to secure remittances and diversify their sources of income, which may then be used to support consumption and investment in agriculture (Lucas and Stark, 1985). This viewpoint is more applicable to the realities of emerging agrarian economies since it takes into consideration both financial limitations and the agency of rural households in deciding to migrate.

The remittance channel, which can improve agricultural production and welfare by reducing liquidity constraints, and the lost labor channel, which can lower productivity and welfare by depleting the household's labor force, are the two opposing channels through which migration affects rural livelihoods. This is one of the main contributions of the NELM framework. The fact that migration has both positive and negative effects is highlighted by this dual effect. This study uses the NELM framework to analyse the effects of migration on crop productivity in the study area, since it is appropriate for examining both the advantages and disadvantages of migrationin rural settings.

Empirical review

Results from empirical studies on the connection between migration and agricultural productivity in rural areas that send migrants are wildly inconsistent. Rural–urban migration can improve agricultural production or welfare in origin areas, according to a number of studies, including de Brauw and Giles (2017) in China, in Ethiopia, and Chamberlin et al. (2020) in Zambia. These studies frequently attribute these gains to remittances that finance agricultural inputs or raise household living standards. Similar benefits are reported by Agza et al. (2023) in southern Ethiopia and Bassie et al. (2022) in Ethiopia, where migration was found to increase welfare and productivity. These results imply that migration may serve as a stimulant for rural development in specific circumstances, such as when remittance flows are put back into farming.

On the other hand, several studies document negative effects of migration, which are frequently connected to a lack of workers in rural areas. For example, studies conducted in Nigeria (Sennuga et al., 2023), Ghana (Kaur and Kaur, 2021), and the Philippines (Morales and Villaronte, 2022) show that migration dramatically lowers agricultural productivity by reducing household labor and increasing reliance on hired labor, which can raise production costs. Out-migration can reduce technical efficiency and production, according to broader multi-country data from Khanal et al. (2015), Imran et al. (2016), Goldsmith et al. (2017), Adaku (2019), and Sauer et al. (2013), as well as similar negative relationships in Vietnam. These findings demonstrate how the effects of migration vary greatly depending on the context, being influenced by factors such as remittance utilization, labor market conditions, and automation levels.

The relationship between migration and productivity is not consistently positive or negative, but rather fluctuates depending on the type of crop, household characteristics, and institutional setting, according to another study. Research conducted in Nigeria by Odozi et al. (2020) and Albania indicates that households with migrants can attain greater technical efficiency or production, perhaps as a result of improved access to money and technology. But as Mesfin et al. (2021) in Ethiopia and Iheke et al. (2013) in Nigeria show, technical efficiency gains are not always accompanied by a decrease in the availability of on-farm labor. This heterogeneity implies that it might be deceptive to extrapolate generalizations from single-country research without taking into account variations in migratory trends, agricultural systems, and remittance usage.

Overall, there is disagreement in the empirical literature over whether migration increases or decreases agricultural output. This is partially because of methodological variations (such as propensity score matching, instrumental variables, and panel models)and contextual diversity (such as Asia, Africa, and Eastern Europe). The mechanisms by which migration influences agricultural outcomes, such as the trade-off between labor loss and remittance investment, are not well studied, nor are situations directly compared. There is also little evidence from the Oromia region of Ethiopia, where agricultural systems and migratory trends may be very different from those researched areas. The present study, which attempts to evaluate the effect of migration on agricultural output while taking local institutional and socioeconomic variables into consideration, is necessary because of this gap. By explicitly addressing these mechanisms, the study seeks to clarify the conditions under which migration can be beneficial or detrimental to rural agricultural systems.

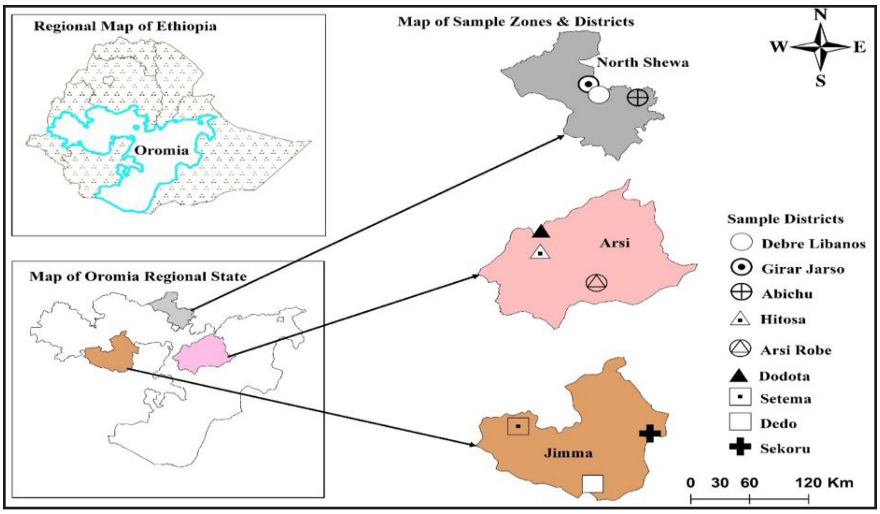
Materials and methods

The study area

This study was conducted in three selected zones of the Oromia regional state of Ethiopia in 2023. Oromia is the largest region in Ethiopia, both in terms of population size and land area. The region shares borders with all regions of Ethiopia except the Tigray region. The 2022 projected population of the region was 42,647,632 2022 (Regional Plan and Development Commission (RPDC), 2022), of which 35,453,080, or 83 percent of the population, live in rural areas. The total area of the region is 363,375 square kilometres. Administratively, the region is divided into 21 administrative zones, 23 town administrations, 294 rural districts/ woredas, and 29 towns. From the total of 21 zones[1] in Oromia National Regional State, three major rural-urban and international migrant sending zones, namely, Arsi, Jimma, and North Shawa, were covered by this research.

According to the latest population estimates, the total population of Jimma was 3,568,782; Arsi was 3,894,248, and the North Shewa zone was 2,100,331 (Ethiopian Statistical Service (ESS), 2022). Jimma zone has 20 woredas, while Arsi and North Shawa have 24 and 11 woredas, respectively. From each Zone, 3 woredas were selected for this study, namely: Setema, Sekoru, and Dedo woredas from Jimma zone; Dodota,

[1] Zones in Oromia include Arsi, Bale, Bedele, Borena, East Haraghe, East Shewa, East Welega, Guji, West Guji, Horo Gudru Welega, Illubabor, Jimma, Kelem Welega, North Shewa, Southwest Shewa, West Arsi, West Hararghe, West Shewa, West Welega, Adama Special Zone, Jimma Special Zone and Oromia-Finfinnes Special Zone.



Source: Authors' preparation (2023)

Figure 1: Map of sample zones and districts of Oromia Regional State

Arsi Robe, and Hitosa woredas from Arsi zone; and Girar-Jarso, Abichu, and Debre Libanos woredas from North Shewa zone were purposively selected for this study (see the Figure 1 above).

Data sources and data collection instruments

This study primarily employed a quantitative-dominated mixed research design to generate a better understanding of the causes and impacts of rural out-migration on the economy of migrant-sending rural areas of the Oromia region. Structured questionnaire, key informant interview, and focus group discussion were used to gather primary data. Primary data on demographic characteristics, sources of rural out-migration, use of remittances, agricultural production, and welfare of rural sample households were gathered from 384 households between January 20 to February 20/2023, from nine sample districts in the Oromia region. Besides, secondary data on the trends of rural out-migration, major sources and destinations of rural-urban and international migrants were obtained from the Central Statistical Service. To support the results from the quantitative data, qualitative data on causes of migration, positive and negative effects of participation in rural out-migration were gathered using key informant interviews (KIIs) and focus group discussions (FGDs). Accordingly, 32 key informant interviews were conducted in this study. The participants in the key informant interview were selected purposively based on their expertise and professional contributions to the study from different offices, such as women and children, labour and social affairs, job creation, and the policy commission. Moreover, 8 focus group discussions were held, and 5-8 participants were included in each FGD. The participants in the FGDs also include elders, development agents, religious leaders, cultural leaders, youth and women group leaders, school principals, community representatives (local police), and return migrants.

Sampling procedures and sample size

A multistage sampling technique was employed to select sample households for this study. First, three sample zones, namely Jima, Arsi, and North Shewa, were selected for this study purposively from 20 zones[2] in the Oromia region. This is because the three zones are the primary sources of both rural-urban and international migrants in the Oromia region. Second, nine major migrant-sending sample woredas were chosen from the three sample zones. As a result, Setema, Dedo, and Sekoru woredas were selected from Jimma zone, while Dodota, Arsi Robe, and Hitosa woredas were chosen from Arsi zone. Likewise, Girar Jarso, Debre Libanos, and Abichu woredas were selected from the North Shewa zone. Third, two kebeles from each of the woredas in Jimma and Arsi Zones, while one kebele from each of the woredas in the North Shewa zone was selected. That means a total of 15 major migrant-sending sample *Kebeles*[3] were selected for this study. Fourth, the sample households were allocated among the three zones using Probability Proportional to Size (PPS). The samples were further allocated for migrant sending and non-migrant sending households, with 2/3 allocated for migrant-sending households from each woreda and kebele, while 1/3 was allocated for non-migrant sending households. The overall sample size is determined using Cochran's (1963) sample determination formula as follows (Equation 1):

$$\boldsymbol{n} = \frac{\boldsymbol{Z^2pqN}}{\boldsymbol{e^2(N-1)+Z^2pq}} \qquad (1)$$

Where *e, p, q, n, N*, and *Z* are the measures of precision, the assumed level of variability in the population, one minus the level of variability in the population, the sample size of the study, the total population, and the value of the standard normal distribution, respectively. The total households (N) in the three sample zones, the degree of variability, and the level of precision in this study are 1249711, 0.5, and 0.05, respectively. Based on the above formula, a sample size of 384 is determined for this study. Hence, quantitative data were collected from 384 rural households on the causes and impact of rural out-migration in the region using a survey questionnaire in the year 2023. Participants in this study were divided into three groups, namely: households without migrants, households with international migrants, and households with rural-urban migrants. However, households with both international and rural-urban migrants were grouped under households with international migrants due to their few numbers (9).

[2] There are (21) zones in Oromia region and they include Arsi, Bale, Bedele, Borena, East Haraghe, East Shew, East Welega, Guji, West Guji, Horo Gudru Welega, Illubabor, Jimma, Kelem Welega, North Shewa, South West Shewa, West Arsi, West Hararghe, West Shewa, West Welega, Adama Special Zone, Jimma Special Zone and Oromia-Finfinnes Special Zone.

[3] The sample Kebeles include Asandabo and Habe Dangazela (Arsi Robe), Dirre Kiltu and Dodota Alem (Dodota), Jawi Chilalo and Sero Ankato (Hitosa), Seo Sidisa and Karti Wosorbi (Dedo), Yera Docha and Chafeta (Setema), Yabbu and Haro Kake (Sekoru), Ano Akabdo (Abichu), Wartu (Girar Jarso) and Wakene (Debre Libanos).

Method of data analysis

Both descriptive and inferential methods of data analysis were applied in this study. The descriptive methods include percentages, bar graphs, frequencies, means, standard deviations, and time series graphs. The inferential analyses, such as mean difference test, analysis of variance (ANOVA), and multinomial endogenous switching model, were employed to answer the research objectives. Since the problem of self-selection biases due to observed and unobserved factors is a common problem in migration analysis, this study employed a multinomial endogenous switching regression model to evaluate the impact of participation in rural-urban and international migration on the crop production of migrant-sending rural households in origin areas. The quantitative data were analysed using STATA 17 and SPSS 23 Statistical Software.

Model specification

To examine the impact of rural out-migration on crop productivity of migrant-sending rural households, wheat output per hectare and teff output per hectare were used as outcome variables. The treatment variable is rural out-migration, which is a nominal variable with three categories, namely, households without migrants (j = 0), with rural-urban migrants (j = 1), and international migrants (j = 2). But there is a problem of self-selection into migration due to both observed and unobserved factors. Put differently, participation in migration is not random, and households with similar characteristics may participate in rural-urban migration or international migration. To account for this selection problem, this study employed the multinomial endogenous switching model. The multinomial endogenous switching model was developed by Deb and Trivedi (2006) to control for endogeneity due to observed and unobserved factors. Based on the concept of expected utility maximization, rural households may participate in rural out-migration if the expected utility from rural out-migration is higher than the expected utility without participation. Following Deb and Trivedi (2006), the latent variable model, which describes the behaviour of rural households in choosing one alternative among the three alternatives to maximize its expected utility, is given by (Equation 2):

$$Y_{ij}^{*} = \beta_i Z_i + U_{ij} \tag{2}$$

Where Y_{ij}^{*} is the latent variable that measures the expected utility of the i^{th} household from choosing among the j^{th} alternative, i = 1,2,3…384, j = 0,1,2, Z_i is a vector of exogenous covariates, β_i is a vector of parameters to be estimated and U_{ij} is an error term. In the multinomial endogenous switching model, a household has j choices, and the latent outcome variable is given by (Equation 3):

$$Y_{ij} = \begin{cases} 1 \ iff \ Y_{i1}^{*} > \max\limits_{k\neq 1}(Y_{i1}^{*}), & U_{i1} < 0 \\ \vdots & \vdots \\ M \ iff \ Y_{iM}^{*} > \max\limits_{k\neq M}(Y_{i1}^{*}), & U_{iM} < 0 \end{cases} \tag{3}$$

where Y_{ij} is the observed value of the outcome variable for the i^{th} household choosing alternative j, U_{i1}, U_{i2} … U_{iM} are error terms of the outcome equation, i = 1,2,3…384, j = 0,1,2, and Y_{i1}^{*} is the latent variable. Given the assumption that U_{ij} is independently and identically distributed or the independence of the irrelevant alternatives (IIA) assumption, the selection model of migration leads to the following multinomial logit model, where the likelihood of choosing alternative j *(Equation 4)*:

$$P_{ij} = \Pr(U_{ij} < 0/Z_j) = \frac{e^{\beta_i Z_i}}{\sum_{k=1}^{M} e^{\beta_i Z_i}} \tag{4}$$

In the second stage, a multinomial endogenous switching regression model is used to quantify the impact of participation in rural out-migration on the crop productivity of migrant-sending rural households. Rural households without a migrant family member, j = 0 is the base category in this study. Hence, the output per hectare of wheat or teff (Q) for the i^{th} household is defined as m regime:

Regime 0: $Q_{i0} = X_i\gamma_0 + \varepsilon_{i0}$, if $j = 0$ (5)

Regime 1: $Q_{i1} = X_i\gamma_1 + \varepsilon_{i1}$, if $j = 1$ (6)

Regime 2: $Q_{i2} = X_i\gamma_2 + \varepsilon_{i2}$, if $j = 2$ (7)

where Q_{ij} is the output per hectare of wheat or teff in kilograms of the i^{th} household in regime j, i = 1,2,3…384, j = 0,1,2, X_i is a vector of covariates, and ε_{ij} is the unobserved factor. Based on equations (5), (6), and (7), the selection bias-corrected outcome equations are given (Equations 8, 9, 10):

Regime 0:

$$Q_{i0} = X_i\beta_0 + \delta_0\left[\rho_0 m(P_{i0}) + \sum_j \rho_j\, m(P_{ij})\left(\frac{P_{ij}}{P_{ij}-1}\right)\right] + $$
$$+\varepsilon_{i0}\,,\ \ if\ \ j=0 \tag{8}$$

Regime 1:

$$Q_{i1} = X_i\beta_1 + \delta_1\left[\rho_1 m(P_{i1}) + \sum_j \rho_j\, m(P_{ij})\left(\frac{P_{ij}}{P_{ij}-1}\right)\right] + $$
$$+\varepsilon_{i1}\ \ if\ \ j=1 \tag{9}$$

Regime 2:

$$Q_{i2} = X_i\beta_2 + \delta_2\left[\rho_2 m(P_{i2}) + \sum_j \rho_j\, m(P_{ij})\left(\frac{P_{ij}}{P_{ij}-1}\right)\right] + $$
$$+\varepsilon_{i2}\ \ if\ \ j=2 \tag{10}$$

where P_{ij} is the probability that the i^{th} rural household chooses the j^{th} alternative, ρ_j is the degree of correlation between the error term of the participation equation, U_{ij} and the error term of the outcome equation, ε_{ij} and $m(P_{ij})$ is the inverse transformation for the normal distribution function. The multinomial endogenous switching regression model is also used to estimate the counterfactual data to quantify the impact of rural out-migration on crop productivity of migrant-sending rural households. Following the work of Bourguignon et al. (2007) and assuming households without migrants, $j = 0$ as the base category, values of output per hectare of wheat or teff in kilograms for households with migrants are given by (Equation 11 and 12):

$$E(Q_{i1}/j=1) = X_i\beta_1 +$$
$$+\delta_1\left[\rho_1 m(P_{i1}) + \sum_{k=1}^{M}\rho_k m(P_{ik})\left(\frac{P_{ik}}{P_{ik}-1}\right)\right] \tag{11}$$

$$E(Q_{i2}/j=2) = X_i\beta_2 +$$
$$+\delta_2\left[\rho_2 m(P_{i2}) + \sum_{k=1}^{M}\rho_k m(P_{ik})\left(\frac{P_{ik}}{P_{ik}-1}\right)\right] \tag{12}$$

Moreover, once the actual mean values of output per hectare of wheat or teff in kilograms for rural households are determined using the above two equations, the mean values of output per hectare of wheat or teff in kilograms for households from the counterfactual data are given by (Equation 13 and 14):

$$(Q_{i0}/j=1) = X_i\beta_0 +$$
$$+\delta_0\left[\rho_0 m(P_{i1}) + \rho_1 m(P_{i0})\left(\frac{P_{i1}}{P_{i1}-1}\right) + \rho_1 m(P_{i1})\left(\frac{P_{i3}}{P_{i3}-1}\right)\right] \tag{13}$$

$$(Q_{i0}/j=2) = X_i\beta_0 +$$
$$+\delta_0\left[\rho_0 m(P_{i2}) + \rho_2 m(P_{i1})\left(\frac{P_{i1}}{P_{i1}-1}\right) + \rho_1 m(P_{i0})\left(\frac{P_{i3}}{P_{i3}-1}\right)\right] \tag{14}$$

Lastly, the conditional average treatment effects on treated (ATT) could be computed by subtracting Equations (13) and (14) from equations (11) and (12), respectively. The positive and significant values of ATT imply that participation in rural out-migration promotes crop productivity of migrant-sending rural households via the remittance channel.

Description of variables and hypotheses

In the first stage regression of the multinomial endogenous switching model, the dependent variable is rural out-migration, which is a nominal variable with three categories, namely, households without migrants (j = 0), households with rural-urban migrants (j = 1), and households with international migrants (j = 2). The occurrence of drought may induce rural out-migration by reducing the income of rural farm households. Studies conducted by Ma et al. (2019) and Abeje (2021) on determinants of rural-urban migration in northern Ethiopia found that the occurrence of drought is positively and significantly related to the propensity of rural out-migration, but the land size and rural-urban migration are negatively and significantly associated (Table 1).

Similarly, family size is considered a pushing factor for labour out-migration from rural areas. For instance, studies conducted by Alarima (2019) and Ma et al. (2019) on factors affecting rural-urban migration using cross-sectional data found that family size and years of schooling of household heads are positively and significantly related to the likelihood of rural out-migration. Besides, a study conducted by Kefelegn (2020) found that family size, years of schooling of household head, being female-headed households, and drought are positively and significantly related to rural out-migration. This study also hypothesizes that family size, education, being a female-headed household, and the occurrence of drought are positively associated with rural out-migration. Added to these, a study conducted by Tegegne and Penker (2016) also found that age, education

Variables	Description	Measurement	Sign
AGE	Age of household head	Continuous	+
EDUC	Education of the household head	Continuous	+
FS	Family size	Discrete	+
LS	Land size in hectares	Continuous	-
TLU	Tropical livestock unit	Continuous	-
Male	Sex of household head	Male = 1 & Female = 0	±
IRR	Use of irrigation	Users = 1 & non-users = 0	-
DPR	Dependency Ratio	Continuous	+
RM	Presence of Return Migrants	Presence = 1 & 0 otherwise	+
NFP	Non-Farm Participation	Participant = 1 &0 Otherwise	-
DR	Drought in the last five years	Occurrence = 1 & 0 otherwise	+
LR	Participation in land renting out	Renting = 1 & 0 otherwise	+
EXTN	Frequency of extension visits	Discrete	-
PSNP	Productive Safety Net Program	Users = 1 & 0 otherwise	±
Oromo	Dummy for Ethnicity of Household	Oromo = 1 & 0 otherwise	+
Arsi	Place a dummy for the zone	Arsi = 1 & 0 otherwise	+
Jimma	Place a dummy for the zone	Jimma = 1 & 0 otherwise	±
Muslim	Dummy for the religion of households	Muslim = 1 & 0 otherwise	+

Source: Authors' preparation (2023)

Table 1: Description, measurement, and expected signs.

level of household heads, and being female-headed households are positively and significantly related to the likelihood of rural-urban migration. However, Wondimagegnhu and Zeleke(2017) examined the determinants of rural-urban migration in Ethiopia and found a negative and significant association between tropical livestock units and rural-urban migration.

Further, a study conducted by Ajaero et al. (2018) on determinants of rural out-migration found that being male-headed households, family size, and age of household head are positively and significantly related to participation in migration. Similarly, Khatir and Rezaei-Moghaddam (2014) conducted a study on predictors of rural out-migration of youth in Iraq and indicated that family size, age of household head, being male-headed household, number of active male family members, and the occurrence of drought are positively and significantly related to the likelihood of rural out-migration while the land size and frequency of extension visits are negatively and significantly associated with rural out-migration.

The outcome variable in the second-stage regression of multinomial endogenous switching regression is the output per hectare of wheat or teff in kilograms. The independent variables in the second stage regression include all independent variables in the first stage regression, less two instrumental variables such as religion and return migrants. Religion and the presence of return migrants served as exclusionary restriction variables in this study. Religion shapes cultural values, social norms, and access to migration networks, all of which either promote or impede migration patterns (Feliciano, 2018). Importantly, after controlling for other socioeconomic and agro-ecological factors, religion is unlikely to have a direct impact on agricultural output. Similarly, the presence of return migrants reduces the costs and uncertainties of migration by providing vital knowledge, financial resources, and logistical support (Gao et al., 2019; Martínez and Rubio, 2022). Return migrants may not directly affect agricultural productivity, but they do have an impact on the probability of migrating. A falsification test further confirmed the validity of the selected instruments in this study.

The treatment variable in the second-stage regression of the multinomial endogenous switching model is rural out-migration, which is a nominal variable with three categories, namely, households without migrants, with rural-urban migrants, and international migrants. Regarding the impact of participation in rural out-migration on the crop productivity of migrant-sending households, there are dichotomous results. While studies conducted by Odozi et al. (2020), and Mesfin et al. (2021) found a positive and significant association between rural-urban migration and crop productivity, studies conducted by Bryan et al. (2014)Khanal et al. (2015), Imran et al. (2016), and Adaku (2019) found a negative and significant relationship.

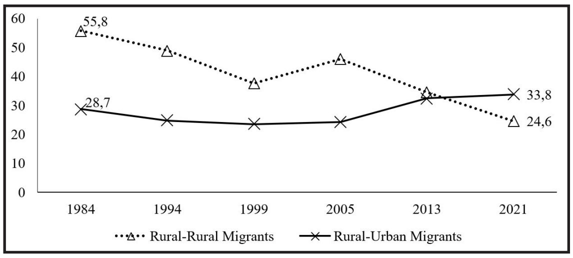
Results and discussion

This section presents both the descriptive and inferential results of data analysis. The descriptive results include the dynamics of rural-urban migration in the Oromia region and the characteristics of sample households in the study area. The inferential results include the independent samples t-test, which compares the means of two groups, and the econometric analysis on the drivers of rural out-migration, and the impact of rural-urban and international migration on crop productivity of migrant-sending households in rural areas of the Oromia region of Ethiopia.

Dynamics of rural-urban migration in Oromia region

The share of migrants from the total urban population was 27.2 percent in the Oromia region in 1999, and this figure increased to 49.2 percent in 2021, as indicated in Table 2. As shown in Table 2, the share of migrants from the total urban population of the Oromia region has been increasing over time, and is relatively higher compared to other regions such as the Amhara and SNNP regions. However, the share of migrants in the total rural population in Oromia regional state decreased between 1999 and 2021 from 15.5 to 12.5%. The implication is that both the urban and rural areas of Oromia are the major migrant receiving areas compared to Amhara and SNNP regional states.

The high share of migrants in the total urban population is primarily contributed to by the current high wave of rural-urban migration. Of all types of internal migration, rural-urban migration has recently become the most dominant type of internal migration in Ethiopia in general and the Oromia region in particular. As indicated in Figure 2, the rate of rural-rural migration decreased from 55.8 to 24.6 percent while the rate of rural-urban migration increased from 28.7 to 33.8 percent between 1984 and 2021. Furthermore, a study on rural outmigration highlights that ongoing land shortages and youth aspirations increasingly drive rural to urban movement, underlining why rural to urban migration is now the most significant internal migration stream. The implication is that rural-urban migration involves the transfer of labour from a place where job creation is easy, rural areas, to a place where job creation is difficult, urban areas. That means job creation in rural areas requires less capital compared to job creation in urban areas. Migration from rural to urban areas is crucial for structural change because it lowers the number of farmers while increasing employment in the urban industrial sector. However, rural-urban migration may result in high urban unemployment and a shortage of agricultural products if it is primarily driven by push factors, occurs before agricultural development, and is from the less productive agricultural sector to the least productive service sector.



Source: Authors' computation from different rounds of NLFs and population census

Figure 3: percentage of rural-rural and rural-urban migrants in Ethiopia over time.

Further, data on the total number of migrants in the Oromia region, as well as the reasons for migration, were gathered from Ethiopia's last four national labor force surveys. According to Table 3, the main reasons for out-migration in the Oromia region are a lack of jobs in rural

Year	Urban				Rural			
	1999	2005	2013	2021	1999	2005	2013	2021
Ethiopia	49.3	39.4	44.4	39.9	15.0	12.5	8.5	10.1
Oromia	**27.2**	**27.8**	**46.6**	**49.2**	**15.5**	**15.2**	**9.1**	**12.5**
Amhara	25.1	38.5	50.0	42.9	13.8	9.5	8.0	8.3
SNNP	20.6	30.6	42.1	31.6	13.6	10.4	7.9	11.8

Source: Authors 'Computation from ESS, 1999; 2005; 2013 and 2021, Labour Force Surveys

Table 2: Proportion of migrants in the total urban Pppulation of Oromia, Amhara and SNNP.

Years	1999	2005	2013	2021
Migrants in the Oromia Region	897,429	1,338,864	1,464,194	1,833,481
Reasons for Migration				
Search for Work	**15.4**	**8.8**	**26.3**	**29.4**
Job Transfer	4.5	4.2	5.5	7.8
Land Shortage	2.4	19.1	1.8	2.0
Education	8.8	10.3	7.9	9.9
Marriage Related Factors	**16.2**	**7.5**	**10.9**	**13.6**
Conflict & Drought	1.4	6.8	1.6	3.2
Join Family or Relatives	**37.6**	**35.4**	**29.2**	**29.4**
Health	0.5	1.3	2.1	1.1
Others	13.0	6.6	14.7	3.6

Source: Authors' Computed from the 1999, 2005, 2013, and 2021 NLFS

Table 3: Total migrants and the distribution of reasons for migration in Oromia region.

areas, the presence of family members or relatives in destination areas, and marriage-related factors. As shown in Table 3, the presence of a relative or family member in the destination area is the main pulling factor of migration in the study area, while a lack of job opportunities in the migrant-sending origin area is the main pushing factor of migration.

Still, as shown in Table 3, the percentage of migrants migrating due to marriage-related reasons has been increasing in the Oromia region. This descriptive result implies that creating viable farm and non-farm employment opportunities in origin areas will help to reduce the current wave of rural out-migration. Urban-biased development policies, on the other hand, would fuel rural out-migration and exacerbate the current high rate of urban unemployment. Therefore, the pull factors in destination areas and push factors in sending areas are the primary drivers of the movement of people from rural areas to urban areas in the Oromia region. A qualitative reviews show that marriage acts as an important pull factor for male migrants heading to urban areas, alongside economic incentives like job opportunities and higher expected income (Fassil and Mohammed, 2017). In Oromia and other rural parts of Ethiopia, limited employment options, including inadequate public sector jobs and constrained livelihoods, are identified as primary push factors, encouraging migration, while better access to jobs and services in urban areas serves as a pull factor (Tesfaye et al., 2023; Fente et al., 2023).

Characteristics of sample households in study area

The majority of the sample households (87%) are male-headed, and 65% of them are Muslims. Of the total sampled households (384), 40% (153) have no migrant, while 21% (82) and 39% (149) of the sample households have at least one rural-urban and international migrant, respectively. Besides, 22% of sample households are participants in off-farm activities in the Oromia region. But the national rate of participation in off-farm activities is 25% in Ethiopia. As can be seen from Table 4, about 10% of sample households are productive safety net users, while nearly 14% of sample households are irrigation users in the Oromia region.

Qualitative Variables	Categories	Frequency	Percent
Sex	Female	50	13.0
Migration	No Migrants	153	39.8
	Rural-Urban	82	21.4
	International	149	38.8
Sample Zones	Arsi	156	40.6
	Jimma	144	37.5
	North Shawa	84	21.9
Religion	Muslim	250	65.1
	Orthodox	125	32.6
	Others	9	2.3
Ethnicity	Oromo	370	96.4
	Amhara	11	2.9
	Others	3	0.8
Irrigation	Users	53	13.8
Productive Safety Net Program	Users	40	10.4
Credit	Users	38	9.9
Off-Farm Participation	Participants	84	21.9

Source: Authors' Computation, 2023

Table 4: frequency distribution of some qualitative variables of sample households.

The independent samples t-test was conducted to test the mean of continuous variables by migration status, and the result is presented in Table 5. The average age of the household head is about 49 years, with average family size and years of schooling being 6.64 and 3.98, respectively. The average age of household heads with migrant household members is higher compared to the average age of household heads without migrants, and the mean difference is significant at 1% level. This could explain a relationship between the age of the household head and migration in the sense that older household heads could have a large family size, which, on the other hand, encourages migration because of resource sharing among members of the household. Likewise, the average family size of households with migrants is higher compared to households with no migrants in the study areas, and the mean difference is also significant at 1% level. This could be because households with large family sizes are more likely to participate in rural out-migration as predicted by the push and pull factors theory of migration. According to Lee (1966), large family size and limited access to agricultural assets are the main push factors. Previous research has demonstrated that, depending on the environment and compensating mechanisms, migration can both increase productivity (via remittances and hired labor) and decrease it (through home labor shortages) (Kokeb et al., 2021; Adino et al., 2022). Given that migration is endogenous and linked with plot and household factors, these mean differences should be interpreted as associations rather than causes. To assess causal influences, one must use techniques like instrumental variables, endogenous switching regression, or difference-in-differences when panel data is available to account for both observed and unobserved confounders.

As indicated in Table 5, the average land size per household in the study area is 1.88 hectares, which is relatively higher than the national average which is 1.15 hectares per household. The average tropical livestock unit (TLU) is 5.35 per household in the study area, while the average TLU for households with and without migrants is 5.36 and 5.34, respectively. The mean dependency ratio is 0.62 in the study area. High dependency ratio reduces household consumption per capita and makes households more vulnerable to shocks. As shown in Table 5, the mean dependency ratio of households without migrants is higher compared to households with migrant members. The implication is that the dependency ratio may tend to reduce participation in rural out-migration.

Econometric results

As specified in the methodology section, the multinomial endogenous switching model estimates two equations simultaneously: the participation equation and the outcome equation. The first stage regression uses the multinomial logistic model and is employed to examine the drivers of rural out-migration in the Oromia region. The dependent variable in the first stage regression is a nominal variable with three categories: non-migrant households ($J = 0$), rural-urban migrants ($J = 1$), and international migrants ($J = 2$). The base category is non-migrant households. The Wald test result is statistically significant at 1% implying that the survey data fit the model well.

As shown in Table 6, the coefficient of age square is favorably and significantly correlated with both rural-urban and international migration, but the estimated coefficient of household head

	Mean				Std. Error	t-value
	Grand Mean	With Migrants	Without Migrants	Mean Difference		
Age of Household Head	48.73	51.31	44.84	6.48	1.21	5.34***
Adult Equivalent	5.51	6.03	4.72	1.31	0.22	5.94***
Dependency Ratio	0.62	0.51	0.77	-0.26	0.07	-3.99***
Education	3.98	3.37	4.85	-1.48	0.40	-3.70***
Family Size	6.64	7.23	5.76	1.46	0.27	5.37***
Land Size	1.88	1.9	1.86	0.04	0.19	0.20
Tropical Livestock Unit	5.35	5.36	5.34	0.03	0.40	0.06
Asset per AE	8129.3	6048.01	11271.54	-5223.53	4130.07	-1.26
Productivity, wheat	1597.22	1725.27	1427.413	297.86	189.66	1.57
Productivity, teff	713.39	778.01	602.24	175.77	84.35	2.08**

Note: DPR and AE refer to dependency ratio and adult equivalent, respectively
Source: Authors' computation, 2023

Table 5: Mean difference test results for some continuous variables.

age is negatively correlated and significantly correlated with both. According to this non-linear relationship, the likelihood of having migrant household members rises with increasing age, but it is lower for younger household heads. Ajaero and Onokala (2013) in Nigeria and Wondimagegnhu and Zeleke (2017) in Ethiopia also noted this pattern, pointing out those middle-aged families with the networks and means to facilitate migration are more likely to migrate. In a similar vein, the head's years of education have a positive and statistically significant impact when it comes to rural-urban migration. This supports the findings of Tegegne and Penker (2016) and de Brauw et al. (2014), who show that higher education levels raise the likelihood of migration by improving access to information and urban employment opportunities and by influencing preferences for urban amenities and services that aren't available in rural areas. There are also notable differences in migration participation by the household head's gender. According to this study, migration is more likely to occur in rural households headed by women than by men. Tacoli and Mabala (2010) have observed similar findings, attributing this to women's restricted access to land, livestock, extension services, and non-farm revenue streams.

The irrigation dummy coefficient is negative and statistically significant at a 5 percent level of significance. This could be because the use of irrigation by rural households increases their farm income, builds resilience to poverty and vulnerability, and could reduce their likelihood of participating in rural out-migration. The study also revealed that family size increases the likelihood of participation in rural-out migration. A similar study is that of Wondimagegnhu and Zeleke (2017), who found that the probability of a household engaging in rural-out migration increases with family size. This suggests that a family with a large size may be forced to share resources like land and other agricultural assets, which may not be enough to make their living.

Multinomial Logistic Regression				Pseudo R square = 33.2		
Log Pseudolikelihood = -233.8592				Wald chi^2 (48) = 132.181		
Number of observations: 384				Prob > chi^2 = 0.000		
Independent Variables	Rural-Urban Migration			International Migration		
	Coefficient (Std. Error)	t-value	p-value	Coefficient (Std. Error)	t-value	p-value
Male	-1.273 (0.65)	-1.94	0.052	-1.607 (0.702)	-2.29	0.022
Age	-0.215 (0.128)	-1.69	0.092	-.323 (0.112)	-2.88	0.004
Age Square	0.003 (0.001)	2.24	0.025	0.003 (0.001)	3.15	0.002
Dependency Ratio	-0.785 (0.384)	-2.04	0.041	-1.522 (0.313)	-4.86	0.000
Education	0.26 (0.177)	1.46	0.143	.368 (0.134)	2.74	0.006
Education Square	-0.053 (0.019)	-2.76	0.006	-0.031 (0.012)	-2.59	0.010
Family Size	0.264 (0.099)	2.67	0.008	0.497 (0.089)	5.58	0.000
Oromo	0.209 (0.674)	0.31	0.757	1.432 (0.831)	1.72	0.085
Land Size	-0.381 (0.121)	-3.15	0.002	-0.363 (0.148)	-2.46	0.014
Extension Visits	0.009 (0.006)	1.54	0.122	0.003 (0.006)	0.44	0.657
NFP	-0.716 (1.064)	-0.67	0.501	0.789 (0.798)	0.99	0.323
TLU	-0.017 (0.054)	-0.31	0.760	-0.12 (0.058)	-2.07	0.039
Drought	-0.615 (0.403)	-1.52	0.128	0.166 (0.352)	0.47	0.638
PSNP	-1.381 (0.78)	-1.77	0.077	-1.063 (0.547)	-1.94	0.052
Arsi	-0.279 (0.58)	-0.48	0.633	2.77 (0.565)	4.91	0.000
Jimma	1.155 (0.453)	2.55	0.011	1.697(0.551)	3.08	0.002
Irrigation	-1.418 (0.62)	-2.30	0.022	-0.959 (0.558)	-1.72	0.086
Muslim	-0.807 (0.462)	-1.75	0.080	1.088 (0.481)	2.26	0.024
Return Migrants	0.182 (0.096)	1.90	0.057	0.20(0.093)	2.15	0.032
Land Renting	0.923 (0.526)	1.76	0.079	0.163 (0.449)	0.36	0.717
Constant	3.915 (3.447)	1.14	0.256	2.571 (2.939)	0.88	0.382

Note: Values in the parentheses are standard errors

Source: Own survey, 2023

Table 6: Estimation results of the drivers of rural out-migration in Oromia region.

As a result, members of the family engage in rural-urban or international migration. Therefore, a large family size is one pushing factors of rural-urban and international migration in Oromia.

The result in Table 6 also revealed that being followers of the Muslim religion reduces participation in rural-urban migration, while it increases participation in international migration. Putting it differently, followers of the Muslim religion are more likely to participate in international migration compared to followers of other religions. This indicates that religion is both a push and a pull factor in migration. More importantly, it shows that Muslim migrants are more likely to be pulled by Muslim countries in migrant-receiving destination areas (Ahsan, 2022). Agricultural assets like rural land size and number of tropical livestock units reduce household participation in migration in the study area. A similar finding was that of Wondimagegnhu and Zeleke (2017), who found number of tropical livestock units that the household owns reduces the probability of a household participating in rural-urban migration in Ethiopia. Abdullah (2022) has also found that land size negatively affects rural out-migration in Bangladesh. This suggests that better access to agricultural assets such as land, livestock, and capital in rural areas has the potential to reduce the rural out-migration in the Oromia region.

Besides, the presence of return migrants in a household is positively and statistically significantly associated with both rural–urban and international migration, ceteris paribus ($p < 0.05$). This finding is consistent with recent evidence indicating that returnees bring back financial capital, human capital, and social capital resources that reduce migration costs and enhance migration propensity. For example, return migrants return with accumulated savings, skills, and networks that finance and facilitate new migration opportunities (Bossavie et al. 2024) and promote entrepreneurship and urban engagement in their home communities (Yu et al., 2022). Comparing the variation across zones, sample households from Jimma and Arsi are more likely to participate in international migration compared to those from the North Shewa zone, as indicated in Table 6. The existence of dependents in a family (indicated by DPR) reduces participation in migration in the study area. That means families with unproductive members are less likely to participate in rural out-migration mainly because of the extra responsibility that the dependents pose to the family. This finding is in agreement with a study conducted by Zakir (2016) found a negative and significant relation between the dependency ratio and the rural-urban migration of households.

Likewise, participation in a productive safety net program also reduces participation in migration. This may be because participation in a productive safety net program might build the resilience of users to poverty and food insecurity and reduce their participation in rural out-migration. More so, the findings from the KIIs and FGDs showed that limited access to agricultural land, large family size, lack of employment, and credit constraints by rural youth are the major pushing factors of rural out-migration in the study area. They added that peer pressure, brokers, the presence of return migrants in the village, underage marriage, conflicts in the family, and divorce are also contributing to rural out-migration. Further, the participants in KIIs and FGDs also reported that rural youth are not interested in agricultural activities, and rural traditional life, and they are rather attracted by public services in urban areas.

The second stage regression of the multinomial endogenous switching model quantified the impact of participation in rural out-migration on output of wheat or teff per hectare in kilograms, and the results are presented in Table 7. As reported in Table 7, the actual mean output of wheat is 1,905 and 1,641 kilograms per hectare for households with rural-urban migrants and international migrants, respectively. But the counterfactual mean output of wheat is 1,564 and 934 kilograms per hectare for households with rural-urban migrants and international migrants, respectively. Accordingly, the conditional average treatment effects on treated (ATT) of wheat for households with rural-urban migrants and international migrants are 341 and 707 kilograms per hectare. This suggests that participation in international migration significantly increases wheat output per hectare in the study area, though the impact is also positive for rural-urban migration. This could be due to the fact that international migrants tend to send larger remittances, which enable households to invest more substantially in improved agricultural inputs, technology, and labor

Similarly, the actual mean output of teff is 996 and 820 kilograms per hectare for households with rural-urban migrants and international migrants, respectively. But the counterfactual mean output of teff is 494 and 563 kilograms per hectare for households with rural-urban migrants and international migrants, respectively.

Outcomes	Choices	Decision to Participate in Migration		Average Treatment Effect on Treated (ATT)
		Participation	Non-participation	
		Actual	Counterfactual	
Wheat Productivity	Rural-urban	1904.90	1563.62	341.28 (215.02)^{c}
	International	1641.42	934.20	707.21 (107.29)^{a}
Teff Productivity	Rural-urban	996.01	493.96	502.05 (84.92)^{a}
	International	819.82	562.77	257.04 (45.68)^{a}
Heterogeneity Effects		**BH_1**	**BH_0**	**TH**
Wheat Productivity	Rural-urban	-963.86 (305.83)^{a}	166.99 (128.73)	-1130.86 (306.88)^{a}
	International	-3.274 (141.93)	-462.41 (91.74)	459.13 (139.02)^{a}
Teff Productivity	Rural-urban	839.19 (142.82)^{a}	-96.20 (31.82)^{a}	935.39 (144.92)^{a}
	International	93.70 (62.14)^{a}	-27.38 (28.56)^{a}	121.09 (66.96)^{c}
Falsification Test Result: F – statistics = 1.61			Probability > F = 0.206	

Standard errors are in parentheses. $^{a, b,}$ and ^{c} denote significance level at 1%, 5%, and 10%.
Source: Authors' Computation, 2023

Table 7: Estimation results of the impact of migration on productivity of wheat and teff.

Accordingly, the conditional average treatment effects on treated (ATT) of teff for households with rural-urban migrants and international migrants are 502 and 257 kilograms per hectare, and significant at the 1 percent level. Despite the fact that the estimated yield effects are significant, some values like for wheat are unusually high when compared to typical input related gains. These effects, therefore, should be interpreted cautiously and viewed as indicative rather than exact, reflecting both the empirical context and model assumptions.

This suggests that the substitution of labor and capital between the rural agricultural sector and urban non-agricultural sectors promotes the productivity of wheat and teff producers in the Oromia region. This could be because the transfer of capital in the form of remittances from urban areas to capital-constrained rural areas will enhance agricultural production by lessening the credit constraints and the risk aversion level of households. This finding supports the credit and risk hypotheses of the new economics labor migration theory, which claims that migration increases agricultural investment and productivity by reducing the risk aversion level and the credit constraints of migrant-sending rural households. Furthermore, since households with limited land size, tropical livestock unit, fragmented land size, and large family size participated in rural out-migration in the study area, as evidenced from the descriptive results, participation in rural out-migration may not necessarily lead to a reduction in agricultural production via the lost labor effect. Moreover, about half of the migrants in the study area are female migrants, and the opportunity costs of female migrants to agricultural production are lower compared to male migrants since male family members are more likely to participate in agricultural activities. Yet, migration may increase output per hectare of remittance-receiving rural households through the remittance channel if rural households spend remittances from migrants on the purchase of agricultural inputs and livestock. This result, therefore, implies that migration promotes crop productivity through the remittance channel by lessening the liquidity constraint and increasing agricultural investment in the region.

Generally, remittances give families access to financial resources that can improve total production capacity, reduce liquidity constraints, and encourage investments in agricultural inputs. These advantages could, however, come at the price of young, talented people leaving the rural areas, which would leave less productive labor available for use in the origin areas. According to this labor-remittance trade-off, remittances can boost productivity by providing financial capital, but these benefits may be somewhat countered by the concurrent loss of human capital, especially in labor-intensive farming systems where having physically fit workers on hand is essential.

In sum, while the human capital theory of migration considers the expected wage differential between rural areas and urban areas as the primary cause of rural out-migration, the new economics labor migration theory insists that rural out-migration is mainly caused by the inefficiency in capital and insurance markets in rural areas. Besides, the new economic labor migration theory also assumes that migration affects the welfare and production of migrant-sending areas via two channels: the lost labor channel and the remittance channel. Hence, rural outmigration is a two-handed transaction, and it gives with one hand and takes with the other hand. The impact of migration on welfare and production of migrant-sending households, therefore, depends on the relative strength of the remittance effect and the lost labor effect. However, in this study, remittances from migrants have been shown to increase agricultural investments, food expenditure, non-food expenditure, and kilocalories per adult equivalent per day of migrant-sending households in Oromia regional state. This result, therefore, supports the new economics labor migration theory, which assumes that participation in migration increases the welfare of migrant-sending households.

The impact of rural out-migration on agricultural output in rural areas that send migrants has been the subject of several prior studies, with mixed results. On the one hand, Khanal et al. (2015), Abdi (2021), Imran et al. (2016), Goldsmith et al. (2017), and Adaku (2019) report a negative and considerable impact, which they frequently attribute to a lack of labor and a decrease in household capacity to oversee farm operations. Conversely, Odozi et al. (2020), Mesfin et al. (2021) and Yu et al. (2024) indicate a positive and significant effect, indicating that migrant workers' remittances can improve productivity, allow the adoption of new technologies, and increase access to agricultural inputs. Likewise Bassie et al. (2022), using cross-sectional data from China, conclude that participation in migration can promote crop productivity, underscoring that migration’s impact on agriculture varies by context, resource availability, and the way remittances are utilised.

The base heterogeneity for participants (BH_1), the base heterogeneity for non-participants (BH_0) and the transitional heterogeneity (TH) effects are computed, and the results are presented in Table 7. The base heterogeneity for participants (BH_1) is the difference between the output per hectare of participants minus the output per hectare of non-participants if they had participated. But the base heterogeneity for non-participants (BH_0) is the difference between the output per hectare of participants if they had not participated minus the output per hectare of non-participants. Hence, a positive value of TH suggests that the productivity-enhancing impact of migration is higher for participants compared to non-participants, whereas the negative value of TH implies that the productivity-enhancing impact of migration is higher for non-participants had they participated in migration compared to participants. This study utilized religion and dependency ratio as exclusion restriction variables. The validity of these instruments was evaluated using a falsification test, the results of which confirmed their appropriateness for the analysis. Once migration choices are taken into consideration, religion has little direct impact on household productivity; however it may influence migration decisions by influencing social networks, cultural norms, and migration preferences. Similarly, return migrants capture exposure to migration networks and information flows that enable out-migration, but apart from their impact on migration behavior, they have no direct impact on production levels.

Conclusion

Rural out-migration involves the flow of labor from the rural agricultural sector to the urban non-agricultural sectors and the transfer of cash in the form of remittances from urban receiving areas to rural sending areas. Migration does not occur in a vacuum; it affects agricultural practices in regions that transfer migrants via labor-loss and remittance channels. However, empirical evidence on the impact of rural out-migration on agricultural production in Ethiopia remains limited. By using cross-sectional survey data from 384 households and applying the New Economics of Labour Migration (NELM) theory along with a multinomial endogenous switching model, this study helps close this gap by investigating the factors that influence rural out-migration and its impact on productivity among wheat and teff producers in the Oromia region.

According to the descriptive findings, the percentage of migrants in Oromia's total urban population increased significantly from 27.2% in 1999 to 49.2% in 2021. This suggests that people migrate from rural areas with comparatively better job creation and lower unemployment to urban areas with high unemployment and few jobs. This trend is consistent with the human capital theory of migration, which holds that improved access to public amenities like clean water, power, transportation, healthcare, and education attracts migrants in addition to job opportunities.

The econometric findings show that while larger family sizes, female-headed households, and older household heads increase the likelihood of migration, land size, irrigation use, livestock holdings, dependency ratio, participation in productive safety net programs, and education level decrease that likelihood. More crucially, the study shows that the productivity of wheat and teff producers is greatly increased by both international and rural-urban migration. In particular, teff yields rose by 502.05 kg/ha and 257.04 kg/ha, respectively, and wheat yields rose by 341.28 kg/ha and 707.21 kg/ha, respectively, as a result of rural–urban and international migration. These results provide support for the NELM theory's credit and risk assumptions, which contend that by reducing risk aversion and loosening credit constraints, migration might boost agricultural investment.

From a policy standpoint, the findings imply that, as long as remittances are efficiently directed toward profitable agricultural ventures, the benefits of rural out-migration on agricultural output may exceed any possible labor shortages. This emphasizes how crucial it is to have laws that make it easier to use remittances for technology adoption and agricultural inputs. Enhancing land access and tenure security, increasing the availability of cheap finance, upgrading irrigation systems, offering pre-migration financial literacy and agribusiness training, and more are examples of practical initiatives. Furthermore, enhancing rural public services and establishing supplementary off-farm job prospects can guarantee that migration choices are advantageous and strategic rather than solely motivated by misery. In sum, these results have wider implications for rural development plans in Ethiopia and other comparable situations.

Rather than perceiving rural out-migration only as a loss of labour, it should be acknowledged as a potential engine of agricultural transformation if supported by enabling institutions, targeted policies, and mechanisms that enhance the productive use of remittances.

This study does, however, have some limitations. First, because it is based on cross-sectional data, dynamic impacts over time are not captured. Second, results are not as comparable or generalizable across nations due to the diversity of socioeconomic and agro-ecological environments. Third, without accounting for factors like labor, seed variety, market accessibility, or input quality, productivity is expressed in kilos per hectare. Panel or longitudinal data should be used in future studies to better understand the temporal dynamics of migration's impacts on agriculture. In order to offer a more profound understanding of the ways in which migration impacts agricultural practices, labor allocation choices, and remittance utilization, researchers may also utilize a qualitative method.

Acknowledgments

This study was conducted as part of the collaborative work between the Oromia Plan and Development Commission (OPDC) and the Ethiopian Economics Association (EEA) with financial support from OPDC. The authors would like to express gratitude to the participants of the Validation Workshop organized by the EEA in Ethiopia for their valuable insights and feedback.

Corresponding author:
Fassil Eshetu
Department of Economics, Arba Minch University
Arba Minch, Ethiopia
Email: bekatfech@gmail.com

References

[1] Abdi, M. (2021) "Remittance and Poverty in Somalia: Propensity Score Matching Approach", *Journal of Research in Economics*, Vol. 5, No. 1, pp. 50-68. ISSN 2636-8307. DOI 10.29228/jore.3.

[2] Abeje, A. (2021) "Causes and Effects of Rural-Urban Migration in Ethiopia: A Case Study from Amhara Region", *African Studies,* Vol. 80, No. 1, pp. 77-94. ISSN 0002-0184. DOI 1010.1080/00020184.2021.1904833.

[3] Abire, B. and Sagar, G. (2016) "Determinant Factors of Illegal Migration to South Africa and Its Impacts on the Society in Case of Gombora District, Hadiya Zone in Ethiopia", *IOSR Journal of Mathematics*, Vol. 12, No. 3, pp. 51-65. E-ISSN 2278-5728.

[4] Adaku, A. (2019) "The Effect of Rural-Urban Migration on Agricultural Production in the Northern Region of Ghana", *Journal of Agricultural Science and Applications*, Vol. 2, No. 4, pp. 193-201. ISSN 2227-6475.

[5] Adino, K., Alamirew, B. and Goshu, D. (2022) "The Impact of Rural–Rural Migration on Crop Production in North-western Ethiopia: An Application of Endogenous Switching Regression Model", *Ethiopian Journal of Development Research*, Vol. 41, No. 2, pp.1-24. ISSN 0378-0813.

[6] Adugna, G. (2019) "Migration Patterns and Emigrants' Transnational Activities: Comparative Findings from Two Migrant Origin Areas in Ethiopia", *Comparative Migration Studies*, Vol. 1, pp. 1-28. ISSN 2214-8590. DOI 10.1186/s40878-018-0107-1.

[7] Agza, M., Alamirew, B. and Shibru, A. (2023) "Determinants of Rural-Urban Migration and its Impact on Migrant-Sending Households' Livelihood Security in Gurage zone, Ethiopia", *Cogent Social Sciences*, Vol. 9, No. 1. E-ISSN 2331-1886. DOI 10.1080/23311886.2023.2190253.

[8] Ahsan, U., Ahmed, S. and Arju, A. (2022) "Religion in the Age of Migration", Politics, Religion and Ideology, pp. 1-26. E-ISSN 2156-7697. DOI 10.1080/21567689.2022.2057476.

[9] Ajaero, C. K. and Onokala, P. C. (2013) "The Effects of Rural–Urban Migration on Rural Communities of Southeastern Nigeria", *International Journal of Population Research*, ISSN 2090-4029. DOI 10.1155/2013/610193.

[10] Alarima, C. I. (2019) "Factors Influencing Rural-Urban Migration of Youths in Osun State, Nigeria", *Agro-Science*, Vol. 17, No. 3, pp. 1-34. E-ISSN 3023-7017. DOI 10.4314/as.v17i3.6.

[11] Al-Maruf, A., Kanak Pervez, A. K. M., Sarker, P. K., Rahman, M. S. and Ruiz-Menjivar, J. (2022) "Exploring the Factors of Farmers' Rural–Urban Migration Decisions in Bangladesh", *Agriculture*, Vol. 12, pp. 1-16, E-ISSN 2077-0472. DOI 10.3390/agriculture12050722.

[12] Bassie, H., Sirany, T. and Alemu, B. (2022) "Rural-Urban Labor Migration, Remittances, and Its Effect on Migrant-Sending Farm Households: Northwest Ethiopia", *Advances in Agriculture*, E-ISSN 2314-7539. DOI 10.1155/2022/4035981.

[13] Bossavie, L., Görlach, J.-S., Ozden, Ç. and Wang, H. (2024) "Capital Markets, Temporary Migration and Entrepreneurship: Evidence from Bangladesh", *World Development,* Vol. 176, p. 106505. ISSN 0305-750X. DOI 10.1016/j.worlddev.2023.106505.

[14] Bourguignon, F., Fournier, M. and Gurgand, M. (2007) "Selection Bias Corrections Based on the Multinomial Logit Model", *Journal of Economic Surveys*, Vol. 21, No. 1, pp. 174-205. E-ISSN 1467-6419. DOI 10.1111/j.1467-6419.2007.00503.x.

[15] de Brauw, A. and Giles, J. (2017) "Migrant Opportunity and the Educational Attainment of Youth in Rural China", *Journal of Human Resources*, Vol. 52, No. 1, pp. 272-311. E-ISSN 1548-8004, ISSN 0022-166X.

[16] de Brauw, A., Mueller, V. and Woldehanna, T. (2014) "Does internal migration improve overall well-being in Ethiopia?", *Journal of African Economies*, Vol. 27, No. 3., pp. 347-365. E-ISSN 1464-3723. DOI 10.1093/jae/ejx026.

[17] Bryan, G., Chowdury, S. and Mobarak, M. (2014) "Underinvestment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh", *Econometrica*, Vol. 82, No. 5, pp. 1671-1748. E-ISSN 1468-0262.

[18] Ethiopian Statistical Service (ESS). (2018) "*Ethiopian Population Projection for All Regions*", Addis Ababa: Ethiopian Statistical Agency. [Online]. Available at: https://ess.gov.et/download/population-projection-of-ethiopia-for-all-regions-at-wereda-level-from-2014-2017-1/ [Accessed: Sept. 17, 2024].

[19] Chamberlin, J., Jayne, T. S. and Sitko, N. J. (2020) "Rural in-migration and agricultural development: Evidence from Zambia", *Agricultural Economics*, Vol. 51, No.4, pp.491–504, ISSN:0169-5150, ISSN 0139-570X. DOI 10.1111/agec.12567.

[20] Chen, D., Gao, H., Luo, J. and Ma, Y. (2020) "The effects of rural–urban migration on corporate innovation: Evidence from a natural experiment in China", *Financial Management*, Vol. 49, No. 2, pp. 521-545. E-ISSN 1755-053X. ISSN 0046-3892. DOI 10.1111/fima.12280.

[21] Cochran, W. G. (1963) "*Sampling Techniques*", 2nd ed., New York: John Wiley and Sons, Inc., ISBN 978-0471162407.

[22] Deb, P. and Trivedi, P. (2006) "Maximum simulated likelihood estimation of a negative binomial regression model with multinomial endogenous treatment", *Stata Journal*, Vol. 6, pp. 246-255. ISSN 1536-867X.

[23] Fassil, A. E. and Mohammed, B. (2017) "Dynamics and determinants of rural-urban migration in Southern Ethiopia", *Journal of Development and Agricultural Economics*, Vol. 9, No.12, pp. 328-340. ISSN 2006-9774. DOI 10.5897/JDAE2017.0850.

[24] Feliciano, C. (2020) "Immigrant Selectivity Effects on Health, Labor Market, and Educational Outcomes", Annual Review of Sociology, Vol. 46, No. 1, pp. 315-334. E-ISSN 1545-2115. DOI 10.1146/annurev-soc-121919-054639.

[25] Fente, M. F. and Abebe, B. G. and Woldeamanuel, M. G. (2023) "Exploring the nexus of migration dynamics and urban expansion: Key drivers of horizontal spatial growth in Woldia Township, Ethiopia", *Frontiers in Human Dynamics*, Vol. 6. E-ISSN 2673-2726. DOI 10.3389/fhumd.2024.1517341.

[26] Food and Agriculture Organization (FAO). (2019) "*Migration Framework Migration as a choice and an opportunity for rural development*", Rome. [Online]. Available: https://openknowledge.fao.org/server/api/core/bitstreams/5f28145b-307f-4114-98fb-aafc5e4b0a21/content [Accessed: Nov. 20, 2024].

[27] de Haas, H. (2011) "Migration and development", *International Migration Review*, Vol. 33, pp. 227-264. ISSN 0197-9183. DOI 10.1111/j.1747-7379.2009.00804.x.

[28] Harris, J. and Todaro, M. (1970) "Migration, unemployment & development: A two-sector analysis", *American Economic Review*, Vol. 60, No. 1, pp.126-142. ISSN 0002-8282.

[29] Horwood, C. (2009) "*In pursuit of the southern dream: Assessment of the irregular movement of men from East Africa to South Africa*", Geneva: International Organization for Migration. [Online]. Available: https://publications.iom.int/system/files/pdf/iomresearchassessment.pdf [Accessed: Nov. 25, 2024].

[30] Hossain, M. Z., Riad, M. M. and Ahmed, J. U. (2016) "Rural–urban migration in Bangladesh and its nexus with some socioeconomic indicators at origin and destination", *International Journal of Social Relevance & Concern*, Vol. 4, No. 11, pp. 1-14. ISSN 2347-9698.

[31] Iheke, O. R., Nwaru, J. C. and Onyenweaku, C. E. (2013) "The impact of migrant remittances on the technical efficiency of arable crop farm households in South-Eastern Nigeria", *Fourth International Conference*, September 22-25, 2013, Hammamet, Tunisia 161263, African Association of Agricultural Economists (AAAE). DOI 10.22004/ag.econ.161263.

[32] ILO (International Labour Organization, 2022) "*Practical guide on developing labour migration policies*". ISBN 978 92 2 035423-0. [Online]. Available: https://www.ilo.org/sites/default/files/wcmsp5/groups/public/@ed_protect/@protrav/@migrant/documents/publication/wcms_832194.pdf. [Accessed: Dec. 14, 2024].

[33] Imran, M., Khuda, B. and Hassan, S. (2016) "Rural to urban migration & crop productivity in Pakistani Punjab", *Mediterranean Agricultural Sciences*, Vol. 29, No. 1, pp. 17-19. E-ISSN 2528-9675. [Online]. Available: https://dergipark.org.tr/en/pub/mediterranean/issue/31214/339631 [Accessed: Nov. 25, 2024].

[34] IOM (International Organization for Migration, 2024) "World Migration Report 2024", Geneva: International Organization for Migration. [Online]. Available: https://publications.iom.int/books/world-migration-report-2024 [Accessed: Nov. 25, 2024].

[35] Kaur, G. and Kaur, J. (2021) "The Impact of Rural-Urban Migration on Agriculture", *Journal of Critical Reviews*, Vol. 8, No. 3, pp. 506-510. ISSN 2394-5125. DOI 10.31838/jcr.08.03.61.

[36] Kefelegn, H. (2020) "Determinants of Youths Rural to Urban Migration in Ethiopia (Case of Addis Ababa", *Academic Journal of Research and Scientific Publishing*, Vol. 2, No. 15, pp. 5-7. ISSN 2706-6495.

[37] Khanal, U., Alam, K., Khanal, R. C. and Regmi, P. P. (2015) "Implications of out-migration in rural agriculture: A case study of Manapang village, Tanahun, Nepal", *The Journal of Developing Areas*, Vol. 49, No. 1, pp. 331-352. E-ISSN 0022-037X. DOI 10.1353/jda.2015.0012.

[38] Khatir, A. and Rezaei-Moghaddam, K. (2014) "Evidence from predictors of rural youth's migration intentions in agricultural communities: the Fars province, Iran", *Migration and Development*, Vol. 3, No. 2, pp. 219-238. E-ISSN 2163-2332. DOI 10.1080/21632324.2014.926612.

[39] Lagakos, D., Marshall, S., Mobarak, A.M., Vernot, C., Waugh, M. (2020) "Migration costs and return to migration in developing world", *Journal of Monetary Economics*, Vol. 113, pp. 138-154. E-ISSN 1873-1295, ISSN 0304-3932. DOI 10.1016/j.jmoneco.2020.03.013.

[40] Lee, E. S. (1966) "A Theory of Migration", *Demography*, Vol. 3, No. 1, pp. 47-57. ISSN 0070-3370.

[41] Lewis, W. (1954) "Economic development with unlimited supplies of labor", *Manchester School of Economic and Social Studies*, Vol. 22, pp. 139-191. ISSN 1463-6786. DOI 10.1111/j.1467-9957.1954.tb00021.x.

[42] Lucas, R. (1987) "Emigration to South Africa's Mines", *The American Economic Review*, Vol. 77, No.3, pp. 313-330. ISSN 0002-8282.

[43] Lucas, R. and Stark, O. (1985) "Motivations to remit: Evidence from Botswana", *Journal of Political Economy*, Vol. 93, No. 5, pp. 901-918. E-ISSN 1537-534X, ISSN 0022-3808.

[44] Ma, L., Chen, M., Che, X., and Fang, F. (2019) " Farmers' rural-to-urban migration, influencing factors and development framework: A case study of the village of Gansu, China", *International Journal of Environmental Research and Public Health*, Vol.16, No. 5. ISSN 1660-4601. DOI 10.3390/ijerph16050877.

[45] Massey, D. S., Arango, J., Hugo, G., Kouaouci, A., Pellegrino, A. and Taylor, J. E. (1998) "*Worlds in Motion. Understanding International Migration at the End of the Millennium*", Oxford: Clarendon Press. E-ISBN 9781383018639 DOI 10.1093/oso/9780198294429.001.0001.

[46] Mesfin, A., Bamlaku, A. and Admasu, S. (2021) "Crop producers' technical efficiency and its determinants in Gurage zone, Ethiopia: A comparative analysis using rural-urban migration as a parameter", *Cogent Social Sciences*, Vol. 7, No. 1, pp. 1-15. E-ISSN 2331-1886. DOI 10.1080/23311886.2021.1995996.

[47] Morales, G. Jr., Villaronte, R. K., Yap, M. C. (2022) "The Relationship Between Rural-Urban Migration and the Agricultural Output of the Philippines", *International Journal of Social and Management Studies (IJOSMAS)*, Vol. 3, No.1 , pp. 62-74. E-ISSN 2775-0809. DOI 10.5555/ijosmas.v3i1.88.

[48] Odozi, J., Adeniyi, O. and Yusuf, S. (2020) "Production efficiency in agriculture in Nigeria: Do migrant remittances matter?", *Ekonomika poljoprivrede*, Vol. 67, No. 2, pp. 2315-327. ISSN 2334-8453. DOI 10.5937/ekoPolj2002315O.

[49] Ravenstein, E. G. (1885) "The laws of migration", *Journal of the Statistical Society of London*, Vol.48, No. 2, pp. 167-235. ISSN 0952-8385. DOI 10.2307/2979181.

[50] Sauer, J., Gorton, M. and Davidova, S. (2013) "Migration and Agricultural Efficiency: Empirical Evidence for Kosovo", *53rd Annual Conference*, Berlin, Germany, September 25-27, 2013 156098, German Association of Agricultural Economists (GEWISOLA).

[51] Sennuga, S. O., Barnabas, T. M., Alabuja, F. O., Dokubo, E. M. and Bankole, O.-L. (2023) "Effect of rural-urban migration among the youths and its impacts on agricultural development in Kuje Area Council, Abuja, Nigeria", *Science and Technology*, Vol. 4, No. 2, pp.12-27. ISSN 2709-0854.

[52] Stark, O. and Bloom, D. E. (1985) "The new economics of labor migration", *The American Economic Review*, Vol. 75, No. 2, pp. 173-178. E-ISSN 1944-7981, ISSN 0002-8282.

[53] Tacoli, C. and Mabala, R. (2010) "Exploring mobility and migration in the context of rural–urban linkages: why gender and generation matter", *Environment and Urbanization*, Vol. 22, No. 2, pp. 389-395. E-ISSN 1746-0301. DOI 10.1177/0956247810379935.

[54] Taylor, J. E. (2001) "Migration: New Dimensions and Characteristics, Causes, Consequences and Implications for Rural Poverty", In Stamoulis, K. (ed.) "*Food, Agriculture and Rural Development: Current and Emerging Issues for Economic Analysis and Policy Research*", FAO. p. 167-201. ISBN 92-5-104566-6.

[55] Taylor, J. E. and Wyatt, T. J. (1996) "Shadow Value of Migrant Remittances, Income, and Inequality in a Household-Farm Economy", *The Journal of Development Studies*, Vol. 32, No. 6, pp. 899-912. E-ISSN 1743-9140, ISSN 0022-0388. DOI 10.1080/00220389608422445.

[56] Tegegne, A. D. and Penker, M. (2016) "Determinants of rural out-migration in Ethiopia", *Demographic Research*, Vol. 35, No. 1, pp. 1011-1044, ISSN 1435-9871. DOI 10.4054/DemRes.2016.35.34.

[57] Todaro, M. P. (1969) "A Model of Labor Migration and Urban Unemployment in Less Developed Countries", *American Economic Review*, Vol. 59, No. 1, pp. 138-148. ISSN 0002-8282.

[58] UN (United Nations). (2016) "*The World's Cities in 2016*", United Nations. Book Series: Statistical Papers - United Nations (Ser. A), Population and Vital Statistics Report. 27 p. ISBN (PDF): 9789210582766. DOI 10.18356/8519891f-en.

[59] Wondimagegnhu, B. and Zeleke, M. (2017) "Determinants of Rural Out-Migration in Habru District of Northeast Ethiopia", *International Journal of Population Research*, pp. 1-8. E-ISSN 2090-4037. DOI 10.1155/2017/4691723.

[60] World Bank (2023) "*Migrants, Refugees, and Societies*", World Development Report 2023, Washington, DC: World Bank. [Online]. Available: https://www.worldbank.org/en/publication/wdr2023 [Accessed: March 10, 2024].

[61] World Bank (2024). "*Personal remittances, received (% of GDP) - Ethiopia*", World Bank Open Data", [Online]. Available: https://data.worldbank.org/indicator/BX.TRF.PWKR.DT.GD.ZS?locations=ET [Accessed: March 10, 2024].

[62] WFP (World Food Program) (2015) "*Global Food Security Update*", World Food Programme, Iss. 17. [Online]. Available: https://sustainabledevelopment.un.org/content/documents/1720Global%20Food%20Security%20Update.pdf [Accessed: March 10, 2024].

[63] Witten, M. (2007) "The Protection of Land Rights in Ethiopia", *Afrika Focus*, Vol. 20, No. 1-2, pp. 153-184. [Online]. Available: https://openjournals.ugent.be/af/article/61094/galley/185499/view/ [Accessed: Feb. 22, 2024].

[64] Yoshino, N., Taghizadeh-Hesary, F., Otsuka, M. (2017) "International remittances and poverty reduction: Evidence from Asian Developing Countries", *Journal of Comparative Asian Development*, Vol.17, No. 2, pp. 21-42. E-ISSN 2150-5403.

[65] Yu, L., Li, Z. and Liu, D. (2024) "Return-Migrant Urbanisation in Inland China: The Case of Hubei Province", *Land*, Vol. 13, No. 2, ISSN 2073-445X. DOI 10.3390/land13020190.

[66] Zahonogo, P. (2011) "Migration and agricultural production in Burkina Faso", *African Journal of Agricultural Resource*, Vol. 6, No. 7, pp. 1844-1852. ISSN 1991-637X. DOI 10.5897/AJAR10.1003.

Volume XVII Number 4, 2025

Farmer Involvement in Irrigation Agriculture: Evidence from the Anambra-Imo River Basin Irrigation Scheme, Nigeria

Ubon Essien[1], Okwudili Ibeagwa[1], Igwe Ukoha[1], Sylvanus Ben[2], Onyekachukwu Ejike[3]

[1] Department of Agricultural Economics, Federal University of Technology, Owerri, Nigeria

[2] Department of Agricultural Technology, Akwa Ibom State College of Science and Technology, Ikono, Nigeria.

[3] Department of Agricultural Economics, Federal University of Technology, Owerri, Nigeria

Abstract

This study was conducted in the Anambra catchment of the Anambra-Imo River Basin Development Authority(AIRBDA), Nigeria. The aim was to analyse the involvement of farmers in irrigation agriculture as a key component of public agricultural project performance. A multi-stage sampling procedure was adopted in selecting ninety(90) farmers from the catchment of the AIRBDA. Descriptive statistics provided initial insight into operational and structural characteristics, while relevant visualizations were produced using Python and Excel. The Logit estimate identified factors influencing farmers' involvement in the irrigation schemes, thereby offering empirical evidence relevant for project appraisal and management. Results showed that 15.6% of farmers reported non-participation, while about 84.4% were active participants in the scheme. The estimated model reported a Wald chi² of 39.65 and a log pseudolikelihood of –281.37084. Farm experience, household size, major occupation, farm income, and membership in the Water Users Association (WUA) significantly influenced farmers' involvement in the irrigation scheme. It recommends strengthening Participatory Irrigation Management (PIM) systems, whereby farmers manage routine water allocation, while the River Basin management provides technical oversight, with a member of the Water Users Association as a part of its team.

Keywords

Irrigation scheme, project, involvement, farmers, logit model, Anambra-Imo-River-Basin.

Essien, U., Ibeagwa, O., Ukoha, I., Ben, S. and Ejike, O. (2025) "Farmer Involvement in Irrigation Agriculture: Evidence from the Anambra-Imo River Basin Irrigation Scheme, Nigeria", *AGRIS on-line Papers in Economics and Informatics*, Vol. 17, No. 4, pp. 57-66. ISSN 1804-1930. DOI 10.7160/aol.2025.170405.

Introduction

The River Basin Development Authorities (RBDAs) play a pivotal role in the agricultural development of every country, particularly through its irrigation project investments. As a public agricultural project, they are crucial in enhancing agricultural productivity, mitigating the effects of seasonal variations in rainfall, and ensuring food security (Abah and Nankiel, 2019; Christopher, 2016). However, the effectiveness of these irrigation schemes depends in part on the participation of the farmers who utilize them. Understanding the dynamics of farmer involvement in these schemes is essential for optimizing outcomes and ensuring sustainable agricultural development. Various factors, ranging from socio-economic, structural to institutional and environmental factors influence the engagement of the farmers in the irrigation project of the Anambra Imo River Basin Development Authority (AIRBDA).

One significant aspect influencing farmer involvement is the availability of the Water Users Associations (WUAs). These associations, which serve as cooperatives, are platforms for collective decision-making, resource management, and conflict resolution among farmers utilizing irrigation facilities in the River Basin Development Authorities (Aarnoudse et al., 2019). The effectiveness of the WUAs in fostering farmer participation hinges on factors such as leadership quality, transparency in decision-making processes, and equitable distribution of water resources.

Further, many farmers are poor and subsistent, therefore the cost of irrigation services provided by the AIRBDA can influence farmers' decisions to participate in the schemes and deter small-scale farmers with limited financial resources from accessing irrigation facilities, even if those farmers belong to the Water User Associations. It is equally interesting to note that regardless of the long practice of irrigation farming in the River Basin Development Authorities, most farmers still find it difficult to identify the irrigation type that is suitable for a particular crop. This education or awareness creation should be the role of the extension agents (Raji et al., 2024) or technical staff of the River Basins. Sadly, access to extension services tend to be limited.

There is equally limited access to water in some of the projects. This development has hampered all year-round production in some sites. Despite the huge amount of funds dedicated to the irrigation scheme and the efforts of the River Basin Development Authority management, farmer involvement remains a problem in most areas. Although a large proportion of the farmers under the scheme have many years of farming experience, quite a number of them depend on seasonal production occasioned by rainfall.

This research, therefore, aims at identifying the types of irrigation systems available in the area, the level of involvement of the farmers in these schemes, and the factors that affect farmer involvement. These issues are pertinent to providing the direction for this study and pivotal to developing policy options that would be beneficial to the River Basin and the farmers.

Materials and methods

This study was carried out in the Anambra catchment of the Anambra-Imo River Basin Development Authority, AIRBDA Nigeria. Anambra is in the Southeastern part of Nigeria, with a projected population of about 5.95 million (National Beareau of Statistics (NBS) (2022).

It is on the latitude (6.2209° N) and longitude (6.9370° E) of the equator. It has a tropical wet and dry season, with a yearly average temperature of 28.99°C (84.18°F). Anambra typically receives about 2,553mm/year mean precipitation and has about 150-180 rainy days annually (Omoja et al., 2021). Agriculture is an important part of the economy of Anambra. The AIRBDA derived its name from the Anambra River and Imo River, which flow through the area and are tributaries of the famous River Niger.

A multi-stage sampling technique was used for the study. In the first stage, Ayamelum Local Government Area (LGA), which is the host L.G.A. of the AIRBDA irrigation scheme, was purposely selected because of the irrigation activities of the River Basin that take place in this region. In the second stage, three communities—Omor, Umumbo, and Umuerum—were purposely selected from the Ayamelum LGA due to a large proportion of farmers in the area, who are involved in various types of cropping activities. The third stage involved the random selection of 30 crop farmers from each of the three communities, giving a total of 90 farmers, who were investigated in this study. The list of the crop farmers was obtained from the management of the AIRBDA.

The data obtained were analyzed, using descriptive statistics such as means, frequencies, and percentages. Specifically, doughnuts, clustered bars, and boxplots were used to visualize the data and draw insights on the level of involvement, type of irrigation used by farmers, and total and irrigable land cultivated by the farmers.

Also, the logistic regression model was employed to analyze involvement in irrigation agriculture by the farmers.

The implicit functional form of the logit model employed is stated, as:

$$Y = \text{Ln}\ (Pi/1\text{-}Pi) = \alpha_0 + \alpha_1 X_1 + \ldots + \alpha_k X_k + ei \quad (1)$$

Explicitly, the model is specified as:

$$\begin{aligned} Y = Ln(Pi/1\text{-}Pi) &= \alpha_0 + \alpha_1 AGE + \alpha_2 EXT + \\ &+ \alpha_3 SIZ + \alpha_4 EDU + \alpha_5 EXP + \alpha_6 CST + \\ &+ \alpha_7 HHS + \alpha_8 MRJ + \alpha_9 FINC + \alpha_{10} WUA \quad (2) \end{aligned}$$

Where Y is the dichotomous dependent variable, which can be explained as $Y = 1$ if the farmer is involved in irrigation agriculture and 0 otherwise. The *Xs* are as defined below:

- *AGE* = Age of farmer (measured in number of years a farmer has lived)
- *EXT* = Extension agent (dummy, farmer has access to an extension agent = 1, otherwise = 0)
- *SIZ* = Farm Size (indicates the size of the farmer's land in hectares)
- EDU = Education (number of years of formal education in years)
- *EXP* = Experience (years of experience in farming)
- *CST* = Cost of irrigation services (in Naira)

- *HHS* = Household Size (number of people living and depending on the farmer for their livelihood)
- *MRJ* = Major occupation (dummy, farming = 1, otherwise = 0)
- *FINC* = Farm income (income from farming (in Naira)
- *WUA* = Membership of Water User Association (if member, 1; 0 = otherwise)

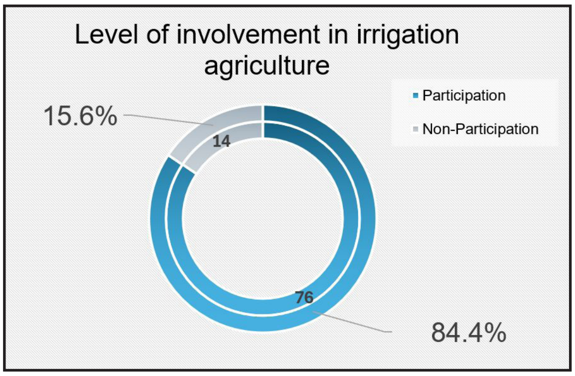
Results and discussion

Operational and structural characteristics of respondent farmers

Results are presented for selected operational and structural characteristics of the farmers such as level of involvement in irrigation agriculture, irrigation methods and usage by farmer, total and irrigable land area cultivated and irrigation type and usage by gender.

Level of involvement in irrigation agriculture

The Figure 1 below shows percentage participation and non-participation in the scheme by the farmers. Fifteen point six percent of the farmers are not involved in irrigation agriculture, while 84.4% of the farmers are fully engaged in irrigation activities. Many of the farmers in the study area were therefore involved in one form of irrigation activity or another. This statistic corroborates the study carried out by Adekunle (2015), where about 80 percent of farmers participated in irrigation farming.



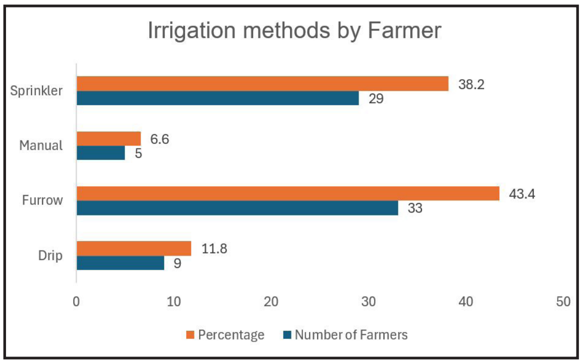
Source: Field survey data, 2023

Figure 1: Level of involvement in irrigation agriculture.

Saddiq et al. (2025) in examining impact of participation in irrigation projects on the livelihood of small-scale rice farmers in selected states of North-Western Nigeria reported 37% and 43.8% of participating and non-participating farmers being within the age range of 41-50 years. Also, 84.6% and 97.7% of the participating and non-participating farmers were of male gender, while 94.3% and 96.9% of participants and non-participants were married. These perecentages reveal that a good number of farmers where involved in irrigation agriculture at the study location. Irrigation has generally been reported to improve yield, and farmer practice of irrigation agriculture is great incentive to economic development.

Irrigation methods and usage by farmer

Irrigation types prevalent in the area and farmer usage were examined, using clustered bar as presented in the Figure 2 below. The bar reveals that 38.2 percent of the farmers in the region make use of sprinkler irrigation, while 11.8 percent make use of drip irrigation. Furthermore, only about 6.6 percent of the farmers use a manual method of irrigation, while 43.4 percent use the furrow method. This result corroborates the study by Wicaksono (2024), which reveals that the majority of the respondents in the study area practised surface irrigation, which accounted for 57 percent of the types of irrigation considered in the study. Further, 26 percent of the farmers in that study made use of sprinkler, while 13 percent used the drip irrigation method.



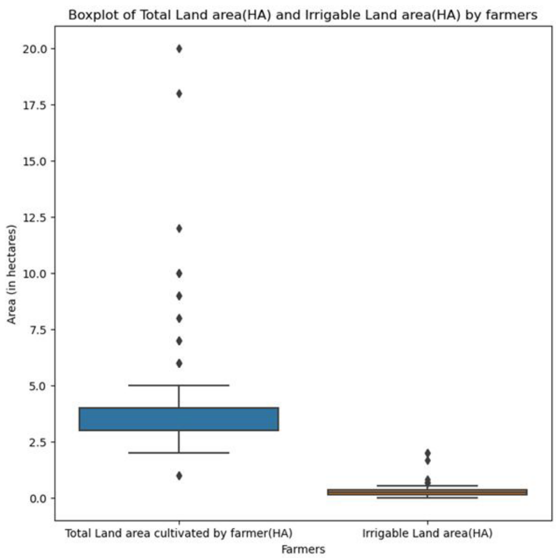
Source: Field survey data, 2023

Figure 2: Irrigation methods used by farmers.

Ballas (2024) in the book "Irrigation Methods" outlined several irrigation methods. Some of these are traditional methods, such as surface irrigation, which rely on gravity to distribute water. This method is somewhat cost-effective. Modern techniques cited in the book are drip irrigation and sprinkler systems which have been use to revolutionized water management. Other advanced methods of irrigation are subsurface irrigation and automated precision irrigation systems, which leverage technology to optimize water distribution based on soil moisture and crop needs (Balas et al., 2024).

Total land area and irrigable land area cultivated by farmers

To further examine farmer participation in irrigation activities within the catchment of Anambra-Imo River Basin, the total land area cultivated by each farmer and the corresponding irrigable land were examined, using boxplot (see Figure 3 below) Large variabilities were observed in total area of land owned and cultivated by the farmer compared to the available irrigable land as shown in the outliers. Total land area may have more variability due to land quality, topography (Smith, 2009), or historical use, while irrigable land area may be more consistent among farmers, most especially if the farmers have similar access to water sources or irrigation infrastructure. Variations in total land area could arise from differences in landownership or tenure rights. The adoption of improve agricultural technologies, most especially among small holder farmers would likely increase with farm size (Ebrahim and Toy, 2024).



Source: Field survey data, 2023

Figure 3: Boxplot showing total land area and irrigable land area by farmer.

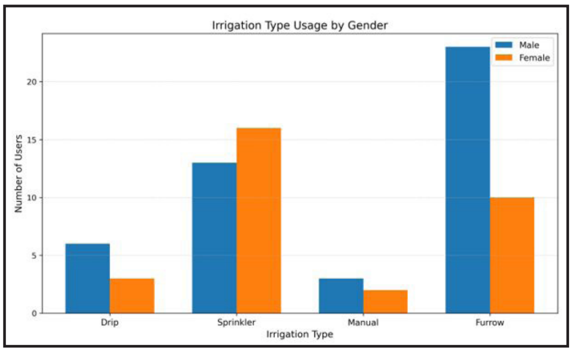
Irrigation type and usage by gender

The Table 1 and Figure 4 shows that the sprinkler and furrow irrigation methods are mostly used by both genders. More male farm owners tend to make use of the furrow method. Manual irrigation was least used by both genders. The sprinkler irrigation is the second most used method. The method simulates rainfall and it is perhaps preferred for its moderate efficiency and crop types cultivated by the farmers in the region (Shankar et al., 2015). Furrow is the most common among the male crop farmers and the second most used among the female farmers. This method, which involves making furrows along farm fields and channeling water through them, might be cost-effective to operate. However, the notable gender difference may be because it is somewhat labor-intensive thereby attracting more males than females (Radovic-Markovic et al., 2020). Gender power relations in irrigation resource access among farmers remains a critical concern (Mwalyigial et al., 2025).

	Percentage Usage by Gender			
Irrigation type	Male	Percentage	Female	Percentage
Drip	6	13.3	3	9.7
Sprinkler	13	28.9	16	51.6
Manual	3	6.7	2	6.5
Furrow	23	51.1	10	32.2
Total	45	100	31	100

Source: Field survey data, 2023

Table 1: Gender distribution of farmers by irrigation methods used.



Source: Field survey data, 2023

Figure 4: Usage of irrigation types by gender.

Determinants of involvement in irrigation agriculture

The estimated determinants of participation in irrigation agriculture are summarized and presented in Table 2 below: The table reveals that farming experience, Household size, major occupation of the farmer, farm income, being a member of the Water Users Association (WUA) were significant variables that influenced particicpation or involvement in irrigation agriculture. The regression result reported a Wald chi² of 39.65 and a log pseudolikelihood

Variable	Co-efficient	Standard Error	Z	P>|z|
Experience in Farming	.0080661	.0035	2.30	0.021[2]
Household Size	-.0399064	.01785	-2.24	0.025[2]
Major Occupation	-.2292222	.05423	-4.23	0.000[1]
Farm Income	5.64e-08	.00000	2.01	0.044[2]
Member of WUA	-.2150519	.06218	-3.46	0.001[1]
Cost of irrigation	1.74e-07	00000	-1.65	0.100
Age	-.0015055	.00315	-0.48	0.633
Extension Services	.0111409	.01478	0.75	0.451
Farm Size	0104183	.00795	-1.31	0.190
Education	.0077013	.0047	1.64	0.102
No. of observations = 90				
Wald chi^2 (10) = 39.65				
Log pseudolikelihood = -281.37084				

Note: [3, 2, 1] statistically significant at 10%, 5% and 1% respectively

Source: Field survey data, 2023

Table 2: Estimated determinants of involvement in irrigation agriculture.

of -281.37084, providing insights into the model's fitness and overall significance.

Farming experience is significant at a 5% level and positively signed. This implies that there is a direct positive relationship between experience in farming and participation in irrigation schemes by the farmer. This is in line with a prior expectation, as the more experienced the farmer is, the more likely he will be exposed to farming technologies and the more receptive to new innovations. A unit change in years of farm experience will lead to a 0.8% change in the probability of participation in irrigation farming. The study corroborates the work of Ainembabazi and Mugisha (2014), which posited that farming experience is useful in the early stages of adoption of a given technology when farmers are still testing the potential benefits of the technology. The study further emphasized that gradual advances in technology development and continuous retraining of the farmers are very essential for sustainable and lasting adoption of agricultural technologies for certain crops. In an earlier study carried out by Dauda et al. (2009), to assess the roles of irrigation farming in the millennium development goals, the result of the study showed that irrigation experience was among the variables found to have significant effects on profit realized from irrigation farming. This would not have been possible without the farmer being involved in the farming activities, leading to years of experience.

Household size is significant at the 5% level and is negatively signed. This implies that there is an inverse relationship between household size and involvement in irrigation agriculture. In larger households, resources such as land, labor, and capital are most of the time spread thinly among family members. As a result, households may prioritize subsistence activities that require less capital investment (Xie, 2017; Kolawole et al., 2020), such as rainfed agriculture, over engagement in irrigation farming that entails additional costs. Furthermore, a larger household may have more members available for agricultural labour, thereby reducing the need for external irrigation infrastructure to enhance productivity. As a result, these households may rely on more traditional farming methods that do not require irrigation. A unit change in the number of farmers household size will lead to a 3.9% change in the probability of involvement in irrigation agriculture. This result differs from the findings of Afodu et al. (2021), who revealed that a larger household size may lead to more adoption of the modern technologies in plantain production, which could translate into increased productivity of the plantain farmers. However, it corroborates the findings of Olumba (2014) and Mengiste et al. (2025), who found that household size influences farming decisions on management practices.

Major occupation is significant at the 1% level and is negatively signed, reflecting an inverse relationship between the major occupation of the farmer and participation in irrigation agriculture. A unit change in the main occupation will lead to a 22.92% change in the probability of involvement in irrigation farming. The distances from water sources and markets for products, as well as the financial costs associated with irrigation activities, can discourage a farmer from being involved in irrigation activities, especially if their main occupation requires significant time and resources elsewhere. Yin et al. (2016) viewed this differently in his attempt to examine the impacts of off-farm employment on irrigation water efficiency (IWE), an attribute of participation, with a set of household level data collected in Hebei Province in North China. Findings from thee study reveals that households with higher number of labourers working off-farm seem to achieve a higher IWEs.This implies that the higher the number of households involved in non-farm income activities, which is most likely a major occupation due to their number, the more efficient they are in irrigation water use. This sub-study contradicts the result of our findings which is an inverse relationship. Further, off-farm employees tend to achieve higher IWEs, while those with more elderly people as labourers and those with larger farms are associated with lower IWEs (Yin et al., 2016). The study further corroborates that of Rustinsyah (2019) which posited that water management in Bengawan Solo river basin, Indonesia is basically managed by businessmen rather than the WUA due to limited capital. This reflects a clear positive relationship between a major occupation and participation in irrigation scheme.

Farm income is significant at the 5% level with a positive sign. This implies that a unit change in farm income will lead to a 0.00000564% change in the probability of involvement in irrigation agriculture. Farm income has a positive relationship with involvement in irrigation agricultural activities because engaging in irrigation activities can significantly impact household income. Research indicates that participation in irrigation can lead to increased household income and crop production. This corroborates the work of Gadisa and Gebrerufael (2021). This positive relationship underscores the importance of irrigation in improving welfare and agricultural productivity (Yusuf et al., 2023). Further evidence by Li (2020) in a study on impact of access to irrigation on rural incomes and diversification in China, posited that, access to irrigation has a significant and positive relationship on rural incomes and diversification. Attributes and treatment effect of irrigation access were to increase farm income, household income and income diversification by 14, 10 and 107% respectively. Findings from the study revealed that irrigation effects on diversification and rural incomes are diverse and multifarious between small and large scale farmers and between gender. Irrigation generally play a crucial role in augmenting household farm income and improve livelihoods of farmers

WUA membership is significant at the 1% level and negatively signed. This result shows an inverse relationship between membership of WUA and participation in irrigation schemes. A unit change in membership will lead to a 21.5% change in the probability of involvement in the irrigation agriculture. The negative coefficient for membership of a WUA is a pointer that being a member of a WUA has an indirect relationship with participation in irrigation activities. This is contrary to a priori expectations and could be due to various factors such as conflicting interests, association dues, lack of perceived benefits, or challenges in decision-making processes, within the association or the management of the river basin (Nyamulo and Pastory, 2024). This finding goes contrary to the work of Gadisa and Gebrerufael (2021)), whose results showed that cooperative membership has a positive and statistically significant effect on technology adoption. Suraj (2025) in his work on Membership of Water User Associations and food security reveal a significant relationship between being a member of WUA and participation in supplementary irrigation and draught index insurance initiatives. Other factors that had significant effect of WUA were age, marital status, access to extension services, farm size and the asset aspiration gap (Suraj, 2025). The study advocates development policies that strengthens existing WUAs through inclusive approaches that addresses food seccurity challenges.

Cost of irrigation shows a borderline involvement of farmers in irrigation agriculture. Even though this effect is not statistically significant, the marginal outcome ($P = 0.10$) suggests that irrigation costs may still play a role, though weak, in affecting farmers involvement in irrigation agriculture. The sign of the coefficient is positive indicating that as irrigation cost increases, involvement in irrigation agriculture equally increases, although the evidence is not strong enough to draw definitive conclusion. The maginal outcome may further be a reflection of the reality that farmers sometimes rely on informal water sharing arrangements (Manzoor Qadir, 2007; Tang et al., 2025) that reduce sensitivity to irrigation costs. The increased involvement may equally mean that farmers who are commercially benefiting from the gains of irrigation, don't mind commiting more resources because of expected higher returns. Such category of farmers are likely to invest in more pumping hours or regular facility mainenance, leading to slightly higher involvement. This findings is in line with studies in Sub Saharan Africa, showing that irrigation cost affects farmers' participation or water-use decisions (Faye and Von Braun, 2024; Faye and Von Braun, 2025). However, modelling cost of irrigation in the study area as a continuous predictor, provides novel empirical evidence, highlighting that even marginal cost effects can shape farmers' involvment in river basin irrigation schemes.

Non-significant factors: age, extension services, farm size and education did not significantly influence farmer involvement in irrigation agriculture in the catchment of the Anambra-Imo River Basin Development authority. Although these factors should naturally influence or play major roles in shaping farmer involvement in irrigation activities, their non-significant status is a pointer to the fact that other factors may be more central to farmer engagement in the irrigation project than the ones examined in this study, suggesting that engagement in the scheme is shaped less by personal or background characteristics and more by context-specific economic, institutional, and household factors.

Conclusion

A relatively high participation rate was recorded among the sampled farmers in the Anambra Imo River Basin Development Authority irrigation scheme. Major public investments in agricultural water infrastructure such as these are established to enhance productivity of farmers (Oiganji et al., 2025). Therefore, farmer participation is important, and the extent thus far, is quite commendable, although majority are involved in surface irrigation practice. This is not unexpected due to its characteristic low cost of installation and operation (Muchara, 2025), and the fact that there is abundant water availability through the rivers, streams and underground water. Factors such as years of experience in farming, household size, major occupation, farm income, membership of water user association, play a major role in influencing farmer involvement in irrigation agriculture. However, it is startling to note that though being a member of the Water Users Association should be an incentive to involvement in irrigation agriculture, it is rather a disincentive development. This may be due to management and logistic protocols. The findings, however, necessitates a further investigation into the situation. It is recommended that a participatory irrigation scheme, which should be democratically administered by the farmers and the River Basin Management, be adopted. This would enable the farmers to take charge of the daily sourcing and allocation of water, while the River Basin Authority, as a supervisory body, would create the enabling environment and resources, with a member of the water users association serving as an ad-hoc part of its team.

Corresponding author:
Dr. Ubon Essien
Department of Agricultural Economics, Federal University of Technology
PMB 1526, Owerri, Imo State, Nigeria
Phone: +2348034330650, Email: ubon.essien@futo.edu.ng

References

[1] Aarnoudse, E., Closas, A. and Lefore, N. (2019) "Water User Association: A review of approaches and alternative management for Sub-Saharan Africa", Colombo, Sri Lanka: International Water Management Institute (IWMI). Working Paper 180, 43 p. E-ISSN 2478-1134. DOI 10.5337/2018.210.

[2] Adugna, A. G. and Gessesse, M. Y. (2019) "Analysis of rural women's participation in rice production using ordinal logistic regression model", *Journal of Rural Development*, Vol. 38, No. 2, pp. 234-265. ISSN 2582-4295. DOI 10.25175/jrd/2019/v38/i2/146745.

[3] Abah, D. and Nankiel, P. (2019) "River basin development authorities and Nigerias economic development since 1960", *Journal of the Faculty of Arts and Islamic Studies*, Vol. 7, No. 1, pp. 52-66. ISSN 0794-9316.

[4] Adekunle, O. A., Oladipo, F. O. and Busari, I. (2015) "Factors affecting farmers participation in irrigation schemes of the Lower Niger River Basin and Rural Development Authority, Kwara State, Nigeria", *South African Journal of Agricultural Extension*, Vol. 43, No. 2, pp. 42-51. E-ISSN 2413-3221, ISSN 0301-603X. DOI 10.17159/2413-3221/2015/v43n2a353.

[5] Afodu, O. J., Akinboye, O. E., Akintunde, A. O., Ndubuisi-Ogbonna, L. C., Shobo, B. A. and Oyewumi, O. S. (2021) "Assessing the impact of technology adoption on the productivity of plantain farmers in Nigeria", *Agricultural Socio-Economics Journal*, Vol. 21, No. 4. ISSN 2745-6897. DOI 10.21776/ub.agrise.2020.021.4.8.

[6] Ainembabazi, J. H. and Mugisha, J. (2014) "The Role of Farming Experience on the Adoption of Agricultural Technologies: Evidence from Smallholder Farmersin Uganda", *The Journal of Development Studies*, Vol. 50, No. 5. E-ISSN 1743-9140, ISSN 0022-0388. DOI 10.1080/00220388.2013.874556.

[7] Balas, D. B., Thakor, D. and Dhakad, S. (2024) "Irrigation Methods", In Tiwari, N. J., Banjare, C., Shulee Ariina, M. M., Dange, M. M., Sahu, A. (eds.) *"Fundamentals of Vegetable Science",* Emerald Publishing House. ISBN 9789395345286.

[8] Christopher, O. A., Olawale, O. A., Fidelis, B. and Toju B. (2016) "Assessing the impacts of irrigation systems on food security in Southwestern Nigeria", *Conference proceedings of the 37th national conference & annual general meeting of the Nigerian Institution of Agricultural Engineers*, Chemical Engineering Hall, Federal University of Technology, Minna, Niger State, Nigeria. 2016; Vol. 37.

[9] Climate-Data.org (n.d.) "*Temperature and climate graph for Anambra"*. [Online]. Available: https://en.climate-data.org/africa/nigeria/anambra-344/ [Accessed: May 25, 2025].

[10] Dauda, T. O. , Asiribo O. E., Akinbode, S. O., Saka, J. O. and Salahu, B. F. (2009) "An assessment of the roles of irrigation farming in the millennium development goals", *African Journal of Agricultural Research,* Vol. 4, No. 5, pp. 445-450. ISSN 1991-637X.

[11] Ebrahim E. A. and Toy, A. (2024) "Impact of Irrigation Participation on Households Food Security in Rural Areas in Ethiopia: Application of Propensity Score Matching (PSM) Method for Casual Inference", *Frontiers in Sustainable Food Systems,* Vol. 8, p. 1403317. ISSN 2571-581X. DOI 10.3389/fsufs.2024.1403317.

[12] Faye, A. and Braun, J. (2025) "Uptake and profitability of small-scale irrigation in the Sahel: insights from the literature and survey data", *Agricultural Water Management*, Vol. 313, p. 109469. ISSN 1873-2283. DOI 10.1016/j.agwat.2025.109469.

[13] Faye, A. and Braun, J. (2024) "Small-Scale Irrigation in the Sahel: Adoption Trends, Profitability, and Challenges, ZEF Discussion Papers on Development Policy No. 353. Zef Centre for Development Research, University of Bonn. DOI 10.2139/ssrn.5048785.

[14] Gadisa M. and Gebrerufae, I. G. (2021) "Impact of small-scale irrigation on household income in Central Ethiopia: Empirical evidence from Walmara District", *International Journal of Agriculture and Biosciences*, Vol. 10, No. 2, pp. 101-106. E-ISSN 2306-3599, ISSN 2305-6622.

[15] Kolawole, A. O., Oluwatusin, F., Ajiboye, A., Kolawole, A. O., Oluwatusin, F., Oludeye, A. A., Atarumu, A., Abdul-Raheem, K. and Akokoh, F. E. (2020) "Status Analysis of Irrigation Farming Households in Nigeria", *World Rural Observations*, Vol. 12, No. 2. E-ISSN 1944-6551, ISSN 1944-6543.

[16] Li, J., Ma, W., Renwick, A. and Zheng, H. (2020) "Impact of Access to Irrigation on Rural Incomes and Diversification: Evidence from China", *China Agricultural Economic Review*. Vol. 12, No. 4, pp. 705-725. E-ISSN 1756-1388, ISSN 1756-137X. DOI 10.1108/CAER-09-2019-0172.

[17] Mengiste, W., Tsegaye, D., Dagalo, S., and Yitbareke, T. (2025) "Determinants of farmers' use of small-scale irrigation practice in the Achewa, area, Gambella, Southwestern Ethiopia", *Water and Soil Management and Modeling*, Vol. 5, No. 4, pp. 159-173. E-ISSN 2783-2546. DOI 10.22098/mmws.2025.18243.1665.

[18] Mkuna, E. and Wale, E. (2023) "Smallholder farmers' choice of irrigation systems: Empirical evidence from Kwazulu-Natal, South Africa and its implications", *Scientific African*, Vol. 20, p. e01688. E-ISSN 2468-2276. DOI 10.1016/j.sciaf.2023.e01688.

[19] Muchara, B. (2025) "When Simplicity Sustains': Assessing the Viability of Furrow-Based Irrigation Over Sprinkler Systems in South Africa's Smallholder Schemes", *IBC 18th International Business Conference* 1 University of South Africa, Graduate School of Business Leadership, p. 1685, South Africa, pp. 1-15.

[20] Mwalyagile, N, Nshimba, J. and Salanga, R. J. (2025) "Understanding gender power relations in irrigation resource access and decision-making in small-scale irrigation schemes in Mbarali District, Tanzania", *IIMT Journal of Management,* Vol. 2, No. 3, E-ISSN 2976-727X, ISSN 2976-7261. DOI 10.1108/IIMTJM-04-2024-0052.

[21] National Beareau of Statistics (NBS). (2022). [Online]. Available: https://www.nigerianstat.gov.ng/pdfuploads/demographic_bulletin_2022_final.pdf/ [Accessed: April 14, 2025].

[22] Nyamulo, P. A. and Pastory, D.(2024) "Factors Affecting Water User Associations' Performance in Managing Water Resources in Ruvu Catchment, Pangani Basin, Tanzania", *Journal of Policy and Development Studies(JPDS)*, Vol. 17, No. 1, E-ISSN 2814-1091, ISSN 1597-9385. DOI 10.4314/jpds.v17i1.6.

[23] Oiganji, E., Igbadun, H., Amaza, P. S. and Lenka, R. Z. (2025) "Innovative Technologies for Improved Water Productivity and Climate Change Mitigation, Adaptation, and Resilience: A Review", *Journal of Applied Sciences and Environmental Management,* Vol. 29, No. 1. pp. 123-136. E-ISSN 2659-1499, ISSN 1119-8362. DOI 10.4314/jasem.v29i1.17.

[24] Olumba, C. C. (2014) "Productivity of improved plantain technologies in Anambra State, Nigeria", *African Journal of Agricultural Research*, Vol. 9, No. 29, pp. 2196-2204. ISSN 1991-637X. DOI 10.5897/AJAR2014.8891.

[25] Omoja, U. C, Okpalaku, B. N., Uchechukwu, U. N. and Obiekezie, T. N. (2021) "Variability of Rainfall in Awka, Anambra State Nigeria", *IOSR Journal of Applied Physics*, Vol. 13, No. 4, Ser. I, pp. 50-56. E-ISSN 2278-4861. DOI 10.9790/4861-1304015056.

[26] Qadir, M., Wicheins, D., Minhas, P., McCornick, P., Abaidoo, R., Attia, F., El-Guindy, S., Ensink, J., Jimenex, B., Kijne, J., Koo-Oshima, S., Oster, J., Oyebande, L., Sagardoy, J. A. and van der Hoek, W. (2007) "Agricultural use of marginal quality water-opportunities and challenges", In Molden, D. (ed.) "*Water for food, water for life: a Comprehensive Assessment of Water Management in Agriculture"*, Chapter 11. London, UK: Earthscan; Colombo, Sri Lanka: International Water Management Institute (IWMI). DOI 10.22004/ag.econ.158131.

[27] Raji, E, Ijomah, T. I. and Eyieyien, O. G. (2024) "Improving agricultural practices and productivity through extension services and innovative training programs", *International Journal of Applied Research in Social Sciences*,Vol. 6, No. 7, pp. 1297-1309. E-ISSN 2706-9184, ISSN 2706-9176. DOI 10.51594/ijarss.v6i7.1267.

[28] Rustinsyah (2018) "Management of Agricultural Irrigation and Non-Farm Economic Activities in Rural Areas", In *Proceedings of the 4th International Conference on Contemporary Social and Political Affairs (ICoCSPA 2018)*, pp. 86-90. SCITEPRESS – Science and Technology Publications, Lda. ISBN 978-989-758-393-3. DOI 10.5220/0008817100860090.

[29] Saddiq, N. M., Abdullahi, T. U. and Burabe, B. I. (2025) "Impact of participation in irrigation projects on the livelihood of small-scale rice farmers in selected states of North-Western. Nigeria", *Journal of Arid Agriculture*, Vol. 26, No. 2, pp. 48-56. ISSN 0189-7551. DOI 10.63659/jaa.v26i2.86.

[30] Shankar, M. S., Ramanjaneyulu, A. V., Neelima, T. L. (2015) "Sprinkler irrigation: an asset in water scarcity and undulating areas", In: Das, D. J., Anup, Ngachan, Sikka, S. V. A.K., and Lyngdoh, M. (eds.) "*Integrated Soil and Water Resource Management for Livelihood and Environmental Security. Rajkhowa*", pp. 259-283. ICAR Research Complex for NEH Region, Umiam—793 103, Meghalaya, India. [Online]. Available: https://www.researchgate.net/publication/325870811_Sprinkler_irrigation_- [Accessed: April 18, 2025].

[31] Smith, R. J., Raine, S., Alison, C. and McCarthy, N. H. (2009) "Managing spatial and temporal variability in irrigated agriculture through adaptive control", *Australian Journal of Multi-Disciplinary Engineering*. Vol. 7, No. 1. ISSN 2204-2180. DOI 10.1080/14488388.2009.11464801.

[32] Suraj, M. M, Martey, E., Kuwornu, J. K, Apiors, E. K., Hemeze, F. H. and Etwire, P. M. (2025) "Membership of Water User Association and Implications for Food Security", *Journal of Agriculture and Food Research*, Vol. 20, p. 101739. E-ISSN 2666-1543. DOI 10.1016/j.jafr.2025.101739.

[33] Tang, H., Yang, Z., Guo, Z., Yang, C., Huang, F. and Ran, R. (2022) "The Willingness to Pay for Agricultural Irrigation Water and the Influencing Factors in the Dujiangyan Irrigation Area: An Empirical Double-Hurdle Model Analysis", *Frontiers, Sec. Environmental Economics and Management*, Vol. 10. E-ISSN 2813-2823. DOI 10.3389/fenvs.2022.906400.

[34] Wicaksono, K. P., Permanasari, P. N., Della-Aprillia, D. P., Dewi, R. W. K., Ramadhani, R. and Rohman, A. (2024) "The Willingness of Farmers to Adopt Innovations from Farmer -Owned Enterprises", *HABITAT*, Vol. 35, No. 1, pp. 31-39. E-ISSN 2338-2007. DOI 10.21776/ub.habitat.2024.035.1.4.

[35] Xie, H., You, L. and Takeshima, H. (2017) "Invest in Small-Scale Irrigated Agriculture: A National Assessment of Potential to Expand Small-Scale Irrigation in Nigeria", *Agricultural Water Management*, Vol. 193, pp. 251-264. E-ISSN 1873-2283. DOI 10.1016/j.agwat.2017.08.020.

[36] Yin, N., Huang, O. and Yang, Z. (2016) "Impacts of Off-Farm Employment on Irrigation Water Efficiency in North China", *Water*, Vol. 8, No. 10, p. 452. ISSN 2073-4441. DOI 10.3390/w8100452.

[37] Yusuf, A. B., Ibrahim, M. L. and Caleb, A. (2023) "Impact of Irrigation Farming Scheme on Farmers' Income: Evidence from Damasak, Borno State, Nigeria", *International Journal of Accounting, Finance, and Administrative Research*, Vol. 1, No. 1, pp. 1-14. ISSN 3027-2483.

Agris on-line Papers in Economics and Informatics

Volume XVII Number 4, 2025

Relationship Between Rural Poverty and Agricultural Diversification at a Local Scale in Colombia: An Approach through Spatial Effects

Alejandro Mojica Godoy

Universidad Externado de Colombia, Bogotá, Colombia

Abstract

This paper examines the relationship between local agricultural diversification and rural poverty in Colombia. The evidence presented suggests that municipal-scale agricultural diversification is associated with higher levels of rural poverty. The primary mechanism driving this relationship is the loss of external economies of scale in agricultural production. However, when analyzing the spatial effects (autocorrelation and spatial heterogeneity), it was found that this relationship varies across the country. Neighboring municipalities have opposite spillover effects that offset the positive effect. In addition, the presence of small rural producers, greater provision of public goods, the existence of industrial crops, and lower persistence of acts of violence were all found to be associated with lower levels of rural poverty. This study joins the literature on the economic and social effects of agricultural diversification, particularly in the context of its promotion as a means of adapting to and mitigating the effects of climate change

Keywords

Agricultural diversification, poverty, spatial econometrics, spillover effects, regional economics.

Mojica Godoy, A. (2025) "Relationship Between Rural Poverty and Agricultural Diversification at a Local Scale in Colombia: An Approach through Spatial Effects", *AGRIS on-line Papers in Economics and Informatics*, Vol. 17, No. , pp. 67-79. ISSN 1804-1930. DOI 10.7160/aol.2025.170406.

Introduction

In Latin America, the public policy debate of whether to promote diversified agriculture with many products or specialize in one is influenced by the recurring competition for resources between large-scale and family farming. For instance, in countries like Brazil, Uruguay, and Argentina, large-scale agriculture encourages specialization through sustained growth in the cultivation of extensive, high-volume, and low-value crops such as soybeans or sugarcane. In contrast, family farming units typically engage in multi-product cultivation, including grains, cocoa, fruits, and vegetables (Giller et al., 2021).

While some economic literature suggests that diversification may deteriorate social indicators due to the loss of economies of scale, reduced labor productivity, and underutilization of comparative advantages (Lobao and Sharp, 2013), other argue that it can enhance social performance through economies of scope, reduced income volatility for producers, and increased land productivity through better soil conservation (Klasen et al., 2016; Wardhana et al., 2017).

The debate between large-scale agriculture, generally specialized, and family agriculture, traditionally diversified, is reflected in Colombia by the competition for the use of resources, especially land, and for the food safety debates. According to data from the 2014 National Agricultural Census, 70.4% of the Agricultural Production Units (UPAs, for its acronym in Spanish) have less than 5 hectares and occupy 2.0% of the rural area; while 0.2% of the UPAs have 1,000 hectares or more and occupy 73.8% of the rural area censed. Small agricultural units are usually dedicated to the harvest of high-value products, such as legumes and fruits, or for commercialization, such as coffee and cocoa (Berry, 2023). In turn, it is estimated that the proportion of food production that comes from smallholder agriculture ranges between 50 and 68% in Colombia (Perfetti and Leibovich, 2013).

For this study, agricultural diversification is described as 'a process that transforms a simplified cropping system by introducing additional crops over time and space' (Hufnagel et al., 2020). Therefore, diversification is commonly linked to various agricultural practices, including crop rotation, adoption of cover crops, relay cropping, alley cropping, and mixed cropping.

The objective of this paper is to analyze the relationship between agricultural diversification and poverty at the local scale in Colombia, which takes the form of local administrative entities. The scale is chosen because in literature the analysis is usually developed from the perspective of agricultural units (Klasen et al., 2016; Thapa et al., 2018; Sekyi et al., 2021) but this approach ignores that the economic benefits and environmental detriments of specialization transcend the agricultural unit and have local spillover effects and economic externalities.

The paper presents that municipal-scale agricultural diversification is associated with higher levels of rural poverty. The primary mechanism driving this relationship is the loss of external economies of scale in agricultural production. However, when analyzing the spatial effects (autocorrelation and spatial heterogeneity), it was found that this relationship is not uniform across the entire national territory and has opposite spillover effects when analyzing the neighboring municipalities that offset the positive effect.

This paper presents in the Material and Methods section the methodology used to calculate the diversification index, and the econometric models used (autocorrelation and spatial heterogeneity). The Results and Discussion section presents the results of the diversification indicator, and the models calculated, as well as the interpretations of the main findings. The article closes with conclusions and their research and public policy impact.

Materials and methods

This section presents how the agricultural diversification indicator is calculated. Moreover, the model to measure the relationship between rural poverty and agricultural diversification is described.

The diversification index

The species evenness index proposed by Pielou (1966) is applied to measure agricultural diversification. This indicator is used in ecology to measure the uniformity of species within an ecosystem, i.e., to analyze how evenly they are distributed. In a literature review, Waha et al. (2022) find that this is one of the most widely used standardized indices in the agricultural economics literature for the study of agricultural diversification within rural productive units.

In the present work, agricultural diversification measured from Pielou's index calculates how uniform agricultural species are cultivated within a territory. If a local administrative entity has a low score on the index, it means greater agricultural uniformity or, in other words, that the same species tend to be cultivated within that territory. While a high score represents less uniformity and greater species diversity. The calculation uses data from the "Evaluaciones Agrícolas Municipales" (EVA), which are annual statistical records at the municipal level in Colombia with agricultural topics such as sown area, harvested area, production and crop yields, among others.

In this order, the Pielou index, H_i, is calculated for Colombia's 1,105 local administrative entities. For each entity i, p_j represents the proportion of the area planted with a crop j within each local territory. In other words, the numerator, a_j, is the area planted with a particular crop for the local territory and the denominator, A_i, is the total area planted in the respective entity (the sum of all the a_j):

$$p_j = \frac{a_j}{\sum_{j=1}^{s} a_j} \tag{1}$$

Subsequently, the proportion of each crop in each entity is multiplied with its natural logarithm. The products are summed up for the S crops that exist in the territorial area. The result of the addition is multiplied by minus 1 and divided by the natural logarithm of the total S number of crops existing in the territorial area to normalize it in a range between 0 and 1.

$$H_i = \frac{-\sum_{j=1}^{S} p_j \ln(p_j)}{\ln(S)} \tag{2}$$

The EVA databases from 2007 to 2021 were used to calculate the diversification indicator for the local administrative entities of the country in each year. Subsequently, an average per locality of agricultural diversification is calculated for the 15 years of the study. The indicator has shown a stable behavior during the period analyzed, which means the absence of a trend or patterns of behavior (more details in the Appendix).

The model

To quantify the impact of agricultural diversification on rural poverty in the country's municipalities, the following regression model is constructed:

$$MPI = \beta_0 + Pielou\ Index + \sum_{2}^{k} \beta_k X_k + \varepsilon \quad (3)$$

In this model the dependent variable is the rural component of the Multidimensional Poverty Index (MPI) at the municipal (local) level, the main independent variable of interest is the agricultural diversification indicator constructed above. The controls are: 1. the management component of the municipal performance measurement indicator constructed by the Colombian National Planning Department (DNP, for its acronym in Spanish) which is composed of an assessment of the territorial entity in mobilization and execution of resources, own resources, citizen attention and accountability, and management of land use planning instruments (this is our approximation to local authorities policy performance), 2. the percentage of Agricultural Production Units within the municipality that have an extension between 0 and 1 hectares, 3. the percentage of the population living in rural areas, 4. the percentage of UPA within the municipality with access to irrigation systems, 5. number of occurrence of violence threat events in 2021 according the 'Registro Único de Victimas', as a violence indicator proxy, 6. a dummy variable for the presence of industrial and/or exportable crops (coffee, sugar cane, sugar cane, cocoa, flowers, oil palm, cotton, timber and bananas) in the municipality, 7. the percentage of land use that are in conflict due to underutilization, and 8. the altitude of the municipality measured in meters above sea level (more details in the Appendix).

Conceptualization of geographic space

The decision to include spatial effects in the modeling is justified by the findings of Galvis and Meisel (2010) that poverty at the municipal level in Colombia has a high and significant spatial correlation. In other words, there is a relationship between the poverty indicator of a municipality and that of its neighbors, even at a timeless level. Therefore, there are "neighborhoods" in which poverty is concentrated in the country and persist over time.

In an econometric model, the presence of spatial relationships between the variables of interest in the model affects the statistical tests or generates biased estimators (Lesage et. al., 2009). In this order, it is necessary to resort to spatial econometrics, a subfield of econometrics that aims to study spatial interaction effects between different geographical units (Elhorst, 2014). In this paper, the geographical units analyzed are the country's local administrative entities, also known as municipalities, while the economic phenomenon of interest is the relationship between rural poverty and agricultural diversification within and between municipalities. The study relationship is analyzed from two spatial effects: i. spatial autocorrelation and ii. spatial heterogeneity.

Spatial dependence

Spatial dependence or autocorrelation refers to the phenomenon whereby the values of a region i depend on the values of its neighboring observations. That is, the presence of a simultaneous data generation process where the values of y_i depend on the values of y_j and vice versa (Anselin, 1988; LeSage and Pace, 2009).

This phenomenon refers to a cross-sectional relationship where the correlation or covariance structure between random variables depends on their relative position in geographic space (Arbia, 2006; Anselin, 2010). In this paper, a spatial Durbin model to represent the spatial dependence model will be run, which has the following representation:

$$\mathrm{Y} = \rho \mathrm{WY} + \alpha t_N + \mathrm{X}\beta + \theta \mathrm{WX} + \varepsilon \quad (4)$$

The terms that are common with traditional linear regression model are: Y, which is an n x 1 vector representing the dependent variable for each sample unit; t_N, which is an n x 1 vector of ones, used to represent the value of the constant or intercept α of the model; X denotes an n x k matrix of exogenous explanatory variables and β is associated with the k x 1 vector of parameters to be estimated.

In contrast to traditional ordinary least squares models, two types of spatial interactions can be evidenced: i. endogenous interaction effects on the dependent variable (Y) and ii. exogenous interaction effects on the independent variables (X) (Anselin, 1988; Elhorst, 2014). Each type of interaction has a coefficient to be estimated that captures the influence of neighboring observations: ρ, on the dependent variable Y and θ, on the independent variables X. In this paper, the independent variable of interest is the diversification index.

In econometric modeling, the spatial lags are represented by the multiplication between the spatial weighting matrices (W) with the spatial coefficients. Specifically, WY represents the endogenous spatial interaction between the dependent variables, and WX between the independent variables of the different units of analysis (Elhorst, 2014).

As used in the literature studying the relationship between agriculture and poverty, or in rural poverty studies, this paper modelling used as a binary contiguity weighting matrix. (Palmer-Jones and Sen, 2006; Wardhana, et al., 2017; da Silva, et al., 2022; Rahmawati, et al., 2023). In this type of matrix, each element indicates whether two units share a common boundary, with a value of 1 representing adjacency and 0 representing non-adjacency.

Spatial heterogeneity

The hypothesis of spatial uniformity of the effects of explanatory variables is unrealistic (Brunsdon et al. 1996) because global estimators hide spatial variations of the parameters (de Bellefon and Floch, 2018). The spatial heterogeneity is the phenomenon of structural instability that manifests itself in space. In other words, the coefficients of explanatory variables of a model can be the same at different points in the geography, but not have the same grade of effects in each of them. Therefore, spatial heterogeneity characterizes the phenomenon in which the parameters of the model are variable at different points (Anselin, 1988; Lesage et. al., 2009).

The Geographically Weighted Regression model (GWR) allows to analyze the variation in space in a continuous way and is represented as follows (Anselin, 2010):

$$y_i = \beta_o(u_i, v_i) + \sum \beta_k(u_i, v_i)x_{ik} + \varepsilon_i \qquad (5)$$

Where y_i is the dependent variable of observation i, $\beta_k\ (u_i,\ v_i)$ are the coefficients of the regression that vary according to the coordinates $(u_i,\ v_i)$ and ε_i is the error term. In this framework, the coefficients are estimated separately for each location $(u_i,\ v_i)$. This modeling implicitly assumes that observations near location i have greater influence on the estimate of $\beta_k\ (u_i,\ v_i)$ than observations that are farther away from i; thus, nearby values will have relatively similar magnitudes and signs. Therefore, to estimate the parameter at location i, a traditional linear regression with the subset of data near i is employed. For the next observation, a different set of nearby points is used to estimate the parameter (Fotheringham et. al., 2003). The main output of this modeling is a map showing the spatial variability of the behavior of an economic relationship.

There are 3 elements of the kernel function that define the spatial relationships within a GWR model. First, the kernel shape refers to the function that defines the weight given to each observation within the neighborhood surrounding the given observation. Functions can be uniform, Gaussian, or exponential; however, the choice of function only slightly changes the results (Brunsdon et al. 1996). Second, there is a choice between a fixed or adaptive kernel, where the former refers to the extent of the kernel being defined by the distance to the point of interest, whereas, in the latter the extent of the kernel is determined by the number of neighbors of the point of interest. That is, in the adaptive kernel, the lower the density of observations, the smaller the kernel. Finally, the last element is the definition and choice of the bandwidth, which represents the distance from which the relationship between observations will have a value of zero. Thus, the value of the bandwidth h is the parameter of choice that has the most influence on the results, since the greater the bandwidth, the greater the number of observations that will have a weight different from 0 (de Bellefon & Floch, 2018, p. 235).

For the application of the current model, a continuous or Gaussian kernel function was chosen because the dependent variable of the model (rural component MPI) does not take binary or discrete values. Additionally, an adaptive kernel was chosen because this extension is recommended for spatial data that do not have a homogeneous distribution (de Bellefon & Floch, 2018), such as the country's municipalities. Finally, for the calculation of the bandwidth, the Akaike criterion was chosen, i.e., to determine the bandwidth, a set of local regression models is calculated and the Akaike criterion defines the number of optimal neighborhoods from the set of regressions.

Results and discussion

This section presents the results of the calculation of the agricultural diversification indicator and the two models estimated: spatial dependence and spatial heterogeneity.

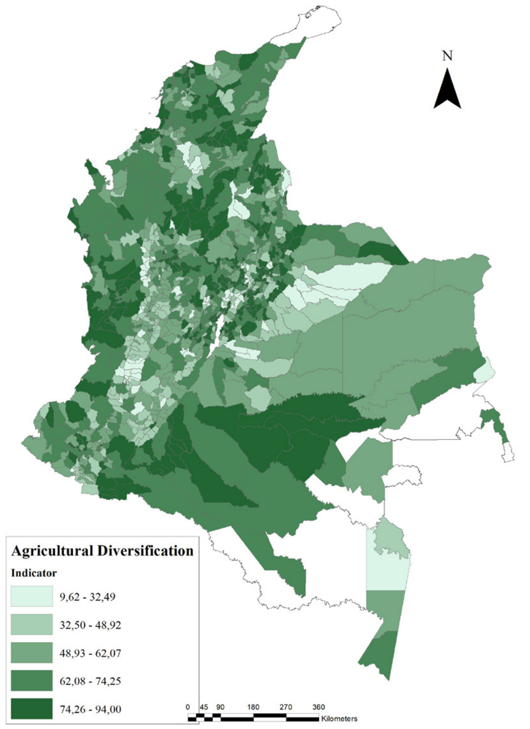
Diversification index

Table 1 presents the local administrative entities of the country with the lowest and highest agricultural diversification for the average of the 15 years of study.

In terms of production, the three products that most promote local agricultural specialization in Colombia are sugarcane, potatoes, and coffee. Sugarcane occupies 93% of the planted areas of Palmira, Valle, and 94.3% of Candelaria, Valle. Potato occupies 90.5% of Tutazá, Boyacá and 91.23% of Sesquilé, Boyacá. Coffee occupies 93.8% of Piendamó, Cauca, and 91.5% of Betania, Antioquia.

Figure 1 shows municipal diversification in Colombia. There is a high agricultural specialization in regions such as 'Altillanura de la Orinoquía' (Rice), middle and upper Cauca River basins (Sugar cane), Colombian Massif (Coffee), Eastern Cundinamarca and western Boyacá (Potato), the Mojana region (Cereals), and Santander, south of Bolivar and Cesar (Oil palm). Therefore, the most agriculturally specialized regions in the country are located mainly in regions with ample access to water resources and in Andean, sub-Andean, or tropical forest ecosystems (ecosystems marked by high rainfall and a tropical climate dependent on altitude).

The most diversified areas are the northern Colombian Pacific, the Amazonian foothills, the Magdalena Medio, and the department of Guaviare. These areas are characterized by their ombrophilous or sub-hygrophilous-hygrophilous equatorial forest ecosystems (ecosystems marked by high humidity and species diversity).



Source: Created by author

Figure 1: Municipal diversification in Colombia.

An analysis of the distribution of municipalities according to their agricultural diversification reveals the existence of marked natural boundaries between highly specialized and diversified regions. The main examples of this characteristic are the Eastern Cordillera, which separates the specialized municipalities of the Cauca basin from the diversified municipalities of the Colombian Pacific; the Central Cordillera

Municipalities with the lowest agricultural diversification in Colombia		Municipalities with the highest agricultural diversification in Colombia	
Municipalities	**Diversification index**	**Municipalities**	**Diversification index**
Sesquilé, Cundinamarca	0.0961846	Fómeque, Cundinamarca	0.8673471
Palmira, Valle del Cauca	0.0976485	Peñón, Antioquia	0.8675603
Puerto Tejada, Casanare	0.1063187	Sáchica, Boyacá	0.8687919
Candelaria, Valle	0.1090595	Rionegro, Antioquia	0.8703443
Villapinzón, Cundinamarca	0.1139393	Tibaná, Boyacá	0.8750623
Betania, Antioquia	0.1209051	Puerto Asís, Putumayo	0.8763599
Tausa, Cundinamarca	0.1259753	Certegui, Chocó	0.8781608
Tutazá, Boyacá	0.1300443	Medio San Juan, Chocó	0.8847447
Piendamó, Cauca	0.1309135	Guapi, Cauca	0.8871211
Padilla, Cauca	0.1357833	Remedios, Antioquia	0.8905849

Source: Author's own calculations

Table 1: Local administrative entities with the lowest and highest agricultural diversification in Colombia from 2007 to 2021.

between the specialized municipalities of the Colombian massif and the diversified municipalities of the Amazonian foothills; and the Western Cordillera between the specialized municipalities of the Altillanura and the diversified municipalities of the department of Cundinamarca. Therefore, a key conclusion of this paper is that geography and natural resource endowment influence diversification-specialization decisions.

Models

The results of the estimations of the models proposed to analyze the relationship between rural poverty and agricultural diversification are presented below, taking into account separately the spatial phenomena of: 1. autocorrelation and 2. spatial heterogeneity. Models with both spatial effects are estimated because first the relationship of interest is quantified and later it will be analyzed if it is uniform for the whole territory of the country.

Spatial dependence

Table 2 presents the modeling results with Durbin's spatial autocorrelation. The calculation was performed using the Maximum Likelihood estimation method as recommended by Anselin (1988). The model is evidenced to possess an explanatory power indicated by a Pseudo R^2 value close to 0.5. Additionally, the p-values of the Wald test indicate that the spatial terms such as the spatial lags of the dependent ($\hat{\rho}$) and independent ($\hat{\theta}$) variable have a statistically significant effect, therefore, the Durbin is preferable over the traditional linear regression model (Elhorst, 2014).

The coefficients are not directly interpretable. In a non-spatial model, the total effect on the dependent variable is the coefficient estimate holding all other variables constant and regardless of their location. Whereas, in a spatial model such as the Durbin Model, the total effect depends on both the neighboring units and the coefficients of the spatial variables (Goméz and Hernan, 2015). Therefore, within a spatial model it is relevant to isolate the effect that spatially lagged variables have on a variable of interest.

VARIABLES	(1) Spatial Durbin Model
Agricultural diversification	0.1883***
	(0.0222)
Municipal Performance	-0.3215***
	(0.0303)
% of agricultural units with an area of less than 1 hectare	-0.1421 ***
	(0.0183)
% of the population living in rural areas	0.1503***
	(0.0163)
% of agricultural units with access to irrigation systems	-0.0546***
	(0.0139)
Altitude	-0.0042***
	(0.0005)
Acts of threats (Violence indicator)	0.0150***
	(0.0036)
% of the municipal productive land underutilized	-0.0832***
	(0.02242)
Intercept	60.8806***
	(2.6560)
$\hat{\rho}$	-0.5668***
	(0.0061)
$\hat{\theta}$	0.0644***
	(0.0083)
Observations	1.084
Pseudo R^2	0.491
p-value	0.00
Wald Spatial Test	0.00

Note: Standard errors in brackets; *** p<0.01, ** p<0.05, * p<0.1
Source: Author's own calculations

Table 2: Spatial Durbin Model estimation.

In this context, Table 3 presents the marginal effects of the Durbin model. Direct effects correspond to the direct impact of an explanatory variable on the dependent variable in each geographic unit, without taking into account spatial neighborhood effects. Therefore, these effects are interpreted similarly to how an OLS model would be interpreted. On the other hand, indirect effects refer to the impact of spatially lagged variables, i.e., spatial neighborhood influencing all units in the model. Finally, the total effect refers to the aggregate impact of the model coefficients, including both direct and indirect effects (Le Saout and Floch, 2018).

There is evidence that an increase of one percentage point in the diversification index of the municipalities is related to an increase of 0.1816 percentage points (p.p.) in the rural

Variables	(1) Direct effect	(2) Indirect effect	(3) Total effect
Agricultural diversification	0.1816***	-0.1195***	0.0620**
	(0.0220)	(0.0175)	(0.0277)
Municipal Performance	-0.3246***	-0,0553***	-0.3800***
	(0.3049)	(0,0102)	(0.0351)
% of agricultural units with an area of less than 1 hectare	-0.1435**	-0.0244***	-0.1679
	(0.0184)	(0.0052)	(0.0218)
% of the population living in rural areas	0.1518***	0.0258***	0.1777***
	(0.0164)	(0.0052)	(0.0195)
% of agricultural units with access to irrigation systems	-0.0551***	-0.0094***	-0.0645***
	(0.0140)	(0.0028)	(0.0164)
Altitude|	-0.0043***	-0.0007***	-0.0050***
	(0.0005)	(0.0001)	(0.0006)
Acts of threats (Violence indicator)	0.0152***	0.0025***	0.0178***
	(0.0036)	(0.0008)	(0.0043)
% of the municipal productive land underutilized	-0.0840***	-0.0143***	-0.0983***
	(0.0226)	(0.0043)	(0.0262)

Note: Standard errors in brackets; *** p<0.01, ** p<0.05, * p<0.1
Source: Author's own calculations

Table 3: Marginal effects of the Durbin model.

component of the Multidimensional Poverty Index; while the spillover effect of neighboring municipalities is close to -0.1195 p.p. of increase in rural poverty. In other words, a higher level of diversification in a municipality is associated with an increase in rural poverty; however, if its neighbors diversify, rural poverty tends to decrease. In fact, the indirect effect almost manages to offset the increase in poverty by the direct effect.

It can be postulated that agricultural diversification can lead to a dispersion of resources and efforts, with the consequent inability of farmers to achieve the economies of scale necessary for competitiveness. This may result in lower efficiency and, consequently, high levels of poverty (Baumgärtner and Quaas, 2010). However, the positive relationship between diversification in the municipality as a whole and rural poverty can be explained, to a large extent, by the loss of economies of scale external to the agricultural unit, and not so by internal economies of scale. If internal economies of scale were significant, the increase in the number of small agricultural units (e.g., 0 to 1 hectares) would have a poverty-increasing effect, because internal economies of scale are present mainly in large farm units (Chavas, 2001).

On the other hand, indirect local effects can be explained through the impact of the specialization of neighboring municipalities in the consolidation and improvement of markets and short marketing circuits. Two possible mechanisms can be observed. First, when neighboring municipalities diversify, they generate a greater variety of products in the local market, thus boosting trade and income with the central municipality. This improves the ability of farmers in the central area to sell their products and access more diverse and potentially lucrative markets. Second, diversification into surrounding municipalities can strengthen resilience to economic and climatic risks, stabilizing markets and providing a more secure economic environment for the central municipality. This economic stability and risk reduction can translate into reduced poverty levels (Klasen et al., 2016; Sotelo, 2020).

This paper shows that the effects of agricultural diversification are dynamic since they affect not only the unitary farm or the local territory itself but also neighboring local units and territories. Therefore, the socioeconomic effects of this type of decision should be analyzed with a regional approach, and not in isolation.

For the remaining variables, a positive relationship was found between rural poverty with the proportion of the population living in rural areas and the occurrence of events involving threats of violence against the population. While, a negative relationship was found with access to irrigation systems (external to the agricultural units), quality of policy management by local authorities, the presence of small production units and percentage of underutilization.

Spatial heterogeneity

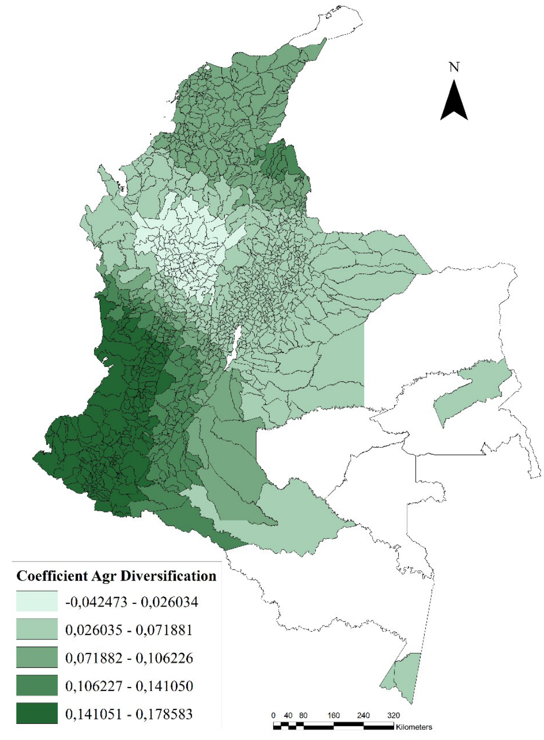
Figure 2 shows the results of the agricultural diversification coefficient of the geographically weighted regression. An increase of between one percentage point in the diversification index of the country's municipalities is related to an increase of between almost 0 and 17 percentage points (p. p.) in the rural component of the Multidimensional Poverty Index. In other words, a higher level of diversification in a municipality is related to an increase in its rural poverty.

The relationship between diversification and poverty is positive and significant for the country, but its intensity varies in different regions. Specifically, in the southwestern regions of the country, such as the Pacific, and some municipalities of the Eje Cafetero. In fact, the further west a municipality is located in the country, the stronger this relationship will be.

This relationship can be interpreted from two levels. First, Valle del Cauca and Cauca are among the departments that allocate the most planted area to agro-industrial products (more than 7.5% of the planted area of the UPAs), such as coffee, African palm, sugarcane, sugarcane, rubber, cotton, etc. (DANE, 2015). Secondly, the departments of Antioquia, Nariño, Tolima, Valle del Cauca, Meta and Cauca represent about 45% of the total harvested area of the country (DANE, 2015). In other words, there is, broadly speaking, a concentration of agroindustrial crops and agricultural production in the western part of the country.

In this sense, the western intensification of the relationship between rural poverty and diversification is explained by the relevance of the external economies obtained by concentrating production in a certain geographic area. External economies of scale are exemplified by shared infrastructure (roads, irrigation, storage, and transportation), dissemination of technical knowledge and technology, and the development of markets and product linkages encouraged by specialization in one or a few agricultural products.

In particular, it is worth highlighting the role of Valle del Cauca in the intensification of the relationship due to the loss of potential benefits from productive linkages when adopting diversification strategies.



Source: Created by the author

Figure 2: Results of the diversification coefficient.

In other words, if a local administrative entity decides to diversify in an environment of constant specialization benefits, it will be more disadvantaged in its social indicators than a municipality that does so in an environment with low external economies of scale due to specialization.

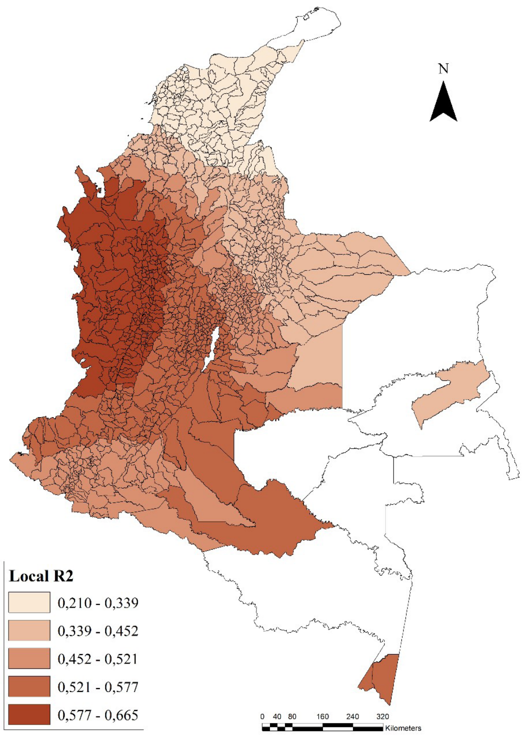
In Table 4, it is presented the diagnosis of the GWR model. In the model presented, the bandwidth is determined by the number of nearest neighbors, ensuring that each local estimation is based on the same quantity of features. Rather than specifying a particular distance, the analysis is conducted using a set number of neighbors. Additionally, it can be noted that a higher R^2 and a R^2 adjusted is preferable.

Variable	Value
Neighbors	585
Residual Squares	118997.7
Effective Number	51.47
Sigma	10.7
AICc	8259.34
R^2	0.5979
R^2 adjusted	0.5783

Source: Author's own calculations

Table 4: Diagnosis of the GWR model.

On the other hand, Figure 3 presents the local R^2 that measures the fit of the model at regression point i and its capacity to replicate the data in the vicinity of point i (de Bellefon and Floch, 2018, p. 216).



Source: Created by the author

Figure 3: Local R^2.

In this map it is evident that the proposed model has a higher explanatory capacity in the northwest of the country with a Local R^2 close to 0.6; while, in the other regions it is close to 0.5.

Conclusion

Economic studies that analyze the relationship between poverty and agricultural diversification usually focus on the analysis of production within an agricultural unit and its same social outcomes (Seo, 2010; Ullah and Shivakoti, 2014; Thapa et al., 2018; Sekyi et al., 2021). This paper, from an innovative approach, analyzes the relationship from the local scale, because it recognizes the importance of spatial effects and externalities that transcend agricultural units.

The evidence presented suggests that municipal-scale agricultural diversification is associated with higher levels of rural poverty. However, when analyzing the spatial effects (autocorrelation and spatial heterogeneity), it was found that this relationship is not uniform across the entire national territory and has opposite spillover effects when analyzing the neighboring municipalities that offset the positive effect.

The positive relationship between rural poverty and agricultural diversification is not uniform throughout the country. Specifically, the relationship is intensified in the western part of the country, including the Pacific region, and some municipalities of the Eje Cafetero. This situation is due, broadly speaking, to a concentration of agroindustrial crops and agricultural production in the western part of the country. In this sense, the western intensification of the relationship between rural poverty and diversification is explained by the relevance of external economies that are obtained by concentrating production in a certain geographic area and are lost by adopting diversification strategies.

The results obtained demonstrate the importance of considering the differences and particularities of each territory in the formulation of sectoral public policies. To this extent, it is essential to approach territorial convergence from the grouping of regions with similar characteristics, opportunities and challenges in "convergence clubs", as discussed in the literature on regional economic convergence (Royuela and García, 2015; Vásquez and Bara, 2009). Although this literature focuses on regional convergence in terms of economic growth, its focus on the formulation and approach of heterogeneous solutions according to "convergence clubs" should be equally addressed in sectoral policies, as in this case agriculture.

The paper's findings are framed by global discussions favoring agricultural diversification as a measure for adaptation and mitigation of climate change effects in agricultural areas (FAO et. al., 2022; IFPRI, 2022). However, it also raises concerns about the economic sustainability and social impacts of this strategy on rural communities, which are already highly vulnerable. Thus, in these discussions around the impact of climate change in rural areas, adaptation and mitigation strategies must be reconciled with their potential social benefits or detriments. A certain degree of specialization in a municipality generates benefits for the farmer and his municipality, provided that there is greater agricultural diversity at the higher scale of analysis (province, subregion, department) to mitigate environmental damage and ensure the sustainability of his activity. Policy recommendations should be oriented toward promoting more tailored and targeted local and regional agricultural policies that coordinate clusters and different levels of specialization; in addition, some social protection measures should be coordinated with agricultural diversification initiatives to mitigate the potential adverse effects of diversification strategies.

Corresponding author:
Alejandro Mojica Godoy
Universidad Externado de Colombia
Calle 12#1-17E, Bogotá, Colombia
Email: alejandro.mojica@est.uexternado.edu.co

References

[1] Anselin, L. (1988) "*Spatial Econometrics: Methods and Models*", Kluwer Academic Publishers, E- ISBN 978-94-015-7799-1. DOI 10.1007/978-94-015-7799-1.

[2] Anselin, L. (2010) "Thirty years of spatial econometrics", *Papers in Regional Science*, Vol. 89, No. 1, pp. 3-25. ISSN 1056-8190. DOI 10.1111/j.1435-5957.2010.00279.x.

[3] Arbia, G. (2006) "*Spatial econometrics: statistical foundations and applications to regional convergence*", Berlin, Springer Berlin Heidelberg. E-ISBN 978-3-540-32305-1. DOI 10.1007/3-540-32305-8.

[4] Baumgärtner, S. and Quaas, M. F. (2010) "Managing increasing environmental risks through agrobiodiversity and agrienvironmental policies", *Agricultural Economics*, Vol. 41, No. 5, pp. 483-496. ISSN 1574-0862. DOI 10.1111/j.1574-0862.2010.00460.x.

[5] Berry, A. (2023) "El papel clave de la pequeña agricultura familiar en Colombia", Editorial Universidad del Rosario. 530 p. ISBN 978-958-500-065-0. (In Spanish).

[6] Brunsdon, C., Fotheringham, S. and Charlton, M. (1996) "Geographically weighted regression", *Journal of the Royal Statistical Society*, Vol. 4, pp. 431-443. ISSN 1467-985X. DOI 10.1111/1467-9884.00145.

[7] da Silva, G. S., Amarante, P. A. and Amarante, J. C. A. (2022) "Agricultural clusters and poverty in municipalities in the Northeast Region of Brazil: A spatial perspective", *Journal of Rural Studies*, Vol. 92, pp. 189-205. ISSN 1873-1392. DOI 10.1016/j.jrurstud.2022.03.024.

[8] Chavas, J.-P. (2001) "Structural Change in Agricultural Production: Economics, Technology and Policy", *Handbook of Agricultural Economics*, Vol. 1A, pp. 263-285. ISSN 1574-0072. DOI 10.1016/S1574-0072(01)10008-3.

[9] Departamento Administrativo Nacional de Estadísticas – DANE. (2014) "Censo Nacional Agropecuario 2014, Sostenibilidad Ambiental en las Unidades de Producción Agropecuaria (UPA), Resultados entrega 08 – octubre 20 de 2015", DANE, Oct. 2015. [Online]. Available: https://www.dane.gov.co/files/CensoAgropecuario/avanceCNA/PPT8-Boletin8.pdf [Accessed: Jul. 24, 2025]. (In Spanish).

[10] De Bellefon, M. and Floch, J. (2018) "Geographically Weighted Regression", In: Loonis, V. (ed.) "*Handbook of spatial analysis: theory and application with R*". pp. 237-239. ISBN 978-2-11-139686-9.

[11] Elhorst, J. P. (2014) "*Spatial econometrics: from cross-sectional data to spatial panels*", Springer. E-ISBN 978-3-642-40340-8.

[12] FAO, IFAD, UNICEF, WFP, and WHO. (2022) "The State of Food Security and Nutrition in the World 2022: Repurposing food and agricultural policies to make healthy diets more affordable", FAO. [Online]. Available: https://openknowledge.fao.org/server/api/core/bitstreams/67b1e9c7-1a7f-4dc6-a19e-f6472a4ea83a/content [Accessed: Jul. 24, 2025]. E-ISSN 2663-807X. ISBN 978-92-5-136499-4.

[13] Floch, J. M., and Le Saout, R. (2018) "Spatial econometrics-common models", In: Loonis, V. (ed.) "*Handbook of spatial analysis: theory and application with R"*. pp. 151-177. ISBN 978-2-11-139686-9.

[14] Fotheringham, A. S., Brunsdon, C., and Charlton, M. E. (2009) "Geographically weighted regression", In: "*The SAGE Handbook of Spatial Analysis*", Vol. 1, pp. 243-254. ISBN 9781412910828. DOI 10.4135/9780857020130.n13.

[15] Galvis, L. A. and Meisel, A. (2010) "Persistencia de las desigualdades regionales en Colombia: Un análisis espacial", *Documentos de Trabajo Sobre Economía Regional y Urbana*, Banco de la República, Vol. 120. ISSN 1692-3715. (In Spanish).

[16] Giller, K. E., Delaune, T., Silva, J. V., Descheemaeker, K., van de Ven, G., Schut, A. G. T., van Wijk, M., Hammond, J., Hochman, Z., Taulya, G., Chikowo, R., Narayanan, S., Kishore, A., Bresciani, F., Teixeira, H. M., Andersson, J. A. and van Ittersum, M. K. (2021) "The future of farming: Who will produce our food?", *Food Security*, Vol. 13, No. 5, pp. 1073-1099. ISSN 1876-4525. DOI 10.1007/s12571-021-01184-6.

[17] Gómez, H. and Hernan, M. (2015) "Econometría espacial usando Stata. Breve guía aplicada para datos de corte transversal", *Documentos de Trabajo del Instituto de Estudios Laborales y del Desarrollo Económico*. Vol. 1. ISSN 1852-1118. (In Spanish).

[18] Hufnagel, J., Reckling, M. and Ewert, F. (2020) "Diverse approaches to crop diversification in agricultural research. A review", *Agronomy for Sustainable Development*, Vol. 40, pp. 1-17. ISSN 1773-0155. DOI 10.1007/s13593-020-00617-4.

[19] International Food Policy Research Institute (IFPRI). (2022) "2022 Global food policy report: Climate change and food systems", International Food Policy Research Institute (IFPRI). DOI 10.2499/9780896294257.

[20] Klasen, S., Meyer, K. M., Dislich, C., Euler, M., Faust, H., Gatto, M., Hettig, E., Melati, D. N., Jaya, I. N. S., Otten, F., Pérez-Cruzado, C., Steinebach, S., Tarigan, S. and Wiegand, K. (2016) "Economic and ecological trade-offs of agricultural specialization at different spatial scales", *Ecological Economics*, Vol. 122, pp. 111-120. ISSN 1873-6106. DOI 10.1016/j.ecolecon.2016.01.001.

[21] LeSage, J. and Pace, R. K. (2009) "*Introduction to Spatial Econometrics*",1st ed., Taylor & Francis. New York, 340 p. E-ISBN 9780429138089. DOI 10.1201/9781420064254.

[22] Lobao, L. and Sharp, J. (2013) "Agriculture and rural development", In. Green, G. P. (ed.) "*Handbook of rural development*", pp. 115-138. ISBN 9781781006702. DOI 10.4337/9781781006719.00016.

[23] Palmer-Jones, R. and Sen, K. (2006) "It is where you are that matters: the spatial determinants of rural poverty in India", *Agricultural Economics*, Vol. 34, No. 3, pp. 229-242. ISSN 1574-0862. DOI 10.1111/j.1574-0864.2006.00121.x.

[24] Perfetti, J. J. and Leibovich, J. (2013) "*Propuesta de estructura orgánica básica del Ministerio de Agricultura y Desarrollo Rural*". [Online]. Available: http://www.repository.fedesarrollo.org.co/handle/11445/381 [Accessed: Jul. 24, 2025]. (In Spanish).

[25] Pielou, E. C. (1966) "Species-diversity and pattern-diversity in the study of ecological succession", *Journal of Theoretical Biology*, Vol. 10, No. 2, pp. 370-383. ISSN 0022-5193. DOI 10.1016/0022-5193(66)90133-0.

[26] Rahmawati, Y., Ichsan, A. K. N., Brintanti, A. R. D. and Jamil, I. R. (2023) "Geo-spatial analysis: the impact of agriculture productivity, drought, and irrigation on poverty in East Java, Indonesia", *Letters in Spatial and Resource Sciences*, Vol. 16, No 1. ISSN 1864-404X. DOI 10.1007/s12076-023-00348-6.

[27] Royuela, V. and García, G. A. (2015) "Economic and social convergence in Colombia", *Regional Studies*, Vol. 49, No. 2, pp. 219-239. ISSN 0034-3404. DOI 10.1080/00343404.2012.762086.

[28] Sekyi, S., Quaidoo, C. and Wiafe, E. A. (2021) "Does crop specialization improve agricultural productivity and commercialization? Insight from the Northern Savannah Ecological Zone of Ghana", *Journal of Agribusiness in Developing and Emerging Economies*, Vol. 13, No. 1., pp. 16-25. ISSN 2044-0847. DOI 10.1108/JADEE-01-2021-0021.

[29] Seo, S. N. (2010) "Is an integrated farm more resilient against climate change? A micro-econometric analysis of portfolio diversification in African agriculture", *Food Policy*, Vol. 35, No. 1, pp. 32-40. ISSN 1873-5657. DOI 10.1016/j.foodpol.2009.06.004.

[30] Sotelo, S. (2020) "Domestic Trade Frictions and Agriculture", *Journal of Political Economy*, Vol. 128, No. 7, pp. 2690-2738. ISSN 0022-3808. DOI 10.1086/706859.

[31] Thapa, G., Kumar, A., Roy, D. and Joshi, P. K. (2018) "Impact of Crop Diversification on Rural Poverty in Nepal", *Canadian Journal of Agricultural Economics*, Vol. 66, No. 3, pp. 379-413. ISSN 1744-7976. DOI 10.1111/cjag.12160.

[32] Ullah, R. and Shivakoti, G. P. (2014) "Adoption of On-Farm and Off-Farm Diversification to Manage Agricultural Risks: Are These Decisions Correlated?", *Outlook on Agriculture*, Vol. 43, No. 4, pp. 265-271. ISSN 0030-7270. DOI 10.5367/oa.2014.0188.

[33] Vásquez, L. F. and Bara, J. L. R. (2009) "Convergencia económica regional: el caso de los Departamentos colombianos", *Ecos de economía*, Vol. 13, No. 28, pp. 167-197. ISSN 1657-4206. (In Spanish).

[34] Waha, K., Accatino, F., Godde, C., Rigolot, C., Bogard, J., Domingues, J. P., Gotor, E., Herrero, M., Martin, Guillaume, Mason-D´Croz, D. and van Wijk, M. (2022) "The benefits and trade-offs of agricultural diversity for food security in low-and middle-income countries: A review of existing knowledge and evidence", *Food Security*, Vol. 33., p. 100645. ISSN 2211-9124. DOI 10.1016/j.gfs.2022.100645.

[35] Wardhana, D., Ihle, R. and Heijman, W. (2017) "Agro-clusters and Rural Poverty: A Spatial Perspective for West Java", *Bulletin of Indonesian Economic Studies*, Vol. 53 No. 2, pp.161-186. ISSN 0007-4918. DOI 10.1080/00074918.2017.1298722.

Appendix

In this appendix, the description of the variables used can be found (Table 5).

Additionally, to ensure that the agricultural diversification index is not a function of time, unit root tests for panel data were calculated using the Fisher Dickey-Fuller Augmented Fisher test (Maddla and Wu, 1999) and the Fisher Phillips-Perron test (Choi, 2001). Table 5 shows that with both tests and all statistics the null hyptothesis of the existence of unit roots is rejected. In other words, municipal agricultural diversification in Colombia has a constant mean and variance over time and does not show a trend in the last 15 years (Banerjee, 1999).

Variable	Description	Year	Source
Rural component of the Multidimensional Poverty Index (MPI) at the municipal (local) level.	The MPI is composed of 15 variables in 5 dimensions (educational conditions, children and youth, health, work, access to public utilities and housing conditions). In 2020, based on information from the National Population and Housing Census, multidimensional poverty was calculated at the municipal level (DANE, 2020).	2020	Departamento Administrativo Nacional de Estadística - DANE
Management component of the municipal performance measurement indicator.	Municipal Performance Measurement is an indicator calculated by the DNP that has two components: management and results: 1. mobilization, 2. execution of own resources, 3. attention to citizens and accountability, and 4. management of land use planning instruments.	2021	Departamento Nacional de Planeación
Agricultural Production Units within the municipality that have an extension between 0 and 1 hectares.	The percentage of agricultural production units in the municipality that have between 0 and 1 hectare of land area.	2014	National Agricultural Census, DANE
Percentage of UPA within the municipality with access to irrigation systems.	Percentage of agricultural production units within the municipality that have access to an irrigation system.	2014	National Agricultural Census, DANE
Percentage of land use that is in conflict due to underutilization.	The percentage of the municipality's land use that is in conflict due to underutilization with respect to the land use vocation.	2012	Departamento Nacional de Planeación based on information from the Instituto Geográfico Agustín Codazzi.
Presence of industrial and/or exportable crops.	Dummy variable that takes the value of 1 if at least 10% of the municipality's cultivated land is destined to the cultivation of industrial and exportable products (subject to industrial transformation and/or export), such as coffee, sugar cane, sugar cane, cocoa, flowers, oil palm, cotton, timber and banana.	2007-2021	Own calculations based on EVAs
Percentage of the population living in rural areas.	Percentage of the municipality's population living in rural areas	2022	Departamento Administrativo Nacional de Estadísticas - DANE
Altitude of the municipality measured in meters above sea level.	Dummy variable that takes the value of 1 if at least 10% of the municipality's cultivated land is destined to the cultivation of industrial and exportable products (subject to industrial transformation and/or export), such as coffee, sugar cane, sugar cane, cocoa, flowers, oil palm, cotton, timber and banana.	2007-2021	Own calculations based on EVAs

Source: Author's own elaboration

Table 5: Description of the variables employed.

FISHER CHI-SQUARE - ADF			FISHER CHI-SQUARE - PP		
Test	Value	P-value	Test	Value	P-value
Inverse chi-squared	4295.6982	0.0000	Inverse chi-squared	4485.0296	0.0000
Inverse normal	-20.8552	0.0000	Inverse normal	-19.7097	0.0000
Inverse logit t	-23.8588	0.0000	Inverse logit t	-23.9772	0.0000
Modified inv. chi-squared	31.7276	0.0000	Modified inv. chi-squared	34.5858	0.0000

Source: Author's own elaboration

Table 6: Tests of stationarity of the municipal agricultural diversification index.

Volume XVII Number 4, 2025

Data-Driven Optimisation of Irrigation Dose Using Machine-Learning Ensembles for Sustainable European Agriculture

Zuzana Palková[1,2,3], Miroslav Žitňák[1], Jan Valíček[1,2], Marta Harničárová[1,2], Miroslav Holý[1], Daniel Levák[1], Hakan Tozan[4], Karol Görči[1]

[1] Institute of Electrical Engineering, Automation, Informatics and Physics, Faculty of Engineering, Slovak University of Agriculture in Nitra, Slovakia

[2] Department of Mechanical Engineering, Faculty of Technology, Institute of Technology and Business in České Budějovice, Czech Republic

[3] New Edu, n.o., Nitra, Slovakia

[4] College of Engineering and Technology, American University of the Middle East, Egaila, Kuwait

Abstract

This study focuses on predicting irrigation doses using digital technologies and statistical modelling to enhance water resource management in agriculture. Conducted as part of the CODECS project in the semi-arid Nitra region of Slovakia, this study aimed to evaluate the effectiveness of various irrigation systems and to develop predictive models for optimal irrigation doses. The methodology integrates environmental sensor data, agronomic models, and machine learning techniques, utilizing IoT sensors alongside Valley and Irriga control software. A significant challenge was the incompatibility of heterogeneous data from different sources, leading to the creation of a unified method-ology for data collection, validation, and analysis. Analytical tools, such as ex-ploratory data analysis, correlation techniques, and regression models, were employed to identify key factors affecting irrigation efficiency, including precipitation, temperature, soil moisture, and energy consumption. The findings aim to inform sustainable irrigation strategies that reduce water usage, enhance crop productivity, and safeguard soil resources under changing climatic con-ditions.

Keywords

Artificial irrigation, digital agriculture, machine learning, data, statistical modelling, smart farming.

Palková, Z., Žitňák, M., Valíček, J., Harničárová, M., Holý, M, Levák, D., Tozan, H. and Görči, K. (2025) "Data-Driven Optimisation of Irrigation Dose Using Machine-Learning Ensembles for Sustainable European Agriculture", *AGRIS on-line Papers in Economics and Informatics*, Vol. 17, No. 4, pp. 81-99. ISSN 1804-1930. DOI 10.7160/aol.2025.170407.

Introduction

The agricultural sector is the largest consumer of water worldwide (SOLAW, 2021; FAO, 2021). Agricultural productivity, coupled with water scarcity, is becoming an increasing problem, particularly in arid areas, often leading to land degradation. Improved on-farm management practices and proper water and soil management contribute to better crop yields, growth, development, and yield, which depend heavily on a sufficient water supply. AI-powered irrigation systems can reduce water usage by up to 25 %, demonstrating the significant potential of technological interventions in water management (Akkem et al., 2023). Khan et al. (Khan et al., 2006) argued that irrigation is also used to improve the quality and diversity of crops, and not only to increase the quantity of crops produced. However, the benefits of irrigation also bring challenges, such as the proper management of water resources, prevention of soil salinization, and issues regarding sustainability.

Precision irrigation scheduling is directed toward efficient water usage for each plant, where and when it is needed, in the right amounts, to compensate for water loss either through evapotranspiration, erosion or deep percolation, while preventing over- and under-irrigation (Ahmed et al., 2024; García et al., 2020; Vallejo-Gómez et al., 2023). Recent advances in precision agriculture have shown that precision irrigation systems can reduce water usage by up to 50 % compared with traditional methods in agriculture (Hamami and Nassereddine, 2020). Furthermore, agricultural evapotranspiration (ET) represents the amount of water that is "consumed" by agriculture, as this water leaves the watershed entirely through evaporation from soil and transpiration by crops (Sithartan et al., 2023).

There is a growing need for advanced irrigation management solutions to ensure efficient use of available water resources and minimize losses. One of the key strategies in this regard is irrigation modelling, which enables more accurate prediction and optimization of irrigation practices. Jensen et al. (1990) defined irrigation modelling as a tool that uses mathematical equations and computational techniques to create a tool for design, manage, and optimize irrigation strategies. There are different types of irrigation models and each has specific applications, methodologies, and benefits. Different models have been used to predict the water flow and manage irrigation systems under unexpected conditions.

Technical innovations have also brought about advances in irrigation modelling, which have increased their computational power. Allen et al. (Al, 2006) revolutionized irrigation planning and management through the development of spatial systems that use geographic information systems (GIS) and remote sensing. Palková and Rodny (Palková and Rodny, 2018) in their article stated that the basic idea of modelling a real process involving irrigation is possible by simulating real processes through stochastic models that incorporate randomness and uncertainty and are thus essential for managing unpredictability, such as climate change.

Advances in irrigation optimization have also been brought about by artificial intelligence (AI), which is increasingly being integrated into these models and enabled by machine learning, increasing their adaptability and decision-making accuracy. Kersebaum et al. (Kersebaum et al., 2007) pointed to the use of machine-learning algorithms that have improved the efficiency of irrigation models, including more accurate predictions of irrigation needs. AI-guided human–machine collaboration can streamline the integration of user needs, allowing customization towards situational farm management adaptation (Wei et al., 2024).

Recent studies confirmed that artificial intelligence (AI) and machine learning (ML) play key roles in irrigation optimization. Umutoni and Samadi (2024) analyzed 16 studies on ML applied to irrigation demand prediction and highlighted the need to integrate physical models with ML for more accurate and robust decision making. Hossain (Hossain, 2023) presented advanced machine-learning techniques combined with ML to optimize irrigation canal operations, which improved water distribution and increased management efficiency. Bellingham et al. (2023) developed ML regression models to predict soil moisture in drip irrigation systems, and achieved high accuracy in predicting soil moisture in the near future. Precision agriculture, a key application of AI in smart farming, involves the use of sensors, drones, and satellite imagery to monitor crop health, soil conditions, and weather patterns (Sharma et al., 2021).

Recent developments in IoT technology have led to advanced applications in agriculture (Kumar Kasera et al., 2024; Prasath and Akila, 2023). García et al. (2020) provided a comprehensive overview of IoT-based smart irrigation systems, emphasizing recent trends in sensors and IoT systems for precision agriculture. According to research reports, the number of connected agricultural devices is expected to increase from 13 million at the end of 2014 to 225 million by 2024 (García et al., 2020). The integration of real-time data collection through IoT devices enables accurate irrigation management and improves decision-making processes (Bwambale et al., 2022).

Advanced wireless sensor networks have been developed for the comprehensive monitoring of agricultural parametres (Karray et al., 2020). The rebound in 2024 suggests a renewed focus on IoT in irrigation management, possibly reflecting emerging challenges, such as climate change adaptation, water resource optimization, integration with AI-driven decision support systems, and the emergence of low-cost solid-state sensors (Abdelmoneim et al., 2025). These investigations demonstrate that the integration of ML models into drip irrigation systems can yield more accurate predictions and greater water resource efficiency, leading to sustainable and efficient agricultural management.

To use AI and ML models effectively, it is essential to obtain high-quality and relevant data that improve the accuracy of predictions and irrigation optimization. Data on climate, precipitation, temperature, humidity, and wind speed are among the key factors for accurate irrigation management, as reported in a study of evapotranspiration (Allen et al., 2006). Understanding and accurately measuring evapotranspiration can significantly improve water use efficiency in agriculture, ensuring sustainability and crop productivity in both the short and long run (Gao et al., 2024)

The application of AI techniques and robotics in agriculture has continued to evolve (Wakchaure et al., 2023). Big data and the AI revolution in precision agriculture present both opportunities and challenges (Bhat and Huang, 2021). Data on soil texture, structure, and moisture content are essential for accurately determining irrigation requirements (Moriasi et al., 2007). It is also important to provide an overview of water sources, quantity, and quality (Bastola et al., 2011). The integration of field measurements and modelling into irrigation management has been analyzed in a recent study that demonstrated effective irrigation management under different climatic conditions.

Significant advances have been made in the development of technologies for data collection in the field of irrigation management. New methods integrate artificial intelligence and IoT sensors for better monitoring of soil moisture and irrigation requirements. Data accuracy and reliability are key for efficient irrigation management. The need for high-quality input data for accurate modelling and decision-making has been emphasized (Crystal-Ornelas et al., 2022)The impact of data quality on the sustainability and ecological suitability of irrigation systems has been highlighted (Jihani et al., 2024).

Research can be more transparent and collaborative using Findable, Accessible, Interoperable, and Reusable (FAIR) principles to publish Earth and environmental science data (Crystal-Ornelas et al., 2022). The implementation of FAIR data practices in irrigation management can significantly enhance the effectiveness of predictive models by improving data sharing, integrating expertise, and developing standardized data formats (Umutoni and Samadi, 2024). For AI and ML models to work effectively, it is important that these data have a uniform structure and are stored in central databases, from which models can process them efficiently and optimize irrigation practices.

A generic smart irrigation system can be schematically represented as consisting of sensors, Internet of Things (IoT) technologies, and algorithms. Sensors sense the environment and the soil conditions, and adapt and send the sensed values to local or remote processing units in which trained machine learning algorithms can forecast agrotechnical indicators useful to decide if irrigation is needed and eventually the optimal amount of water to be provided to the crops (Kaur et al., 2024). This smart irrigation system allows farmers to control and adjust irrigation practices remotely, with systems achieving up to 60% water savings through automated monitoring of soil moisture and environmental conditions (Dong et al., 2024).

Owing to the impact of climate change on agriculture and the emergence of water security issues, proper irrigation management has become increasingly important to overcome these challenges. Internet of Things (IoT) technology is being utilized in agriculture to collect field information and share it through websites in real time (Abioye et al., 2022).

The prediction of irrigation application rates using digital technologies and statistical modelling has gained significant attention owing to the growing need for efficient water management in agriculture. Various approaches have been developed to improve irrigation scheduling and optimize water consumption, particularly by using artificial intelligence (AI) and machine learning (ML). These methodologies not only improve crop yields but also contribute to sustainable agricultural practices.

Machine learning is a rapidly evolving technology for precision irrigation systems because of its ability to mimic human decision-making, while also addressing the complexities of agricultural environments (Abioye et al., 2020). Agricultural water use accounts for 70% of the total water withdrawal worldwide. The evapotranspiration during crop growth is one of the important hydrological processes in the agricultural water cycle (Boser et al., 2024). AI-based digital systems for intelligent irrigation continue to advance (Raouhi et al., 2023). AI and ML techniques, such as support vector machines and random forests, have been effectively used to predict irrigation needs based on factors, such as soil moisture and temperature (Jones, 2024).

Evapotranspiration (ET) isn't just a scientific term, it's the foundation of efficient irrigation. Understanding and applying ET helps maximize yields while improving water productivity (Wakchaure et al., 2023). The use of evapotranspiration data to fine-tune irrigation scheduling helps to avoid wasteful overwatering and protects precious water supplies. Farmers can reduce their water use and energy bills by using precision irrigation based on ET rates (Raouhi et al., 2023).

Modern technologies, particularly remote sensing and accurate hyperlocal weather data, enable farmers to minimize the time and effort required to calculate and monitor evapotranspiration levels, particularly in large fields (Arulraj and Karthikeyan, 2024).

A significant barrier to the adoption of ML in irrigation is the limited availability of quality data, which hinders its accuracy (Umutoni and Samadi, 2024). The complexity of AI models can lead to challenges in understanding their predictions, which requires the development of more interpretable models (Bhat and Huang, 2021). Combining process models with ML can improve the prediction accuracy and address data limitations (Umutoni and Samadi, 2024).

Various machine learning-based irrigation models have been proposed to minimize water waste. AIDSII, an AI-powered digital application that leverages IoT-based precision agriculture and CNN-LSTM models, offers a comprehensive feedback system through mobile and web technologies, enabling farmers to automate, optimize, and streamline their irrigation processes (Arulraj and Karthikeyan, 2024).

Innovative methods, such as those that predict soil moisture levels, can provide early warnings for irrigation and ensure optimal water use (Jones, 2024). Although digital technologies and statistical modelling offer significant advantages for irrigation management, challenges remain, such as the lack of data, interpretability of models, and the need for physics-informed models. Addressing these issues through improved data sharing, integration of expertise, and the development of FAIR data practices can increase the effectiveness of predictive models (Umutoni and Samadi, 2024).

Traditional irrigation practices face several challenges, ranging from inefficient water usage and overwatering to inadequate distribution systems. This results in water wastage, increased costs, and environmental concerns. Technological innovations address these challenges through precision and micro irrigation systems, soil sensors, etc. (Arulraj and Karthikeyan, 2024). Standardization, power efficiency, security, scalability, cost reduction, and user-centered design are among the critical factors that determine the success of IoT-enabled irrigation solutions (Gao et al., 2024).

The integration of AI, IoT, and precision agriculture technologies represents a paradigm shift toward sustainable water management in agriculture. Future research should focus on developing cost-effective IoT solutions, strengthening security frameworks, and enhancing user-friendly interfaces in order to promote mass adoption. By fostering interdisciplinary research collaborations and leveraging advancements in sensor technology, data analytics, and machine learning, the future of smart irrigation can be both technologically robust and economically viable, ensuring sustainable agricultural water management for years to come.

Materials and methods

The objective of this study was to analyze the efficiency of selected types of irrigation systems under realistic conditions based on sensor data and operational records and to develop prediction models that can reliably predict irrigation efficiency based on the recommended application rate, applied rate, and natural precipitation. As a result, key factors influencing irrigation decision making were identified, and a methodology for their efficient collection and processing was proposed.

The research is carried out within the CODECS (Maximising the CO-benefits of Agricultural Digitalisation Through Conducive Digital Ecosystems) project at the Slovak Living Lab. Within the CODECS project, the Living Lab is a network of farmers, knowledge brokers, stakeholders, and policymakers who come together around an emerging problem in a specific application scenario and are willing to develop solutions through collaboration.

The Slovak Living Lab is a part of the Gamota Group, which operates in agriculture, forestry, fish farming, and foreign trade. The company’s main activities include premium production, GMO-free virgin soybean oil (GamoSoy), and high-quality soybean expellers (SoyProFat). The company was founded in 2011 as a small family business that focused on purchasing and exporting agricultural products and integrating smaller businesses providing input products (fertilizers and chemicals) and services such as fertilization, tillage, and harvesting. Owing to its rapid development and strategic investments, the company has become one of the leading players in the Slovak agricultural sector with the aim of developing sustainable solutions for the agrosector’s future.

Gamota Group has become a key partner in the Living Lab for Irrigation Management in Slovakia, which focuses on the research and application of sustainable irrigation strategies that simultaneously help minimize water consumption, increase crop production efficiency, and protect soil resources.

The research was conducted in the Nitra region in the town of Hurbanovo, which is characterized by a semi-arid climate and requires supplementary irrigation to ensure optimal crop growth. This location was chosen because of the existing irrigation infrastructure and the interest in introducing innovative solutions for water management. Figure1 shows Living Lab Gamota research farms, where digital technologies, IoT sensors, and machine learning are being tested for precise irrigation management. The goal of this study was to minimize water consumption by predicting irrigation application rates based on current environmental conditions, thereby contributing to greater cropping efficiency and agricultural sustainability in the region. The field study was conducted throughout the 2023 growing season (spring-autumn), allowing seasonal changes in irrigation demand and their impact on soil moisture to be recorded.

The main methodology involved the integration of environmental sensor data, agronomic models, and machine-learning techniques to develop a model capable of predicting optimal irrigation application rates. The experimental design included field trials using IoT sensors and Valley and Irriga control software. Individual sensors provided real-time data on soil moisture levels, which were correlated with meteorological variables such as solar radiation, rainfall, humidity, temperature, and wind speed, obtained from local weather stations.

Pivot irrigation systems have been installed on most plots in Living Lab Gamota and are the dominant technology in precision irrigation. The other fields used pilot systems or were managed without irrigation. The fields were mainly maize (48 %), followed by peas (31 %), and wheat (11 %), with barley, beans, and buckwheat making up the remainder. Digital solutions are implemented on the farm, including automated irrigation systems and Valley and Irriga Global software, which enable optimized irrigation management based on sensory and meteorological data. The analysis confirmed that the Valley and Irriga Global systems provide heterogeneous and hardly compatible data that require complex pre-processing before they can be effectively used in decision-making processes. Therefore, this research focused on designing a unified methodology for data collection, validation, and analysis usable in artificial intelligence (AI) and machine learning (ML) predictive models, specifically for the purpose of determining optimal irrigation application rates. Several analytical tools have been applied to support this objective, including exploratory data analysis (EDA), correlation techniques, and regression models. Statistical analysis allows the identification of key factors influencing irrigation efficiency, such as atmospheric precipitation, temperature, soil moisture, recommended and applied rates, and energy consumption. Correlation analysis helps to reveal the links between variables and supports the selection of the most appropriate predictors. Regression models allow the generation of quantified estimates of irrigation needs based on real conditions.

The methodological framework involves collecting and preprocessing data from different sources, cleaning, normalizing, removing outliers, and integrating them into a common analytical model. Significant attention has been paid to selecting appropriate machine learning models and tuning them for the highest possible prediction accuracy. The main data sources are as follows:

- **Land data:** location, acreage, crops grown, history of interventions.
- **Meteorological data:** microclimatic variables from automatic stations (precipitation, temperature, wind, and humidity).
- **Irrigation systems** (e.g., pivots) include location, system type, frequency, and extent of application.
- **Agrotechnical records:** irrigation rates, measured soil moisture, data from agronomists.
- **Satellite and tabulated values:** NDVI, reference soil moisture, evapotranspiration.

For modelling purposes, daily temporal granularity was chosen, which allowed the synchronization of data from different sources into a single dataset suitable for the statistical analysis and training of machine learning models. Data integration (called data fusion) was designed to respect different measurement frequencies and spatial resolutions, thus ensuring the consistency of the inputs for further calculations.

Within the concept of precision agriculture, in which irrigation benefit modelling is an integral part, the spatial accuracy of the data is extremely important. In practice, this means that the collected data must have a well-defined spatial resolution and must be related to a specific location, typically at the level of a single pivot or field. The extent of the "small neighbourhood, " to which the data are related, depends on the specific application, but in our case, the smallest homogeneous unit is the area farmed by a single irrigation system.

Therefore, when collecting data at the pivot level, it is ideal to collect data for each system separately: crop growth, electricity consumption, irrigation application rate, size and date, microclimatic data, soil moisture values, and other agronomically relevant variables.

In this context, data fusion is the process of combining different types of data (technical, meteorological, agronomic, and satellite) into a single integrated dataset that can then be used for statistical analysis and machine learning. For successful fusion, it is crucial that the data are compatible in terms of both the temporal and spatial resolutions

Results and discussion

Data preprocessing is a key step in ensuring the consistency, quality, and usability of data for statistical modelling. This process involves removing missing values, identifying outliers, transforming the data into an appropriate format, and standardizing the selected variables. The goal was to minimize the noise, reduce the risk of bias, and maximize the reliability of the model outputs.

Because data fusion was implemented in the previous phase, further work was being performed at the individual attribute level. Some variables, such as tabular recommendation values, are left untouched or adjusted in analogy with measured values. The data are primarily used as indicative and do not usually enter directly into the model.

The numerical variables were subjected to a distribution check, search for missing data or outliers, and subsequent consideration for inclusion in the model. The removal or imputation of data depends on the type of attribute, its importance, and frequency of missingness. The decision on model inclusion is based on correlation analysis or by experimentally comparing the performance of models with different combinations of variables.

It is important not only to evaluate the correlation between the target variable (e.g., irrigation rate) and the input variables but also between the inputs themselves because of the risk of multicollinearity. The inclusion of highly correlated variables may negatively affect the accuracy and interpretability of a model.

Numerical variables can be standardized (e.g., change to z-score) or normalized (scaling to the 0–1 interval), particularly when used in scale-sensitive models. This step is particularly important when working with the variables in different units.

Categorical variables are particularly useful for low cardinality (e.g., crop type), where they are clearly interpretable. At higher cardinalities, the information value may be lost, and model complexity may increase. Attributes are encoded using one-hot or label encoding.

Modelling irrigation doses

Modelling represents the final stage of the analytical process, in which statistical and machine learning algorithms are used to create predictive models. The aim was to identify the relationships between the environmental variables and the target value, the optimal irrigation rate, and to use these relationships to make decisions under realistic conditions.

Regression techniques were used to quantify the influence of individual factors while optimizing the input data in terms of predictive accuracy. An important benefit of this phase is the feedback on the data collection process, where the best-performing models indicate which variables have the greatest impact and can be excluded in the future, thus simplifying the entire collection process.

The modelling also allows the simulation of different irrigation scenarios and their impact on efficiency and water consumption, thus supporting decision making in the spirit of sustainability and precision agriculture.

Depending on the problem formulation, two main approaches can be applied in irrigation benefit modelling: regression and classification.

Regression models predict a specific numerical value, such as the required irrigation rate in millimeters (e.g., output = 5 mm).

Classification models decide the category to which a given situation belongs, for example, whether irrigation is required. The classification can be binary (e.g., 0 = do not irrigate, 1 = irrigate) or multiclass (e.g., 0–5 mm = no irrigation, 6–15 mm = moderate irrigation, >15 mm = intensive irrigation).

Transforming the regression problem into a classification problem can simplify decision-making, for example, by setting a threshold of 10 mm below which there is no irrigation and above which there is. The advantages of classification are a wider choice of available algorithms and sometimes better interpretability.

The most commonly used approach is multiple linear regression, if the existence of a linear relationship between the target variable and input attributes is confirmed. The accuracy of the model is most often assessed using the coefficient of determination R^2, which indicates the proportion of variability in the target variable that the model explains (ideally, 1.00).

The model can be optimized using regularization (e.g., Lasso or Ridge), which penalizes less significant variables and reduces the risk of overfitting. The quality of the model can also be visually verified by comparing the predicted and actual values in a dot plot and calculating the average absolute error (e.g., in millimeter irrigation).

A wide range of models is available for classification, from simple (logistic regression, decision trees, and random forests) to complex deep learning architectures. Some models such as support vector machines (SVMs) can be used for both regression and classification.

An important part of classification modelling is hyperparametric tuning, in which different model configurations are tested. Because it is not possible to explore all combinations, a strategy of selecting from a limited set (e.g., grid search or random search) was used.

The models are evaluated using metrics such as:

- accuracy – the proportion of correctly classified cases,
- Precision (precision) and sensitivity (recall) – important in the case of non-uniform classes
- F1-score – harmonic mean of precision and sensitivity.

In the context of irrigation, it is important to minimize false-negative predictions (e.g., the model does not recommend irrigation when needed), which can negatively affect crop yields.

Time-series models can also be used if irrigation is analyzed over time. The best-known are:

- ARIMA – suitable for stationary time series with a relatively stable mean,
- Long Short-Term Memory (LSTM) is a type of recurrent neural network that is suitable for modelling long-term dependencies and nonlinear trends.

These models allow the prediction of future irrigation values based on historical trends and the evolution of environmental factors.

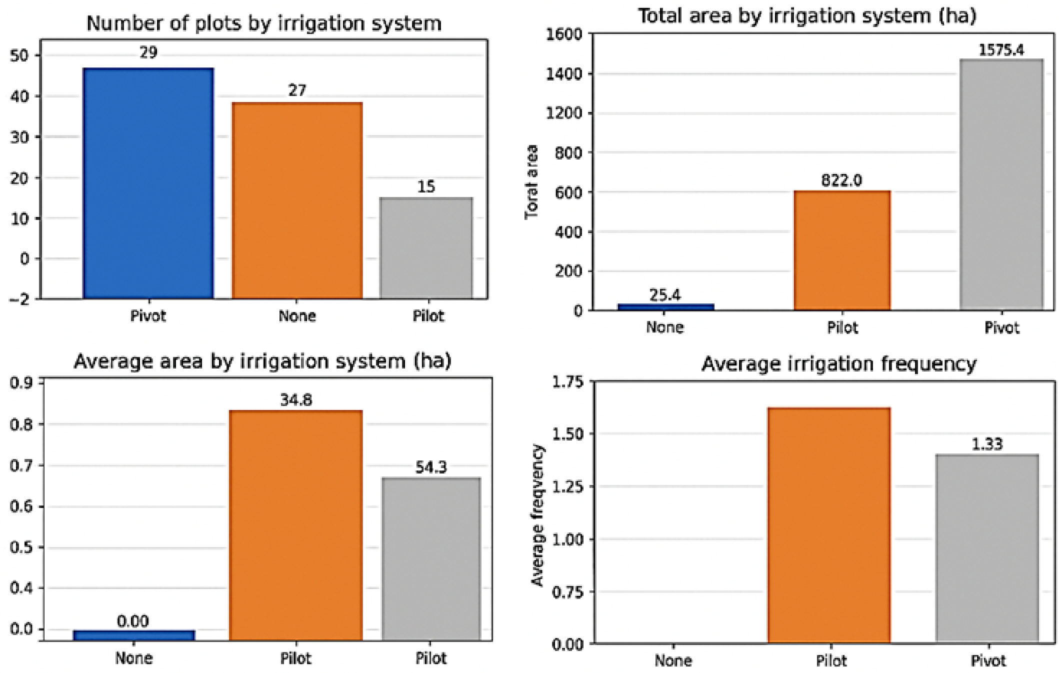
Based on the proposed dataset structure and type of target variable, regression models are more suitable for the accurate prediction of irrigation rates in millimeters. Nevertheless, the problem can be reformulated as a classification problem by categorizing the doses into discrete intervals, according to agronomic recommendations. The choice between regression and classification depends on the model output, data availability, and interpretation requirements.

Efficiency analysis of irrigation systems

Analysis of the spatial distribution and technical parameters of irrigation systems in the study area revealed significant heterogeneity in the implementation of hydromeliorative technologies (Figure 1). Two dominant types of irrigation systems were identified within the set of 71 agricultural plots, pivot and pilot irrigation systems, while 27 plots were managed without the application of artificial irrigation. Pivot systems were installed on 29 plots (40.8 %) with an average irrigated area of 56.17 ha and an average irrigation frequency of 1.38 applications over the study period. Pilot irrigation systems were implemented in 15 plots (21.1 %), with an average irrigated area of 54.16 ha and a higher average irrigation frequency of 2.00 applications per period. Plots without irrigation systems (38.0%) show a significantly lower average area of 0.96 ha, indicating preferential implementation of irrigation technologies on larger agricultural plots.

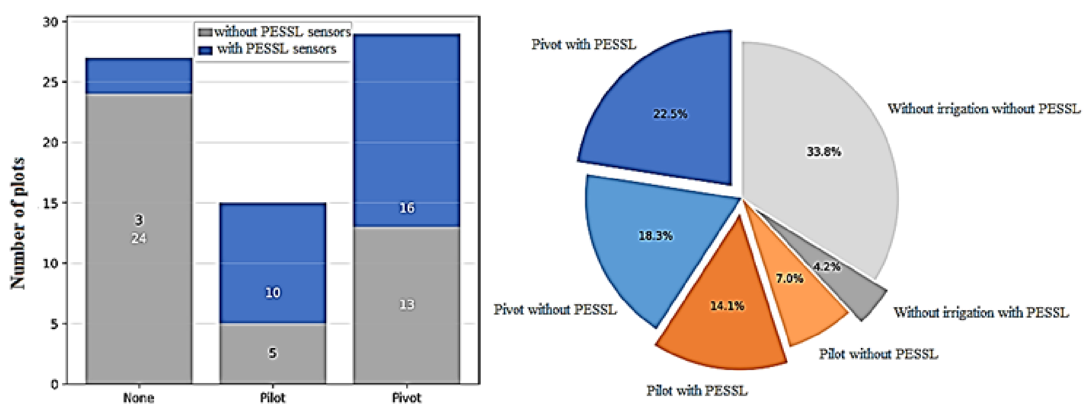
A comparative analysis of the technical parameters of the identified irrigation systems revealed distinctive characteristics of the individual typological categories. Pivot systems, characterized by the rotational movement of the irrigation boom around a central point, show a lower frequency of irrigation water application but cover a slightly larger average area than pile systems. The total area irrigated by pivot systems was 1,575,40 ha, while pilot systems cover 822,05 ha. A significant difference was identified in the average irrigation efficiency, where pivot systems reached a value of 79.64 %, whereas pilot systems showed an efficiency of 75.10 %. This discrepancy can be attributed to the different technical parameters of irrigation water distribution and the different algorithms used to control the irrigation process.

An important aspect of hydraulic improvement management in the analyzed area was the implementation of the PESSL sensor system for monitoring the soil moisture and meteorological parameters (Figure 2).



Source: Own processing

Figure 1: Distribution of irrigation systems by number of plots, total area, average area and frequency of irrigation.



Source: Own processing

Figure 2: Distribution of PESSL sensors by the type of the irrigation system.

A total of 29 plots (40.8%) were equipped with these sensors, with an uneven distribution between the different types of irrigation systems. Of the 29 plots with pivot irrigation, 16 (55.2%) were equipped with PESSL sensors, whereas 10 of the 15 plots (66.7%) in the pilot system category were equipped with sensor technology. A surprising finding was the presence of PESSL sensors in three plots without irrigation systems (11.1%), indicating the potential use of sensor data for optimizing further agronomic operations or planning future implementation of irrigation technologies (Table 1).

Analysis of irrigation efficiency, defined as the ratio between the optimum and applied irrigation water quantity, revealed an average value of 75.53%, with significant variability among plots (Table 2).

A comparative analysis of the recommended irrigation (average 129.55 mm) and the actual amount of water applied (average 95.88 mm) indicated a systematic tendency towards under-application of irrigation water, which may be due to economic constraints, technical limitations of irrigation systems, or a conservative approach to water resource management. A significant factor influencing irrigation efficiency was the distribution of natural rainfall, which averaged 186.50 mm over the study period. The correlation analysis between applied irrigation and efficiency showed a moderately strong positive correlation ($r = 0.7256$), indicating that an increase in the amount of irrigation water applied leads to higher water-use efficiency up to a certain saturation point (Figure 3).

Type of system	Number of plots	Average area (ha)	Average frequency	PESSL sensors
Pivot	29	54.32	1.33	16 (55.2 %)
Pilot	15	54.80	1.85	10 (66.7 %)
Without irrigation	27	0.94	0.00	3 (11.1 %)
Total	71	34.12	0.93	29 (40.8 %)

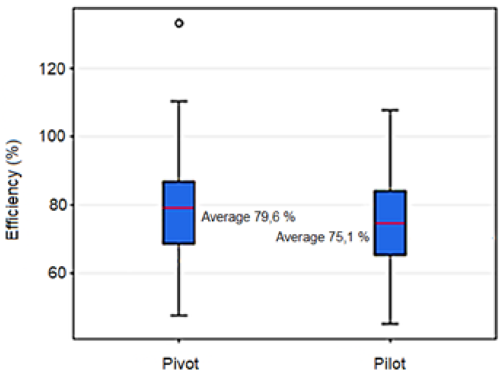
Source: Own processing

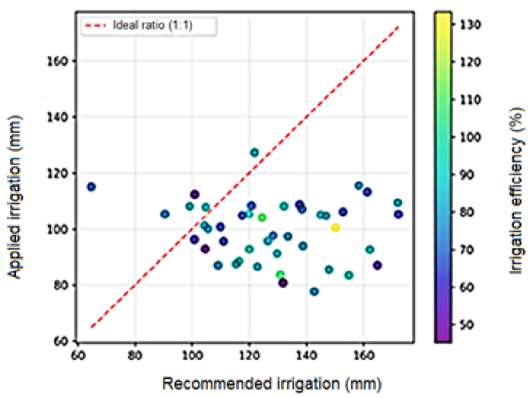
Table 1: Key parametres of irrigation systems.

Type of system	Average efficiency (%)	Recommended irrigation (mm)	Applied irrigation (mm)	Precipitation (mm)
Pivot	79.64	126.86	99.53	180.10
Pilot	75.10	131.00	99.38	201.06
Diameter (irrigated)	78.09	128.28	99.48	187.25

Source: Own processing

Table 2: Irrigation efficiency parametres.





Source: Own processing

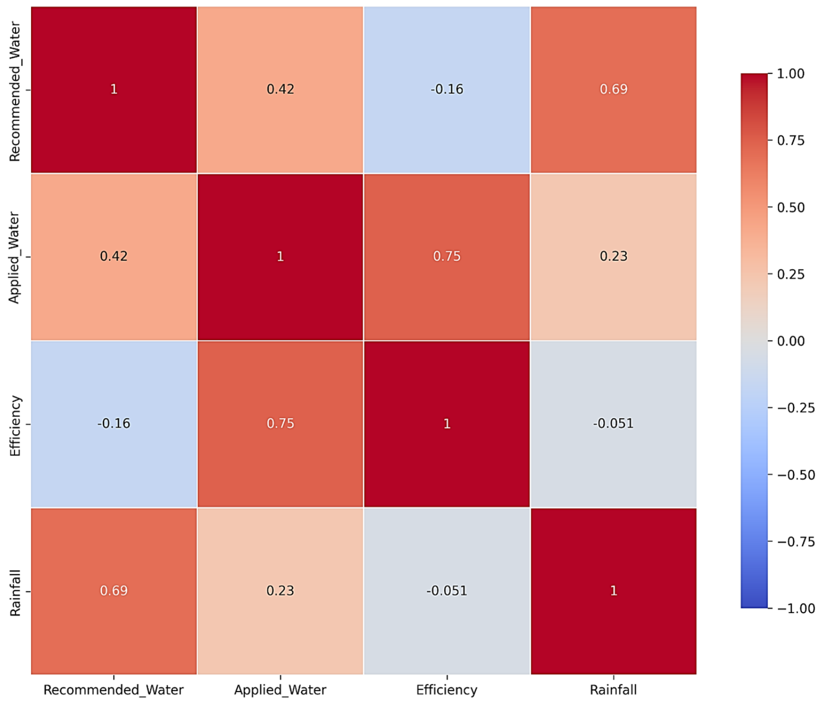
Figure 3: Irrigation efficiency analysis by the system type and relationship between recommended and applied irrigation.

Irrigation efficiency analysis

Quantitative analysis of the hydroremediation parameters based on data obtained from the PESSL sensor system revealed complex relationships between the different components of the irrigation process. The average recommended irrigation rate was 129.55 mm, while the average applied rate was only 95.88 mm, indicating a systematic tendency to under-application of irrigation water in the range of 25.99%. The average irrigation efficiency, defined as the ratio of the optimum use of applied water to the theoretically required amount, was 75.53%. Atmospheric precipitation was a significant factor influencing the irrigation regime, with an average value of 186.50 mm over the study period, representing the dominant component of the total water input to the soil profile.

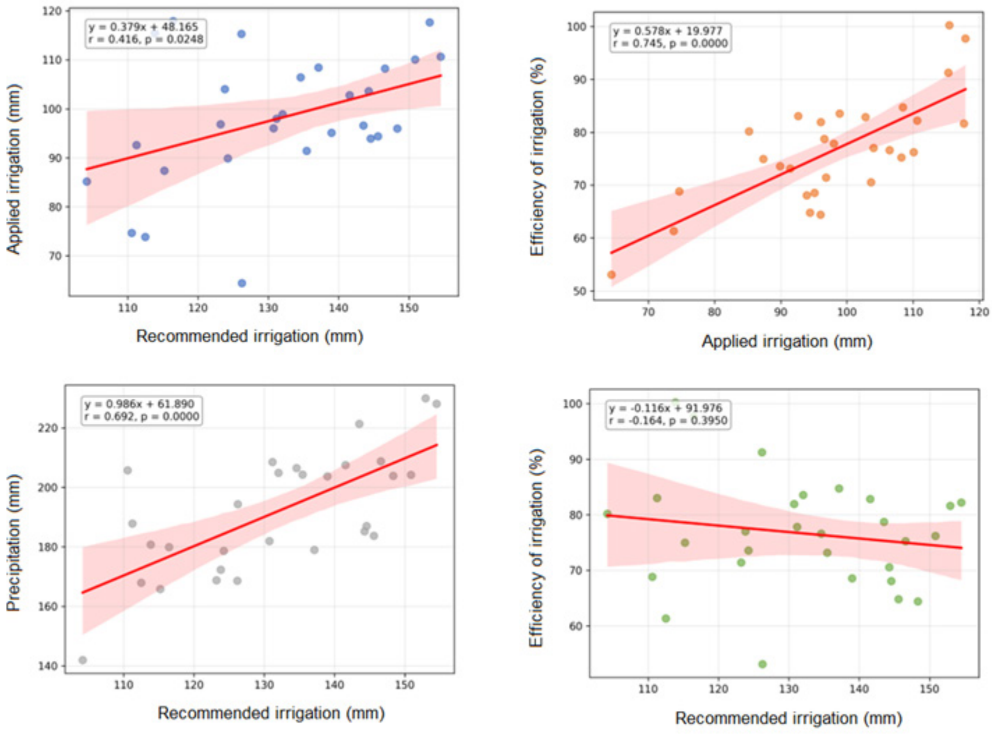
Correlation analysis of the hydromeliorative parameters (Figure 4) revealed several statistically significant relationships. The strongest positive correlation ($r = 0.8583$, $p < 0.001$) was detected between the recommended irrigation and rainfall, reflecting the adaptive nature of the irrigation algorithm that integrates the prediction of rainfall into the calculation of the optimal irrigation rate. The moderately strong positive correlation ($r = 0.7256$, $p < 0.001$) between applied irrigation and efficiency indicates that an increase in the amount of irrigation water applied leads to a proportional increase in water use efficiency, probably because of the surpassing of the soil moisture thresholds necessary for optimal water uptake by the plant root system. The weak negative correlation ($r = -0.1027$, $p = 0.5964$) between recommended irrigation and efficiency suggests a potential overestimation of optimal irrigation rates by the PESSL algorithm, especially under conditions of higher soil moisture. The medium positive correlation ($r = 0.5149$, $p = 0.0042$) between the recommended and applied irrigations reflects the partial implementation of the sensor system recommendations in practical irrigation management.

Regression analysis of the relationship between applied irrigation and efficiency generated a prediction model (Figure 5) characterized by the equation Efficiency = 0.483 × Applied + 29.27 ($R^2 = 0.527$), which explains more than half of the variability in irrigation efficiency. This model implies that each 10 mm increase in applied irrigation rate leads to an average increase in efficiency of 4.83 percentage points. extrapolation of the model indicated that a theoretical achievement of 100% efficiency would require



Source: Own processing

Figure 4: Correlation matrix of irrigation parametres showing the strength and direction of the relationships between the variables.

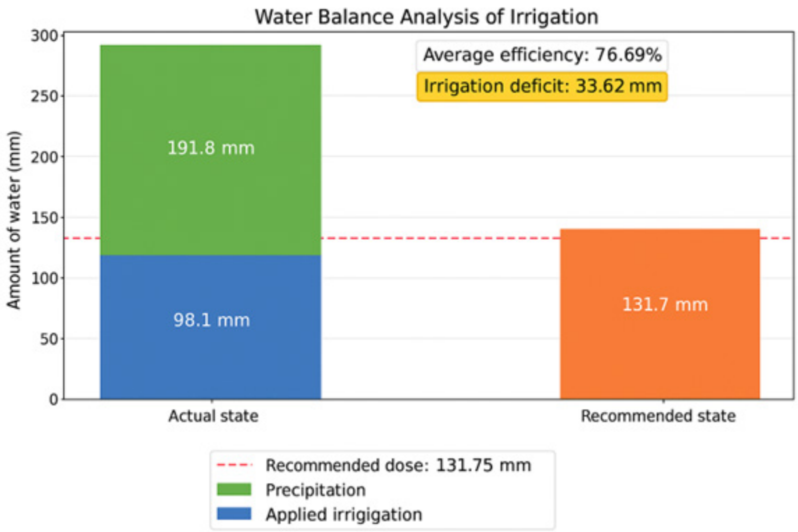


Source: Own processing

Figure 5: Regression analysis of key relationships between irrigation parametres with regression equations and correlation coefficients plotted.

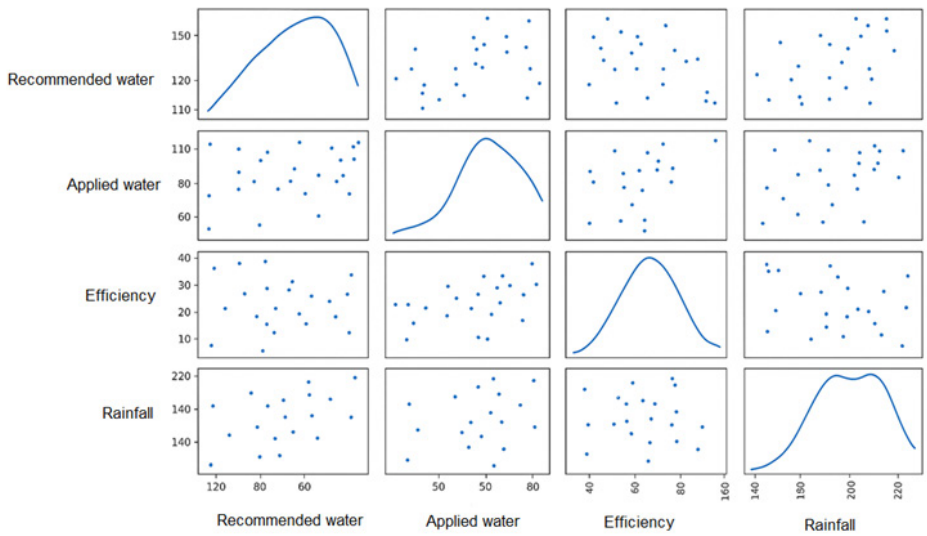
the application of approximately 146.4 mm of irrigation water, which exceeds the average recommended application rate by 13.0%. This discrepancy may be the result of the conservative settings of the PESSL algorithm or specific soil and climate conditions that modify the optimal irrigation parameters. The regression model of the relationship between recommended and applied irrigation (Applied = 0.385 × Recommended + 46.01, R^2 = 0.265) further confirmed a systematic tendency towards underapplication of irrigation water, with the rate of implementation of recommendations decreasing with increasing recommended rate.

The Water balance analysis (Figure 6) revealed that the combination of applied irrigation (95.88 mm) and rainfall (186.50 mm) provided a total water supply of 282.38 mm, which exceeded the average recommended rate (129.55 mm) by 118.0%. However, this apparent surplus must be interpreted in the context of the temporal distribution of rainfall, its infiltration efficiency, and evapotranspiration losses, which significantly reduces the amount of water available to plants. The irrigation deficit, defined as the difference between the recommended rate and the applied rates, was 33,67 mm, which represents 26,0 % of the recommended rate. This deficit is partly compensated by rainfall, but its temporal and spatial variability limits the reliability of this compensation. The implementation of precise irrigation strategies based on continuous monitoring of soil moisture and meteorological parameters has the potential to optimize the efficiency of water use and minimize the irrigation deficit while maintaining the economic efficiency of hydraulic improvement measures. (Tables 3, 4, 5).



Source: Own processing

Figure 6: Regression analysis of key relationships between irrigation parametres with regression equations and correlation coefficients plotted.



Source: Own processing

Figure 7: Dot-plot matrix showing the interrelationships between all the analysed irrigation parametres.

Parameter	Value
Average recommended irrigation rate	131.75 mm
Average applied irrigation rate	98.13 mm
Average irrigation efficiency	76.69%
Average rainfall	191.76 mm
Irrigation deficit	33.62 mm
Total water supply (irrigation + precipitation)	289.89 mm

Source: Own processing

Table 3: Key irrigation parametres and correlations.

Relationship	Correlation coefficient (r)	p-value
Recommended irrigation vs. Precipitation	0.6921	0.0000
Applied irrigation vs. Efficiency	0.7454	0.0000
Recommended irrigation vs. Efficiency	-0.1641	0.3950
Recommended vs. Applied irrigation	0.4159	0.0248
Applied irrigation vs. Precipitation	0.2334	0.2230
Efficiency vs. Collisions	-0.0507	0.7940

Source: Own processing

Table 4: Correlation coefficients between irrigation parametres.

Model	Regression equation	Coefficient of determination (R^2)
Applied irrigation vs. Efficiency	Efficiency = 0.578 × Applied + 19.98	0.556
Recommended vs. Applied irrigation	Applied = 0.379 × Recommended + 48.16	0.173
Recommended irrigation vs. Efficiency	Efficiency = -0.116 × Recommended + 91.98	0.027
Precipitation vs. Recommended irrigation	Recommended = 0.486 × Precipitation + 38.56	0.479

Source: Own processing

Table 5: Regression models for irrigation parametres.

Irrigation efficiency prediction models

For the analysis, four different regression models (Table 6) were developed based on three key variables: the recommended irrigation amount, actual water applied, and natural precipitation. Data were collected using PESSL sensors placed on the selected agricultural plots. The irrigation efficiency was calculated as the ratio between the optimum water use and the actual amount applied and expressed as a percentage.

Analysis of the different prediction models revealed significant differences in their ability to explain variability in irrigation system efficiency. The model based on the recommended irrigation amounts alone (Model 1) showed a low coefficient of determination (R^2 = 0.0918), suggesting that the recommended values alone are not sufficient predictors of actual efficiency. The negative coefficient in equation (-0.2606) suggests an inverse relationship, where higher recommended water amounts are paradoxically associated with lower efficiency, which may be due to a tendency to overirrigate at higher recommended values.

The model based on the actual amount of water applied (Model 2) performed significantly better with R^2 = 0.4121. A positive coefficient (0.4896) indicates that irrigation efficiency increases as the amount of water applied increases, which may reflect the fact that farmers adapt the amount of water to actual conditions better than the automated recommended values.

Surprisingly, the model based on natural precipitation alone (Model 3) showed the lowest predictive ability (R^2 = 0.0114), suggesting that precipitation alone had minimal direct influence on the efficiency of irrigation systems. This result may be due to the fact that farmers already take rainfall into account when making irrigation decisions, thus offsetting its influence.

The multiple regression model (Model 4), which combined all three variables, achieved a significantly higher predictive ability, with R^2 = 0.8770. This model explains approximately 87.7% of the variability in irrigation efficiency, a significant improvement over the simple models. The coefficients in Eq. indicate that, while the recommended amount of water has a negative effect on efficiency (-0.8386), the amount actually applied has a positive effect (0.7266). Interestingly, in the context of multiple regression analysis, rainfall had a slightly positive effect (0.1864), suggesting a complex interaction between these variables.

The results of this study have several important implications for optimizing agricultural irrigation practices. First, the significant difference between the predictive abilities of the simple and multiple

Model	Equations	R^2
Model 1: Recommended irrigation	Efficiency = -0.2606 × Recommended + 115.8477	0.0918
Model 2: Applied irrigation	Efficiency = 0.4896 × Applied + 30.7128	0.4121
Model 3: Collisions	Efficiency = -0.0714 × Precipitation + 94.6646	0.0114
Model 4: Multiple regression	Efficiency = 82.6585 - 0.8386 × Recommended + 0.7266 × Applied + 0.1864 × Precipitation	0.8770

Source: Own processing

Table 6: Regression models.

regression models highlights the complex nature of irrigation efficiency, which is simultaneously affected by the interaction of several factors.

Second, the negative relationship between the recommended water quantity and efficiency suggests the potential shortcomings of the current irrigation recommendation algorithms. These algorithms may be too conservative and lead to overirrigation, which reduces the overall efficiency of the water use.

Third, the strong positive relationship between applied water quantity and efficiency highlights the importance of human factors and farmers' experience in irrigation decisions. This result suggests that a combination of automated systems and expert judgment can lead to a more optimal use of water resources.

The multiple regression model is a robust tool for predicting irrigation efficiency and can serve as a basis for the development of advanced decision support systems in precision agriculture. The implementation of this model can lead to significant water savings and improve the sustainability of agricultural practices.

This analysis showed that the efficiency of irrigation systems is a complex phenomenon that cannot be adequately modelled by a single variable. A multiple regression model that takes into account the recommended amount of irrigation, the actual amount of water applied, and natural rainfall provided the best predictive power and explained approximately 87.7% of the variability in irrigation efficiency.

The results highlight the need for a holistic approach to optimizing irrigation practices that considers both technological aspects (sensors, recommendation algorithms) and human factors (farmers' experience and decision-making processes). Future research should focus on integrating other factors such as soil properties, climatic conditions, and crop specificities into prediction models to further increase their accuracy and applicability.

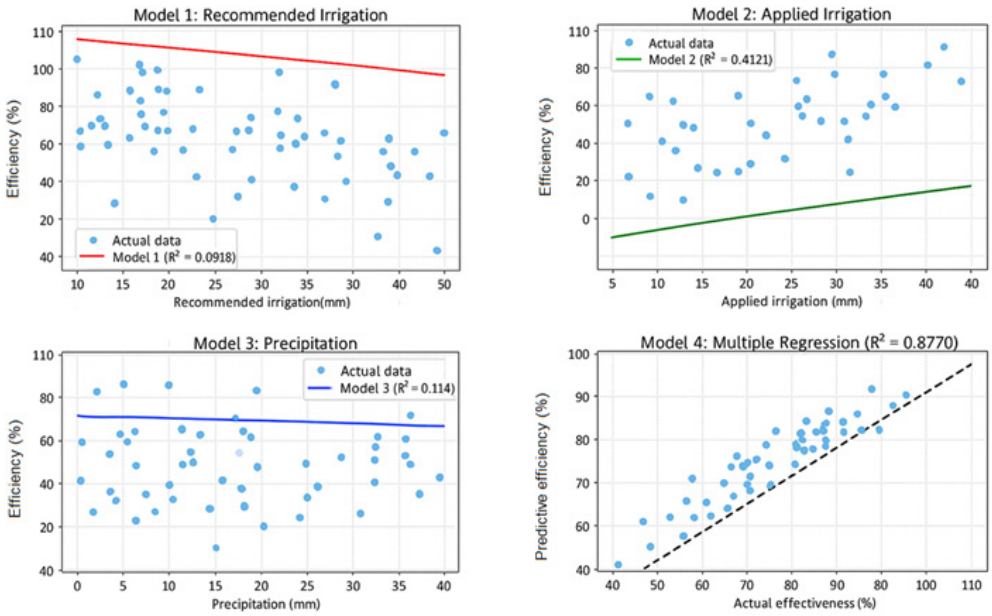
The graphs (Figure 8) show the relationship between each variable and irrigation efficiency for each model. The relationship between the actual and predicted efficiencies in Model 4 (multiple regression) is shown.

The graph (Figure 9) shows the coefficients of determination (R^2) for each model, indicating their ability to explain variability in irrigation efficiency.

A multiple regression model (Model 4) was developed based on three key predictors: recommended irrigation amount, actual water applied, and natural precipitation. This model is defined by the following equation.

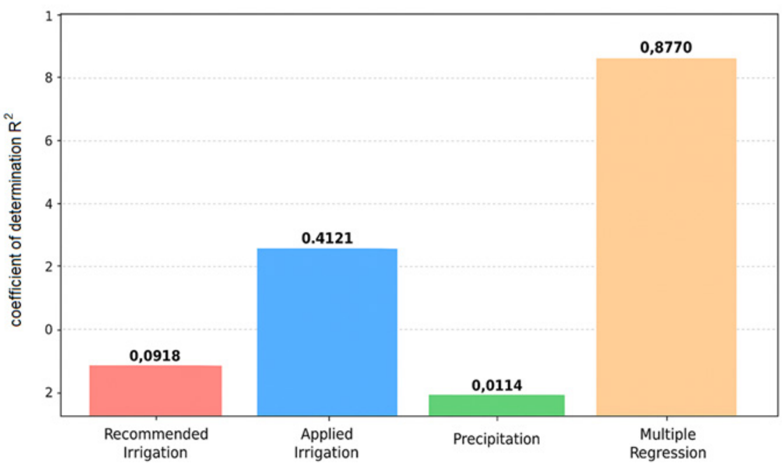
$$Efficiency = 82.6585 - 0.8386 \times Recommended + 0.7266 \times Applied + 0.1864 \times Precipitation.$$

This model achieves a coefficient of determination $R^2 = 0.8770$, which means that it explains approximately 87.70% of the variability in irrigation efficiency. Compared to simpler models based on single predictors, the multiple regression model represents a significant improvement in predictive ability and provides a more comprehensive view of the factors influencing irrigation system efficiency.



Source: Own processing

Figure 8: Comparison of irrigation efficiency prediction models.



Source: Own processing

Figure 9: Comparison of irrigation efficiency prediction models.

Conclusion

The coefficients of the multiple regression model provided valuable information on the relative influence of individual predictors on irrigation efficiency. Each coefficient represents the change in predicted efficiency when the relevant variable is increased by one unit (1 mm), assuming that the other variables remain constant. This ceteris paribus interpretation allows the effects of each factor to be separately isolated and quantified.

The coefficient for recommended irrigation (-0.8386) shows a negative value, indicating an inverse relationship between the recommended water quantity and irrigation efficiency. Specifically, each 1 mm increase in the recommended amount is associated with a decrease in the predicted efficiency of approximately 0.84 percentage points, holding other variables constant. This negative relationship can be interpreted in several ways. A plausible explanation is that irrigation recommendation algorithms may systematically overestimate the amount of water required, leading to inefficient use of water resources. Alternatively, this relationship may reflect the fact that higher recommended values are generated in more challenging environments, where high irrigation efficiency is more difficult to achieve.

The coefficient for applied irrigation (+0.7266) is positive, indicating that a 1 mm increase in the actual amount of water applied is associated with an increase in the predicted efficiency of approximately 0.73 percentage points, holding other variables constant. This positive relationship may reflect the adaptive behavior of farmers who adjust the amount of water applied to actual conditions based on their experience and observations. The results suggest that farmers' decisions regarding the amount of water applied are often more accurate than automated recommendations, highlighting the importance of human factors in the irrigation process.

The coefficient for precipitation (+0.1864) is slightly positive, indicating that a 1 mm increase in natural precipitation is associated with an increase in the predicted efficiency of approximately 0.19 percentage points, holding other variables constant. This relationship is interesting because in the simple regression model (Model 3), precipitation showed a negative coefficient and very low predictive power ($R^2 = 0.0114$). However, in the context of multiple regression, rainfall positively contributed to efficiency, suggesting a complex interaction between rainfall and other factors. The positive effect of rainfall may be explained by the fact that natural rainfall provides a more optimal distribution of water in the soil than artificial irrigation or by the fact that farmers integrate rainfall information more efficiently into their irrigation decisions.

The model constant (82.6585) represents the predicted irrigation efficiency when all predictors are zero. In a practical context, this value is of limited interpretative value because the null values of the predictors lie outside the range of observed data. However, this constant contributes to the overall accuracy of the model and ensures that predictions are calibrated to the correct level of efficiency.

The multiple regression model provides several important practical implications for optimizing irrigation practices in precision agriculture. First, the negative coefficient for recommended irrigation suggests the need to revise the current algorithms to generate recommendations. These algorithms should be calibrated to better reflect the actual needs of crops and minimize the risk of overirrigation. The implementation of adaptive algorithms that learn from historical data and farmer experience could lead to more accurate recommendations and greater water-use efficiency.

Second, the positive coefficient for applied irrigation highlights the value of farmers' experience and knowledge in making irrigation decisions. Decision support systems should integrate both automated recommendations and manual adjustments based on expert judgement. A hybrid approach that combines technological solutions with human factors can lead to more optimal use of water resources and higher irrigation efficiency.

Third, the positive coefficient for rainfall in the context of multiple regression suggests that effective integration of rainfall data into the decision-making process can improve irrigation efficiency. Advanced rainfall monitoring and weather prediction systems should be part of comprehensive irrigation management solutions. These systems should provide not only historical data, but also short- and medium-term forecasts to enable farmers to plan irrigation activities better.

The multiple regression model can be implemented as part of an advanced decision-support system for precision agriculture. These systems should integrate data from various sources, including soil moisture sensors, weather stations, and satellite imagery, to provide farmers with comprehensive information for optimizing irrigation practices. The implementation of such systems can lead to significant water savings, reduced irrigation costs, and increased sustainability in agricultural practices.

Corresponding author:
prof. Ing. Zuzana Palková, PhD.
Institute of Electrical Engineering, Automation, Informatics and Physics, Faculty of Engineering
Slovak University of Agriculture in Nitra, 949 76 Nitra, Slovakia
Department of Mechanical Engineering, Faculty of Technology,
Institute of Technology and Business in České Budějovice, 370 01 České Budějovice, Czech Republic
New Edu, n.o., Kováčikova 12 949 01 Nitra, Slovakia
Phone: + 421 37 6414 765, E-mail: zuzana.palkova@uniag.sk

References

[1] Abdelmoneim, A. A., Kimaita, H. N., Al Kalaany, C. M., Derardja, B., Dragonetti, G. and Khadra, R. (2025) "IoT Sensing for Advanced Irrigation Management: A Systematic Review of Trends, Challenges, and Future Prospects, *Sensors*, Vol. 25, No. 7, p. 2291. ISSN 1424-8220. DOI 10.3390/S25072291/S1.

[2] Abioye, E. A., Abidin, M. S. Z., Mahmud, M. S. A., Buyamin, S., Ishak, M. H. I., Rahman, M. K. I. A., Otuoze, A. O., Onotu, P. and Ramli, M. S. A. (2020) "A review on monitoring and advanced control strategies for precision irrigation", *Computers and Electronics in Agriculture*, Vol. 173, p. 105441. E-ISSN 1872-7107. DOI 10.1016/J.COMPAG.2020.105441.

[3] Abioye, E. A., Hensel, O., Esau, T. J., Elijah, O., Abidin, M. S. Z., Ayobami, A. S., Yerima, O. and Nasirahmadi, A. (2022) "Precision Irrigation Management Using Machine Learning and Digital Farming Solutions", *AgriEngineering*, Vol. 4, No. 1. pp.70-103. E-ISSN 2624-7402. DOI 10.3390/AGRIENGINEERING4010006.

[4] Ahmed, A. A., Sayed, S., Abdoulhalik, A., Moutari, S. and Oyedele, L. (2024) "Applications of machine learning to water resources management: A review of present status and future opportunities", *Journal of Cleaner Production*, Vol. 441, p. 140715. E-ISSN 1879-1786. DOI 10.1016/J.JCLEPRO.2024.140715.

[5] Akkem, Y., Biswas, S. K. and Varanasi, A. (2023) "Smart farming using artificial intelligence: A review", *Engineering Applications of Artificial Intelligence*, Vol. 120. E-ISSN 1873-6769, ISSN 0952-1976. DOI 10.1016/J.ENGAPPAI.2023.105899.

[6] Allen, R., Pereira, L., Raes, D. and Smith, M. (2006) "Parte C. Evapotranspiración del cultivo en condiciones no estándar ET c bajo condiciones de estrés hídrico. Evapotranspiración Del Cultivo Guías Para La Determinación de Los Requerimientos de Agua de Los Cultivos. *ESTUDIO FAO RIEGO Y DRENAJE,* Vol. 56., 48 p. ISSN 0254-5293. (In Spanish).

[7] Arulraj, B. and Karthikeyan, N. (2024) "Machine learning approaches for irrigation scheduling: A comprehensive review", *Computers and Electronics in Agriculture*, Vol. 2018. ISSN 0168-1699.

[8] Bastola, S., Murphy, C. and Sweeney, J. (2011) "The role of hydrological modelling uncertainties in climate change impact assessments of Irish river catchments", *Advances in Water Resources*, Vol. 34, No. 5, pp. 562-576. E-ISSN 1872-9657. DOI 10.1016/J.ADVWATRES.2011.01.008.

[9] Bellingham, K., Thompson, J. and Rodriguez, A. (2023) "Deep learning approaches for soil moisture prediction in precision irrigation", *Computers and Electronics in Agriculture*, Vol. 208. ISSN 0168-1699.

[10] Bhat, S. A. and Huang, N. F. (2021) "Big Data and AI Revolution in Precision Agriculture: Survey and Challenges", *IEEE Access*, Vol. 9, pp. 110209-110222. ISSN 2169-3536. DOI 10.1109/ACCESS.2021.3102227.

[11] Bwambale, E., Abagale, F. K. and Anornu, G. K. (2022) "Smart irrigation monitoring and control strategies for improving water use efficiency in precision agriculture: A review", *Agricultural Water Management*, Vol. 260. ISSN 1873-2283. DOI 10.1016/j.agwat.2021.107324.

[12] Crystal-Ornelas, R., Varadharajan, C., O'Ryan, D., Beilsmith, K., Bond-Lamberty, B., Boye, K., Burrus, M., Cholia, S., Christianson, D. S., Crow, M., Damerow, J., Ely, K. S., Goldman, A. E., Heinz, S. L., Hendrix, V. C., Kakalia, Z., Mathes, K., O'Brien, F., Pennington, S. C., Robles, E., Rogers, A., Simmonds, M., Velliquette, T., Weisenhorn, P., Welch, J. N., Whitenack, K. and Agarwal, D. A. (2022) "Enabling FAIR data in Earth and environmental science with community-centric (meta)data reporting formats", *Scientific Data*, Vol. 9, No. 1, p. 700. ISSN 2052-4463. DOI 10.1038/s41597-022-01606-w.

[13] Dong, Y., Werling, B., Cao, Z. and Li, G. (2024) "Implementation of an in-field IoT system for precision irrigation management", *Frontiers in Water*, Vol. 6, p. 1353597. ISSN 2624-9375. DOI 10.3389/FRWA.2024.1353597/BIBTEX.

[14] Gao, H., Liu, J., Wang, H., Mei, C. and Wang, J. (2024) "Estimation of irrigated crop artificial irrigation evapotranspiration in China", *Scientific Reports*, Vol. 14, Vol. 1, p. 16142. ISSN 2045-2322. DOI 10.1038/s41598-024-67042-5.

[15] García, L., Parra, L., Jimenez, J. M., Lloret, J. and Lorenz, P. (2020) "IoT-Based Smart Irrigation Systems: An Overview on the Recent Trends on Sensors and IoT Systems for Irrigation in Precision Agriculture", *Sensors*, Vol. 20, No. 4, p. 1042. ISSN 1424-8220. DOI 10.3390/S20041042.

[16] Hamami, L. and Nassereddine, B. (2020) "Application of wireless sensor networks in the field of irrigation: A review", *Computers and Electronics in Agriculture*, Vol. 179, p. 105782. ISSN 0168-1699. DOI 10.1016/j.compag.2020.105782.

[17] Hossain, M. S. (2023) "Advanced machine learning techniques for irrigation optimisation in canal systems", *Water Resources Management*, Vol. 37, pp. 3241-3258. ISSN 1573-1650.

[18] Jensen, M. E., Burman, R. D. and Allen, R. G. (eds.) (1990) "*Evapotranspiration and Irrigation Water Requirements: A Manual*", ASCE Manuals and Reports on Engineering Practice No. 70. ASCE Manuals and Reports on Engineering Practice No. 70, The Society, 332 p. ISBN 9780872627635.

[19] Jihani, N., Kabbaj, M. N., Benbrahim, M. and Jerbi, M. (2024) "A systematic review on smart irrigation management systems using machine learning: Current trends and future per-spectives", *Information Processing in Agriculture*, Vol. 11, pp. 306-325. E-ISSN 2214-3173.

[20] Jones H. G. (2024) "Advances in precision irrigation: integrating AI and remote sensing", *Agriculture*, Vol. 14, No. 2. ISSN 2077-0472.

[21] Karray, F., Jmal, M. W., Garcia-Rodriguez, J., Aloui, Z. and Obaid, A. J. (2020) "A comprehensive survey on wireless sensor networks", *Future generation computer systems*, Vol. 109, pp. 1-22. E-ISSN 1872-7115, ISSN 0167-739X. DOI 10.1016/j.comnet.2018.05.010.

[22] Kaur, A., Bhatt, D. P. and Raja, L. (2024) "Developing a Hybrid Irrigation System for Smart Agriculture Using IoT Sensors and Machine Learning in Sri Ganganagar, Rajasthan", *Journal of Sensors*, Vol. 1, p. 6676907. E-ISSN 1687-7268, ISSN 1687-725X. DOI 10.1155/2024/6676907.

[23] Kersebaum, K. C., Hecker, J.-M., Mirschel, W. and Wegehenkel, M. (2007) "Modelling water and nutrient dynamics in soil–crop systems: a comparison of simulation models applied on common data sets", In *Modelling Water and Nutrient Dynamics in Soil–Crop Systems*, Springer, Dordrecht, pp. 1-17. ISBN 978-1-4020-4478-6. DOI 10.1007/978-1-4020-4479-3_1.

[24] Khan, S., Tariq, R., Yuanlai, C. and Blackwell, J. (2006) "Can irrigation be sustainable?", *Agricultural Water Management*, Vol. 80, No. 1-3 Sp. Iss., pp. 87-99. ISSN 1873-2283. DOI 10.1016/J.AGWAT.2005.07.006

[25] Kumar Kasera, R., Gour, S. and Acharjee, T. (2024) "A comprehensive survey on IoT and AI based applications in different pre-harvest, during-harvest and post-harvest activities of smart agriculture", *Computers and Electronics in Agriculture*, Vol. 216, p. 108522. ISSN 0168-1699. DOI 10.1016/J.COMPAG.2023.108522.

[26] Moriasi, D. N., Arnold, J. G., Liew, M. W., Van, Bingner, R. L., Harmel, R. D. and Veith, T. L. (2007) "Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations", *Transactions of the ASABE*, Vol. 50, No. 3, pp. 885-900. E-ISSN 2151-0040. DOI 10.13031/2013.23153.

[27] Palková, Z. and Rodny, M. (2018) "Stochastic modelling of irrigation processes", *Water Resources Management*, pp. 2849-2863. E-ISSN 1573-1650.

[28] Prasath, B. and Akila, M. (2023) "IoT-based pest detection and classification using deep features with enhanced deep learning strategies", *Engineering Applications of Artificial Intelligence,* Vol. 121, p. 105985. ISSN 1873-6769. DOI 10.1016/J.ENGAPPAI.2023.105985.

[29] Raouhi, E. M., Zouizza, M., Lachgar, M., Zouani, Y., Hrimech, H. and Kartit, A. (2023) "AIDSII: An AI-based digital system for intelligent irrigation", *Software Impacts*, Vol. 17, p. 100574. ISSN 2665-9638. DOI 10.1016/J.SIMPA.2023.100574.

[30] Sharma, A., Jain, A., Gupta, P. and Chowdary, V. (2021) "Machine Learning Applications for Precision Agriculture: A Comprehensive Review", *IEEE Access*, Vol. 9, pp. 4843–4873. ISSN 2169-3536. DOI 10.1109/ACCESS.2020.3048415.

[31] Sithartan, R., Rajesh, M., Shanmuganathan, V., Eswaran, V., Eswaran, S. K., Yuvaraj, S., Kumar, A., Raglend, J. and Vengatesan, K. (2023) "A novel autonomous irrigation system for smart agriculture using AI and 6G enabled IoT network", *Microprocessors and Microsystems*, Vol. 101, pp. 104905. ISSN 0141-9331. DOI 10.1016/J.MICPRO.2023.104905.

[32] FAO (2021) "*The State of the World's Land and Water Resources for Food and Agriculture – Systems at breaking point (SOLAW 2021)*", Synthesis report 2021, book, 82 p. ISBN 978-92-5-135327-1. DOI 10.4060/CB7654EN.

[33] Umutoni, L. and Samadi, V. (2024) "Application of machine learning approaches in supporting irrigation decision making: A review", *Agricultural Water Management*, Vol. 294, p. 108710. ISSN 1873-2283. DOI 10.1016/J.AGWAT.2024.108710.

[34] Vallejo-Gómez, D., Osorio, M. and Hincapié, C. A. (2023) "Smart Irrigation Systems in Agriculture: A Systematic Review", *Agronomy*, Vol. 13, No. 2, 342 p. ISSN 2073-4395. DOI 10.3390/AGRONOMY13020342.

[35] Wakchaure, M., Patle, B. K. and Mahindrakar, A. K. (2023) "Application of AI techniques and robotics in agriculture: A review", *Artificial Intelligence in the Life Sciences*, Vol. 3, p. 100057. E-ISSN 2667-3185. DOI 10.1016/J.AILSCI.2023.100057.

[36] Wei, H., Xu, W., Kang, B., Eisner, R., Muleke, A., Rodriguez, D., de Voil, R., Sadras, V., Monjardino, M. and Harrison, M. T. (2024) "Irrigation with Artificial Intelligence: Problems, Premises, Promises", *Human-Centric Intelligent Systems*, Vol. 4, No. 2, pp. 187-205. E-2667-1336. DOI 10.1007/S44230-024-00072-4.

Agris on-line Papers in Economics and Informatics

Volume XVII Number 4, 2025

Enhancing Market Access for Smallholder Farmers in Indonesia: The Role of Managerial Capacity and Member Motivation in Collective Action within Farmer Groups

Rokhman Permadi, Lili Winarti

Department of Agribusiness, Faculty of Agriculture, Darwan Ali University, Central Kalimantan, Indonesia

Abstract

Smallholder farmers are pivotal to global food security yet encounter substantial obstacles in accessing competitive markets. The objective is to investigate the impact of managerial capacity and member motivation on collective action and market access among smallholder farmers in Indonesia. A survey was conducted with 249 kepok banana farmers belonging to farmer groups in Seruyan Regency, Central Kalimantan, Indonesia. Data were collected using a structured questionnaire that included demographic information and perceptions of managerial capacity, motivation, collective action, and market access. Partial Least Squares Structural Equation Modeling (PLS-SEM) was employed to assess the relationships among the constructs. The results indicate that managerial capacity significantly enhances the role of farmer groups ($\beta = 0.494$, $p < 0.001$), while member motivation does not show a significant effect ($\beta = 0.076$, $p = 0.290$). The role of farmer groups significantly influences both collective action ($\beta = 0.616$, $p < 0.001$) and market access ($\beta = 0.240$, $p < 0.001$). Furthermore, collective action has a significant positive effect on market access ($\beta = 0.479$, $p < 0.001$). These findings underscore the critical role of farmer groups in organizing collective strategies to enhance market access. Managerial capacity is pivotal for successful collective action. Policymakers should strengthen farmer group institutions to foster collective action, reduce market barriers, and achieve sustainable agricultural growth.

Keywords

Agricultural development, economic resilience, institutional capacity, PLS-SEM.

Permadi, R. and Winarti, L. (2025) "Enhancing Market Access for Smallholder Farmers in Indonesia: The Role of Managerial Capacity and Member Motivation in Collective Action within Farmer Groups", *AGRIS on-line Papers in Economics and Informatics*, Vol. 17, No. 4, pp. 101-115. ISSN 1804-1930. DOI 10.7160/aol.2025.1704x08.

Introduction

Smallholder farmers play a critical role in ensuring global food security, providing approximately 80% of the world's food supply (FAO and IFAD, 2019; Lowder et al., 2019). Their contributions are particularly significant in developing countries, where agriculture remains the backbone of rural economies and a major source of livelihood. Despite their essential role in food production, smallholder farmers face numerous challenges in accessing competitive markets. These challenges significantly limit their participation in and benefits from broader agricultural value chains, thereby hampering income generation and economic resilience.

A multitude of factors constrain smallholder farmers' access to competitive markets. Limited resources, including inadequate credit, high input costs, and insufficient market information, represent significant barriers (Courtois and Subervie, 2014; Kiveu and Ofafa, 2013; Magesa et al., 2014). These issues are further compounded by poor infrastructure and high transportation costs, which restrict farmers' ability to compete effectively and reduce profitability (Fan and Salas Garcia, 2018; Migose et al., 2018). Limited access to financial services hinders necessary investments in productivity and marketing strategies (Akpa et al., 2023), while geographic isolation diminishes market access and bargaining power (Untari and Vellema, 2022).

In Indonesia, many agricultural markets are characterized by oligopsony structures, where a few dominant buyers exert considerable influence over prices. Such market conditions often result in unfavorable terms for smallholder farmers. Collective marketing through farmer groups can mitigate this imbalance by enhancing negotiation power and improving economies of scale (Mangan and Ward, 2024; Markelova and Mwangi, 2010).

However, in the absence of such collective strategies, smallholder farmers are placed at a considerable disadvantage compared to larger agricultural enterprises, which naturally benefit from scale efficiencies and market access. The fragmented nature of smallholder farming leads to high transaction costs and low bargaining power for individual producers. Moreover, even motivated farmers remain constrained when marketing efforts are uncoordinated, thereby limiting the effectiveness of individual initiative (Rhoads and Shogren, 1999).

Agricultural institutions such as cooperatives and farmer groups have shown significant potential in supporting economic development and enhancing smallholder farmers' competitiveness (Otekunrin et al., 2019; Sitoe and Sitole, 2019). By leveraging collective action, these institutions help smallholders address challenges such as high transaction costs and limited bargaining power. Through pooling resources, coordinating production and marketing activities, and strengthening bargaining positions, collective action facilitates better access to market information and improves farmers' ability to secure higher prices for their products (Barrett, 2008; Markelova et al., 2009). Empirical evidence highlights that collective action enhances market access, thereby promoting economic resilience (Abdul-Rahaman and Abdulai, 2020; Aku et al., 2018).

Managerial capacity and member motivation are critical determinants of successful collective action within farmer groups. Motivation drives active participation in group activities, as members who perceive tangible benefits from their involvement are more likely to contribute meaningfully (Hartwell et al., 2024). However, motivation alone is insufficient. Without effective managerial support and well-defined collective marketing strategies, even highly motivated members may struggle to achieve significant outcomes (Mangan and Ward, 2024). Thus, the interplay between managerial capacity and member motivation is essential for understanding the effectiveness of collective action in enhancing smallholder farmers' market access.

This study explores how the managerial capacity and motivation of farmer group members influence their roles and how these roles impact collective action to enhance market access for smallholder farmers. By focusing on the case of kepok banana (Musa paradisiaca L.) farmers in Seruyan Regency, Central Kalimantan, this research seeks to provide insights into the institutional factors that improve market access for smallholder farmers, contributing to their economic resilience and the sustainability of the agricultural sector.

The findings are expected to contribute to the literature on smallholder market access and collective action models by offering empirical evidence on the specific mechanisms through which managerial capacity and member motivation drive collective action. These insights aim to inform policies and practices that enhance the economic and social outcomes for smallholder farmers globally.

Materials and methods

Study area

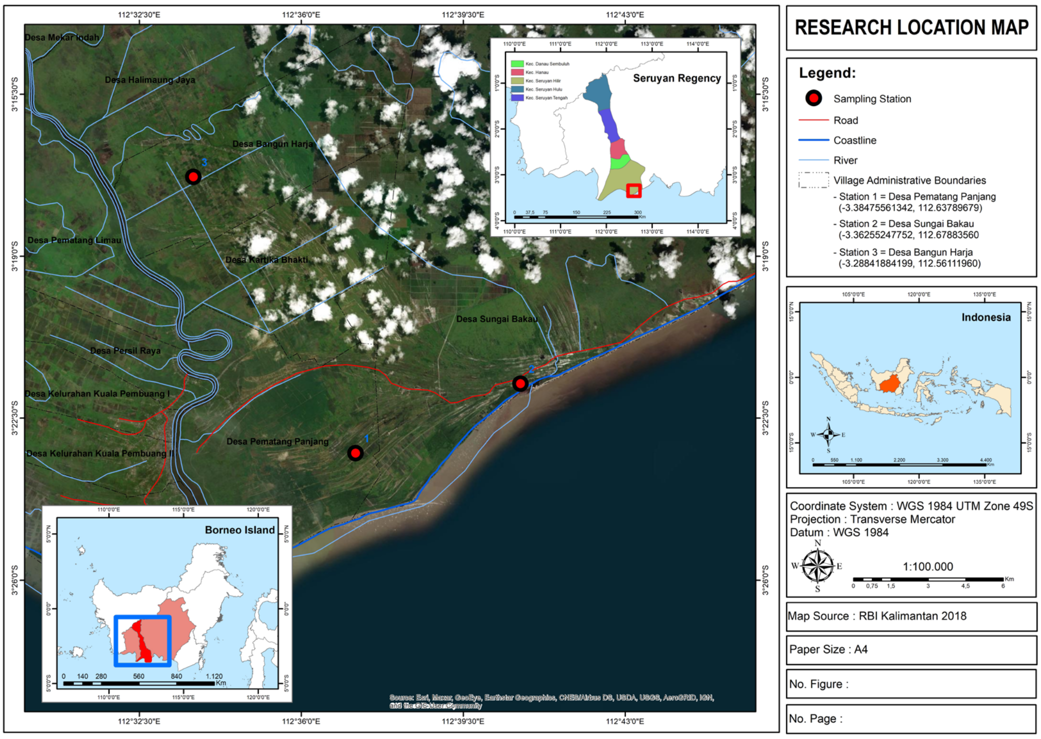
The study was conducted in Seruyan Regency, Central Kalimantan Province, Indonesia, a region recognized for its significant potential in kepok banana production and well-organized farmer groups (Statistics Indonesia, 2023) . This setting provides an ideal context for analyzing the dynamics of collective action and its impact on market access. Geographically, Seruyan Regency is located between latitudes 00°77' to 30°56' South and longitudes 111°49' to 112°84' East (Figure 1).

Sampling procedure

A multistage sampling procedure was employed to select the study participants systematically. In the first stage, Seruyan Regency was purposively chosen based on its significant potential for kepok banana production and marketing activities. In the second stage, three villages (Sungai Bakau, Pematang Panjang, and Bangun Harja) were selected because of their active farmer groups. The final stage involved selecting respondents using Yamane's formula with a 5% margin of error (Yamane, 1973). The formula 1 is as follows:

$$\boldsymbol{n} = \frac{\boldsymbol{N}}{\boldsymbol{1 + N(e)^2}} = \frac{\boldsymbol{655}}{\boldsymbol{1 + (0.05)^2}} = \boldsymbol{249} \qquad (1)$$

Where: n = sample size, N = population size (655 kepok banana farmers), and e = margin of error (0.05). Calculating the sample size yielded 249 respondents. These respondents were randomly selected from the 26 farmer groups within the three villages to ensure a representative sample and minimize selection bias.



Source: Authors´illustration based on the Indonesian Topographic Map (RBI)

Figure 1: Research location.

The farmer groups are semi-formal rural institutions commonly found across Indonesia. They typically operate at the village level and are recognized by local agricultural offices. These groups serve as platforms for collaboration among farmers, focusing on joint planning, training, input procurement, and collective marketing. Each group is usually led by a chairperson and conducts regular meetings to coordinate activities and resolve production or marketing issues collaboratively.

Data collection instrument

Data were collected using a structured questionnaire administered through face-to-face interviews. The questionnaire was organized into two primary sections:

1. Demographic and Socio-Economic Information: Included age, gender, education level, land size, and farming experience.
2. Study Variables: Captured perceptions of managerial capacity, member motivation, the role of farmer groups, collective action, and market access.

Key variables, their associated items, and measurement scales are summarized in Table 1. Responses were recorded on a five-point Likert scale ranging from strongly disagree (1) to strongly agree (5).

Variable	Notation	Item	Scale 1-5
Managerial capacity	Man1	The farmer group consistently formulates an annual plan of activities	strongly disagree – strongly agree
	Man2	The farmer group assigns clear duties and responsibilities to each member	strongly disagree – strongly agree
	Man3	The farmer group ensures that activities are executed in accordance with the pre-established plans	strongly disagree – strongly agree
	Man4	The farmer group monitors and supervises the implementation of activities	strongly disagree – strongly agree
	Man5	The farmer group also conducts evaluations of the activities that have been completed.	strongly disagree – strongly agree

Source: Author processing, 2024

Table 1: Key variables and measurement scales (To be continue).

Variable	Notation	Item	Scale 1-5
Member motivation	Mot1	Did you join the farmer group because you want to receive guidance from the group	definitely not – definitely yes
	Mot2	Did you join the farmer group because you want easy access to the facilities provided by the group	definitely not – definitely yes
	Mot3	Did you join the farmer group because you want to improve my skills and knowledge through the group	definitely not – definitely yes
Role of the farmer group	Rule1	The farmer group always provides the latest information on kepok banana market prices	strongly disagree – strongly agree
	Rule2	The farmer group provides transportation facilities to market kepok bananas	strongly disagree – strongly agree
	Rule3	The farmer group encourages members to use technology in marketing kepok bananas	strongly disagree – strongly agree
	Rule4	The farmer group frequently holds meetings with external parties to discuss marketing cooperation for kepok bananas	strongly disagree – strongly agree
Collective action	Coll1	The farmer group often communicates with other farmer group members about kepok banana marketing strategies	strongly disagree – strongly agree
	Coll2	The farmer group often conducts joint planning for marketing activities	strongly disagree – strongly agree
	Coll3	The farmer group frequently shares resources (tools, information, funds) for kepok banana marketing activities	strongly disagree – strongly agree
	Coll4	The farmer group often organizes activities that involve the entire community or external parties to support kepok banana marketing	strongly disagree – strongly agree
Market access	Acc1	I have found it easier to enter new markets for selling kepok bananas	strongly disagree – strongly agree
	Acc2	I now have better access to distant markets where I can sell my produce	strongly disagree – strongly agree
	Acc3	I can quickly and comprehensively obtain relevant information about kepok bananas	strongly disagree – strongly agree
	Acc4	I am able to efficiently distribute my harvested kepok bananas to the market	strongly disagree – strongly agree

Source: Author processing, 2024

Table 1: Key variables and measurement scales (Continuation).

Data analysis techniques

Data analysis was conducted using Partial Least Squares Structural Equation Modeling (PLS-SEM) with the SmartPLS software. SEM is a robust econometric modeling technique frequently employed for testing complex marketing theories and examining theoretical concepts in real-world situations (Benitez et al., 2019; Martínez-López et al., 2013). PLS-SEM was chosen for its advantages in handling small sample sizes and missing data, as well as its capability to manage both reflective and formative indicators (Hair et al., 2019; Lowry and Gaskin, 2014). This approach has also been adopted in recent empirical studies related to farmer groups and cooperatives. For instance, Iyioku et al., (2024) applied PLS-SEM to analyze the motivational factors and social capital influencing farmer participation in cashew marketing cooperatives in Kenya. Similarly, Wang and Wang, (2024) employed PLS-SEM to examine how collective action capabilities affect economic performance in rural cooperatives in China. These examples underscore the method's relevance in capturing behavioral and institutional dynamics within smallholder farming systems.

The analytical process adopted a two-stage approach to ensure the rigor and validity of the study's findings. The first stage involved evaluating the measurement model to confirm the reliability and validity of the constructs. Indicator reliability was assessed through factor loadings, where values greater than 0.5 were considered acceptable (Hair et al., 2019). Internal consistency reliability was evaluated using Cronbach's alpha and Rho A, with thresholds set at values above 0.7 (Hair et al., 2019). Convergent validity was determined by calculating the Average Variance Extracted (AVE), where values exceeding 0.5 indicated sufficient convergence (Hair et al., 2019). Composite Reliability (CR) was also employed to verify internal consistency, with acceptable values set above 0.7 (Hair et al., 2019). Discriminant validity was confirmed through cross-loadings, ensuring that each item's loading on its designated construct was higher than its loadings on other constructs. The Fornell-Larcker criterion was applied to verify that the square root of each construct's AVE was greater than its correlations with other constructs, and the Heterotrait-Monotrait Ratio (HTMT) was used, with values below 0.90 indicating adequate discriminant validity (Hair et al., 2019).

Once the reliability and validity of the measurement model were established, the second stage focused on evaluating the structural model. Multicollinearity among predictor variables was examined using the Variance Inflation Factor (VIF), with acceptable values set below 3.00 (Hair et al., 2019). The model's explanatory power was assessed through the Coefficient of Determination (R^2), which quantifies the variance explained in the endogenous constructs. The effect size (f^2) was calculated to determine the impact of each exogenous variable on the endogenous variables, while predictive relevance (q^2) was assessed using the blindfolding procedure to evaluate the model's predictive accuracy (Hair et al., 2019).

To test the significance of the hypothesized relationships (Figure 2), a bootstrapping procedure with 5,000 subsamples was performed. This procedure provided estimates of standard errors and t-statistics for the path coefficients, enabling robust inferences regarding the direct and indirect effects among the constructs (Hair et al., 2019). This comprehensive analytical approach ensured the validity and reliability of the results, thereby enhancing the robustness of the study's conclusions.

Figure 2 shows the conceptual framework of the study, outlining the relationships between managerial capacity, member motivation, the role of farmer groups, collective action, and market access. The hypothesized relationships are grounded in organizational behavior and collective action theory, as discussed by Markelova and Mwangi (2010), Hartwell et al., (2024), and supported empirically by Mangan and Ward (2024). These theoretical foundations support the relevance of the model in analyzing the internal drivers of farmer group performance.

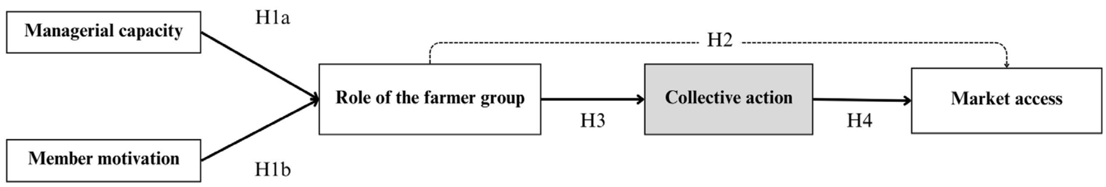
Results and discussion

Demographic characteristics of respondents

The study surveyed a total of 249 kepok banana farmers from three villages in Seruyan Regency, Central Kalimantan: Bangun Harja, Sungai Bakau, and Pematang Panjang. The demographic characteristics of the respondents provide valuable insights into the socio-economic context of the farmers involved in this study, as shown in Table 2.

A significant majority of the respondents were male (86.8%), reflecting the gender dynamics within the agricultural sector in the region. The largest age group was between 36 and 45 years (33.7%), indicating a mature farming population with potentially substantial experience. Educational attainment was generally low, with over half of the respondents (55.0%) having only completed elementary school, and a small fraction (3.2%) possessing university-level education.

Regarding farming experience, half of the farmers (50.60%) had between 6 to 10 years of experience, while 39.76% had five years or less. Landholding sizes were predominantly small, with 51% of farmers owning less than 2 hectares, classifying them as smallholders. These demographics highlight



Source: Authors

Figure 2: Conceptual framework of the study.

Characteristic		Frequency	Percentage (%)
Sex	Male	216	86.75
	Female	33	13.25
Age	< 25 years	24	9.64
	26-35 years	60	24.10
	36-45 years	84	33.73
	46-55 years	54	21.69
	> 55	27	10.84
Farming experience	< 5 years	99	39.76
	6–10 years	126	50.60
	> 10 years	24	9.64
Education level	No formal education	2	0.80
	Elementary school	137	55.02
	High school	102	40.96
	University	8	3.21
Land size	< 2 Ha	127	51.00
	2-4 Ha	88	35.34
	> 4 Ha	34	3.65

Source: Author compilation, 2024

Table 2: Demographic characteristics of survey respondents.

the challenges faced by smallholder farmers in terms of resources and access to education, which may influence their capacity for collective action and market engagement.

Measurement model evaluation

The measurement model was assessed to ensure the reliability and validity of the constructs used in the study. Reliability was evaluated using Composite Reliability (CR) and Cronbach's alpha. According to (Hair et al., 2019), both CR and Cronbach's alpha values should exceed the threshold of 0.70 to indicate high internal consistency. As presented in Table 3, all constructs met this criterion, demonstrating satisfactory reliability.

Construct	Composite Reliability (CR)	Cronbach's Alpha	AVE
Managerial capacity	0.894	0.884	0.743
Member motivation	0.799	0.761	0.630
Role of the farmer group	0.865	0.846	0.682
Collective action	0.887	0.887	0.746
Market access	0.889	0.889	0.750

Source: Author compilation, 2024

Table 3: Reliability and validity assessment of constructs.

Convergent validity was assessed by examining the Average Variance Extracted (AVE) for each construct. Chin (2010) recommends that AVE values should be above 0.50 to confirm that the constructs adequately capture the relevant variance. As shown in Table 3, all AVE values exceeded 0.50, confirming convergent validity.

Discriminant validity was evaluated using the Fornell-Larcker criterion and the Heterotrait-Monotrait (HTMT) ratio. The Fornell-Larcker criterion requires that the square root of each construct's AVE be higher than its correlations with other constructs. The HTMT ratio should be less than 0.90 to indicate adequate discriminant validity (Hair et al., 2019). Both criteria were met in this study (Table 4 and 5), confirming that each construct was distinct from the others.

The cross-loading analysis further supported discriminant validity, which examines whether each measurement item loads more significantly on its assigned construct than on any other constructs (Chin, 2010; Hair et al., 2019). According to the criteria, an indicator should exhibit the highest loading on its own latent variable to confirm discriminant validity. As illustrated in Table 6, all items demonstrated higher loadings on their respective constructs compared to loadings on other constructs. For instance, items measuring Managerial Capacity loaded strongly on that construct, with loadings ranging from 0.65 to 0.98, while their loadings on other constructs were considerably lower.

Constructs	Managerial capacity	Member motivation	Role of the farmer group	Collective action	Market access
Managerial capacity	0.862				
Member motivation	0.051	0.794			
Role of the farmer group	0.498	0.101	0.826		
Collective action	0.449	0.099	0.616	0.864	
Market access	0.307	0.105	0.536	0.628	0.866

Source: Author compilation, 2024

Table 4: Discriminant validity (Fornell–Larcker criterion).

Constructs	Collective action	Managerial capacity	Market access	Member motivation
Managerial capacity	0.499			
Market access	0.707	0.408		
Member motivation	0.091	0.100	0.107	
Role of the farmer group	0.700	0.563	0.603	0.107

Source: Author compilation, 2024

Table 5: Heterotrait–monotrait ratio (HTMT).

Variable	Collective action	Managerial capacity	Market access	Member motivation	Role of the farmer group
Acc1	0.528	0.348	0.845	0.176	0.460
Acc2	0.559	0.275	0.846	0.056	0.426
Acc3	0.545	0.359	0.897	0.072	0.484
Acc4	0.541	0.300	0.875	0.062	0.486
Coll1	0.862	0.428	0.540	0.089	0.563
Coll2	0.865	0.385	0.543	0.127	0.480
Coll3	0.883	0.362	0.546	0.074	0.537
Coll4	0.845	0.374	0.539	0.055	0.547
Man1	0.461	0.891	0.350	0.062	0.450
Man2	0.356	0.898	0.347	0.065	0.437
Man3	0.430	0.859	0.376	0.027	0.466
Man4	0.277	0.795	0.173	0.016	0.350
Mot1	0.117	0.061	0.130	0.979	0.117
Mot2	-0.004	-0.065	-0.034	0.646	0.016
Mot3	0.033	0.043	0.027	0.717	0.032
Rule1	0.499	0.442	0.452	0.053	0.833
Rule2	0.444	0.349	0.331	0.076	0.779
Rule3	0.452	0.361	0.386	0.121	0.811
Rule4	0.612	0.471	0.560	0.089	0.876

Source: Author compilation, 2024

Table 6: Cross-loadings of measurement items.

These results confirm that each item correlates most strongly with the construct it is intended to measure, satisfying the requirements for discriminant validity through cross-loadings (Hair et al., 2019). The distinctiveness of each construct is thereby established, reducing concerns about multicollinearity and ensuring that the constructs are measuring separate concepts within the model.

Collectively, the reliability and validity assessments including internal consistency reliability, convergent validity, and discriminant validity indicate that the measurement model is robust. These validations provide confidence in the constructs used and support the progression to the structural model assessment, where the hypothesized relationships between the constructs can be examined with a solid foundation (Shrestha, 2021).

Structural model evaluation and hypothesis testing

After confirming the measurement model's adequacy, the structural model was evaluated to examine the hypothesized relationships among the constructs. Multicollinearity among the predictor variables was assessed using the Variance Inflation Factor (VIF) and tolerance levels. According to Hair et al., (2011), VIF values exceeding 5 indicate significant collinearity issues, while Collier (2020) suggests that tolerance values greater than 0.10 are preferable. Becker et al., (2015) argue that collinearity can be a concern at VIF values as low as 3. In this study, all VIF values were below 3, as shown in Table 7, indicating that multicollinearity was not a concern and the model's estimates were reliable.

Endogenous variable	Predictor variables	VIF
Role of farmer group	Managerial capacity	1.003
	Member motivation	1.003
Collective action	Role of farmer group	1.000
Market access	Role of farmer group	1.612
	Collective action	1.612

Source: Author compilation, 2024

Table 7: Variance inflation factor (VIF) values.

Once the measurement model was assessed correctly, and the subsequent SEM analysis was deemed reliable (Shrestha, 2021), the structural paths were examined to determine the relationships between the study constructs and their statistical significance. The determination coefficients (R^2) were calculated to assess the amount of variance in the endogenous constructs explained by the predictor variables. According to Hair et al., (2019), R^2 values of 0.25, 0.50, and 0.75 indicate weak, moderate, and substantial explanatory power, respectively. The R^2 values in this study are presented in Table 8. The R^2 values indicate that the model has moderate explanatory power for the endogenous variables. Predictive relevance (Q^2) was assessed using the blindfolding procedure, with values greater than zero indicating adequate predictive accuracy (Hair et al., 2019). All Q^2 values were above zero, confirming the model's predictive relevance.

Effect sizes (f^2) were calculated to evaluate the impact of each exogenous variable on the endogenous variables. (Hair et al., 2019) suggest that f^2 values of 0.02, 0.15, and 0.35 represent small, medium, and large effects, respectively. The effect sizes are presented in Table 9. The results indicate that managerial capacity has a large effect on the role of the farmer group ($f^2 = 0.326$), while member motivation has a small effect ($f^2 = 0.008$). The role of the farmer group strongly influences collective action ($f^2 = 0.612$), and collective action has a medium effect on market access ($f^2 = 0.250$).

The hypothesized relationships between the constructs were tested using path coefficients and their significance levels through a bootstrapping procedure with 5,000 subsamples (Hair et al., 2019). The results presented in Table 10 and Figure 3 indicate that managerial capacity has a significant positive effect on the role of the farmer group ($\beta = 0.494$, $t = 8.631$, $p < 0.000$), supporting the hypothesis that effective managerial skills enhance group functionality. In contrast, member motivation does not significantly influence the role of the farmer group ($\beta = 0.076$, $t = 1.059$, $p = 0.290$), suggesting that motivation alone may be insufficient without opportunities for meaningful engagement

Construct	Q^2	R^2	Adjusted R^2	Criteria
Role of the farmer group	0.166	0.253	0.252	Moderate
Collective action	0.279	0.380	0.379	Moderate
Market access	0.318	0.430	0.428	Moderate

Source: Author compilation, 2024

Table 8: Coefficient of determination (R^2) and Predictive Relevance (Q^2).

Endogenous variable	Exogenous variables	f^2	Effect vize
Role of farmer group	Managerial capacity	0.326	Large
	Member motivation	0.008	Small
Collective action	Role of farmer group	0.612	Large
Market access	Role of farmer group	0.063	Medium
	Collective action	0.250	Medium

Source: Author compilation, 2024

Table 9: Effect sizes (f^2) of predictor variables.

or managerial facilitation. The role of the farmer group significantly influences collective action ($\beta = 0.616$, $t = 12.966$, $p < 0.001$) and market access ($\beta = 0.240$, $t = 3.385$, $p = 0.001$). Furthermore, collective action also exerts a significant positive effect on market access ($\beta = 0.479$, $t = 7.062$, $p < 0.001$). These findings confirm that strong internal group dynamics and coordinated efforts are essential to improving access to competitive markets. The results underscore the pivotal role of managerial capacity in strengthening farmer group functions, which in turn fosters effective collective action and facilitates improved market access for smallholder farmers.

Yanduri and Siddayya, (2024) emphasize that managerial competencies are pivotal determinants of organizational success. Research by Ton et al. (2015) supports this notion, highlighting that managerial capacity fosters effective resource utilization, coordination, and strategic planning, which are critical for collective marketing. Similarly, Prabhavathi et al. (2023) find a significant correlation between managerial skills and improved financial performance and sustainable growth in farmer organizations. By equipping leaders with targeted training programs and practical experience, this study demonstrates the transformative potential of managerial capacity in enhancing group functionality. Moreover, Markelova and Mwangi, (2010) argue that leadership, when combined with enabling environments, drives successful collective action, reducing transaction costs and enhancing market access for smallholder farmers. The significant positive influence observed in this study reinforces the vital role of leadership in overcoming market challenges, fostering cohesive group efforts, and ensuring the sustainability of agricultural practices. Collectively, these findings emphasize that managerial capacity serves as the backbone of effective farmer groups, enabling them to navigate complex market dynamics and secure economic resilience.

Interestingly, while managerial capacity significantly influenced the role of the farmer group, member motivation did not exhibit a statistically significant impact. This finding suggests that motivation alone may not be sufficient to enhance group performance if it is not accompanied by structured opportunities for engagement and institutional support. Motivated members may still be hindered by limited decision-making authority, unclear roles, or a lack of inclusive governance mechanisms within the group. These barriers can suppress the active participation necessary for translating motivation into meaningful contributions.

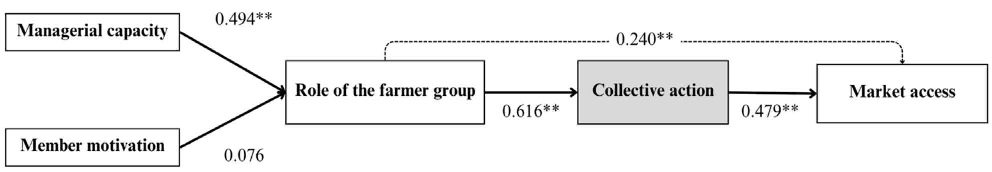
Furthermore, the strong influence of managerial capacity may indicate a compensatory effect, whereby effective leadership and planning structures reduce reliance on individual motivation. In such contexts, a capable management team may drive group functionality through top-down mechanisms, setting objectives and ensuring execution, even if member participation remains passive. While this may ensure short-term effectiveness, long-term sustainability and group cohesion could be at risk if member engagement remains superficial.

These findings align with the observations

Relationship	Path coefients (β value)	t-value	p-value	Hypothesis supported
Managerial capacity > Role of the farmer group	0.494	8.631	0.001	Yes
Member motivation > Role of the farmer group	0.076	1.059	0.290	No
Role of the farmer group > Collective action	0.616	12.966	0.001	Yes
Role of the farmer group > Market access	0.240	3.385	0.001	Yes
Collective action > Market access	0.479	7.062	0.001	Yes

Source: Author compilation, 2024

Table 10: Path coefficient and hypothesis testing.



Note:*and** denote a 5% and 1% significance level, respectively
Source: Authors

Figure 3: Structural model results.

of Mangan and Ward (2024), who argue that collective success depends not just on willingness but on the structural capacity to act. Similarly, Hartwell et al., (2024) emphasize that during periods of institutional stress, the absence of participatory mechanisms can nullify even strong intrinsic motivation. Hence, efforts to strengthen the role of farmer groups should focus not only on managerial skill-building but also on creating systems that empower member voices and engagement.

The study also found that the role of the farmer group significantly influences collective action, which in turn has a substantial positive impact on market access. This highlights the pivotal role of farmer groups as facilitators of collective action, enabling smallholder farmers to overcome market barriers. Konja and Abdulai (2024) identified collective market action as a means of enhancing farm performance, noting that by pooling resources and coordinating efforts, farmer groups can access markets that would otherwise be inaccessible to individual farmers. Similarly, Bikkina et al. (2018) found that collective efforts, such as resource pooling and information sharing, help farmers overcome logistical and financial constraints.

Farmer groups serve as critical platforms for enhancing collective bargaining power, which is essential for negotiating better terms with buyers and reducing transaction costs. According to Markelova and Mwangi (2010), collective action enables smallholder farmers to leverage their collective strength to secure better prices, improve market linkages, and reduce the risk of exploitation by intermediaries. This aligns with findings from Ssajakambwe et al. (2020), who emphasize that coordinated group efforts can help farmers establish long-term contracts with large-scale buyers, fostering greater market stability and income security.

Moreover, strong internal governance within farmer groups plays a crucial role in sustaining effective collective action. Groups with clear operational guidelines, transparent decision-making processes, and active member participation are more likely to achieve their goals (Barham and Chitemi, 2009). This study's findings also emphasize the importance of leadership and institutional maturity in ensuring that collective efforts translate into tangible market benefits. As Fischer and Qaim (2012) suggest, fostering trust and cohesion within groups enhances their ability to address logistical challenges and exploit high-value market opportunities.

However, the success of farmer groups as facilitators of collective action is not without challenges. Weak leadership, lack of resources, and unequal benefit distribution can undermine group effectiveness (McLeay et al., 1996). Addressing these challenges requires targeted interventions, such as capacity-building programs for group leaders and the provision of financial and technical support to enhance group functionality. Bikkina et al., (2018) highlight the need for external support from governments and NGOs to strengthen group structures and provide access to critical resources, ensuring the sustainability of collective market initiatives.

The positive impact of collective action on market access is consistent with findings by Anania and Towo (2016), who reported that farmer groups in Tanzania enhanced market access and improved bargaining power through structured collective strategies. Liu et al., (2019) advocate for capacity-building initiatives within agricultural institutions to enhance smallholder farmers' bargaining power and market access. By investing in leadership development and managerial training, farmer groups can improve their organizational structures and strategic capabilities, leading to better collective action outcomes.

Moreover, the role of collective action in promoting sustainable agricultural practices is significant. Research indicates that participation in collective organizations positively influences farmers' decisions regarding the adoption of soil and water conservation measures, thereby enhancing their resilience to climate change (Jia et al., 2024). This is further supported by findings from the MASIPAG network in the Philippines, which emphasizes the importance of self-reflection and critique among farmer organizations to adapt and thrive in changing agricultural landscapes (Jack et al., 2022).

This study provides valuable insights into the roles of managerial capacity and member motivation in enhancing collective action and market access. However, several limitations must be acknowledged. The research was geographically limited to kepok banana farmer groups in Seruyan Regency, which may reduce the generalizability of the findings to other crops, regions, or organizational contexts. In addition, the study only examined two internal factors—managerial capacity and member motivation—while omitting other potential influences such as financial access, technology adoption, and institutional support from governments or NGOs. Future studies should consider incorporating these variables to better capture the multidimensional nature of collective action.

Another limitation stems from the use of self-reported data, which may introduce response bias, particularly in measuring subjective constructs such as motivation and managerial effectiveness. Although PLS-SEM is robust against measurement error, triangulating results with qualitative or observational data in future research would improve validity. Moreover, the study did not include a gender perspective, despite evidence that collective action often affects men and women differently due to cultural and structural inequalities. The study focused solely on farmers who are active members of producer groups, which may limit the generalizability of the findings to unaffiliated farmers. The relatively small size of farmer groups may influence internal dynamics and group performance, potentially enhancing coordination but limiting economies of scale. Future studies should include both affiliated and non-affiliated farmers to provide a broader understanding of collective action dynamics.

The study contributes to the literature by providing empirical evidence on the specific mechanisms through which managerial capacity influence collective action and market access. It highlights the need for policies and interventions that focus on building the managerial skills of farmer group leaders and encouraging active participation among members. Such efforts can lead to more effective collective strategies that enhance the competitiveness of smallholder farmers in the market. Future research should aim to address the identified limitations by exploring additional factors that may influence collective action and market access, considering different agricultural contexts, and incorporating gender analyses. By doing so, researchers can develop a more holistic understanding of how to support smallholder farmers through effective farmer groups and collective action initiatives.

Conclusion

The findings demonstrate that managerial capacity is a critical determinant of a farmer group's success in facilitating collective action and improving market access. Specifically, effective leadership and management practices significantly enhance the group's functionality, enabling it to organize and implement collective strategies more efficiently.

Conversely, member motivation alone was not found to have a significant impact on the role of the farmer group. This finding supports earlier discussions suggesting that motivation, while important, may not be sufficient in the absence of structured engagement mechanisms, inclusive governance, or clear participation channels. When opportunities for meaningful involvement are limited, even highly motivated members may remain passive or disengaged.

The implications of these findings are substantial for policymakers, development practitioners, and stakeholders involved in agricultural development. Enhancing managerial capacity within farmer groups through targeted training programs and capacity-building initiatives can empower leaders to drive collective action effectively. By focusing on these institutional capacities, interventions can facilitate collective strategies that overcome market barriers, increase bargaining power, and improve the economic resilience of smallholder farmers.

Future research should address the limitations of this study by expanding the geographic scope to include diverse agricultural contexts and crop types. Incorporating additional variables such as access to financial services, technological adoption, and external support can provide a more comprehensive understanding of the factors influencing collective action and market access. Moreover, integrating a gender perspective is crucial, as the effects of collective action on women farmers remain underexplored. Investigating how gender dynamics affect participation and benefits within farmer groups can inform more inclusive and equitable strategies.

Acknowledgments

The authors would like to express their sincere gratitude to the Directorate of Research, Technology, and Community Service (DRTPM), Ministry of Education, Culture, Research, and Technology, for their invaluable support and funding for this research. This study would not have been possible without their generous contribution. We are also thankful for the continuous guidance and encouragement provided throughout the research process, which has significantly contributed to the successful completion of this project.

Corresponding author:
Dr. Rokhman Permadi
Agribusiness Department, Universitas Darwan Ali
Central Kalimantan 74321, Indonesia
Phone: +6281230351487, E-mail: rokhmanpermadi@gmail.com

References

[1] Abdul-Rahaman, A. and Abdulai, A. (2020) "Farmer groups, collective marketing and smallholder farm performance in rural Ghana", *Journal of Agribusiness in Developing and Emerging Economies*, Vol. 10, No. 5, pp. 511-527. E-ISSN 2044-0847, ISSN 2044-0839. DOI 10.1108/JADEE-07-2019-0095.

[2] Akpa, A. F., Osabohien, R., Ashraf, J. and Al-Faryan, M. A. S. (2023) "Financial inclusion and post-harvest losses in West African Economic and Monetary Union", *Agricultural Finance Review*, Vol. 83, No. 2, p. 320-332. E-ISSN 2044-0847, ISSN 2044-0839. DOI 10.1108/AFR-06-2022-0076.

[3] Aku, A., Mshenga, P., Afari-Sefa, V. and Ochieng, J. (2018) "Effect of market access provided by farmer organizations on smallholder vegetable farmer's income in Tanzania", *Cogent Food & Agriculture*, Vol. 4, pp. 1-13. ISSN 2331-1932. DOI 10.1080/23311932.2018.1560596.

[4] Anania, P. F. and Towo, P. E. (2016) "The Contribution of Agricultural Marketing Co-operatives in Service Provision to Members in Tanzania: A Case of Moshi District", *Tengeru Community Development Journal*, Vol. 3, No. 2. ISSN 2665-0584.

[5] Barham, J. and Chitemi, C. (2009) "Collective action initiatives to improve marketing performance: Lessons from farmer groups in Tanzania", *Food Policy*, Vol. 34, pp. 53-59. ISSN 0306-9192. DOI 10.1016/j.foodpol.2008.10.002.

[6] Barrett, C. B. (2008) "Smallholder market participation: Concepts and evidence from eastern and southern Africa", *Food Policy,* Vol. 33, No. 4, pp. 299-317. ISSN 0306-9192. DOI 10.1016/j.foodpol.2007.10.005.

[7] Becker, J.-M., Ringle, C. M., Sarstedt, M. and Völckner, F. (2015) "How collinearity affects mixture regression results", *Marketing Letters*, Vol. 26, No. 4, pp. 643-659. ISSN 0923-0645. DOI 10.1007/s11002-014-9299-9.

[8] Benitez, J., Henseler, J., Castillo, A. and Schuberth, F. (2019) "How to perform and report an impactful analysis using partial least squares: Guidelines for confirmatory and explanatory IS research", *Information & Management*, Vol. 57, No. 2, p. 103168. E-ISSN 1872-753, ISSN 0378-7206. DOI 10.1016/j.im.2019.05.003.

[9] Bikkina, N., Turaga, R. M. R. and Bhamoriya, V. (2018) "Farmer producer organizations as farmer collectives: A case study from India", *Development Policy Review*, Vol. 36, No. 6, pp. 669-687. DOI 10.1111/dpr.12274.

[10] Chin, W. W. (2010) "How to Write Up and Report PLS Analyses", In: Esposito Vinzi, V., Chin, W., Henseler, J., Wang, H. (eds) "*Handbook of Partial Least Squares. Springer Handbooks of Computational Statistics*". Springer, Berlin, Heidelberg. E-ISBN 978-3-540-32827-8. DOI 10.1007/978-3-540-32827-8_29.

[11] Collier, J. E. (2020) "Applied Structural Equation Modeling Using AMOS", 1st ed., Routledge. 366 p. E-ISBN 9781003018414. DOI 10.4324/9781003018414.

[12] Courtois, P. and Subervie, J. (2014) "Farmer bargaining power and market information services", *American Journal of Agricultural Economics*, Vol. 97, No. 3, pp. 953-977. E-ISSN1467-8276, ISSN 0002-9092. DOI 10.1093/ajae/aau051.

[13] Fan, Q. and Salas Garcia, V. B. (2018) "Information Access and Smallholder Farmers' Market Participation in Peru", *Journal of Agricultural Economics*, Vol. 69, No. 2, pp. 476-494. E-ISSN 1477-9552. ISSN 0021-857X. DOI 10.1111/1477-9552.12243.

[14] FAO and IFAD. (2019) "*The future of family farming in the context of the 2030 Agenda*". [Online]. Available: https://openknowledge.fao.org/handle/20.500.14283/ca4778en [Accessed: Oct. 16, 2024].

[15] Fischer, E. and Qaim, M. (2012) "Linking Smallholders to Markets: Determinants and Impacts of Farmer Collective Action in Kenya", *World Development*, Vol. 40, No. 6, pp. 1255-1268. E-ISSN 1873-5991, ISSN 0305-750X. DOI 10.1016/j.worlddev.2011.11.018.

[16] Hair, J. F., Ringle, C. M. and Sarstedt, M. (2011) "PLS-SEM: Indeed a Silver Bullet", *Journal of Marketing Theory and Practice*, Vol. 19, No. 2, pp. 139-151. E-ISSN 1944-7175. ISSN 1069-6679. DOI 10.2753/MTP1069-6679190202.

[17] Hair, J. F., Risher, J. J., Sarstedt, M. and Ringle, C. M. (2019) "When to use and how to report the results of PLS-SEM", *European Business Review*, Vol. 31, No. 1, pp. 2-24. E-ISSN 1758-7107, ISSN 0955-534X. DOI 10.1108/EBR-11-2018-0203.

[18] Hartwell, C. A., Lawton, T. C. and Tingbani, I. (2024) "Obstacles to collective action during a crisis: A meta-organizational perspective", *European Management Review*, Vol. 21, Vol. 2, pp. 296-314. ISSN 1740-4762. DOI 10.1111/emre.12596.

[19] Iyioku, G., Donkor, E., Mazancova, J. and Hejkrlik, J. (2024) "Understanding farmers' motivations for joining cashew marketing cooperatives in coastal Kenya: the moderating role of social capital", *Cogent Food & Agriculture*, Vol. 10, No. 1. E-ISSN 2331-1932. DOI 10.1080/23311932.2024.2433600.

[20] Jack, G., Plahe, J. and Wright, S. (2022) "Development as freedom? Insights from a farmer-led sustainable agriculture non-governmental organisation in the Philippines", *Human Relations*, Vol. 75, No. 10, pp. 1875-1902. E-ISSN 1741-282X. ISSN 0018-7267. DOI 10.1177/00187267221090779.

[21] Jia, R., Shuai, Z., Guo, T., Lu, Q., He, X. and Hua, C. (2024) "Impact of participation in collective action on farmers' decisions and waiting time to adopt soil and water conservation measures", *International Journal of Climate Change Strategies and Management*, Vol. 16, No. 2, pp. 201-227. E-ISSN 1756-8706. DOI 10.1108/IJCCSM-02-2023-0027.

[22] Kiveu, M. and Ofafa, G. (2013) "Enhancing market access in Kenyan SMEs using ICT", *Global Business and Economics Research Journal*, Vol. 2, No. 9, pp. 29-46. ISSN 2302-4593.

[23] Konja, D. T. and Abdulai, A. (2024) "Collective market action, farm performance, and household welfare among maize farmers: the role of outgrower scheme in Northern Ghan", *Applied Economics*. pp. 1-17. ISSN 1466-4283. DOI 10.1080/00036846.2024.2364927.

[24] Liu, Y., Ma, W., Renwick, A. and Fu, X. (2019) "The role of agricultural cooperatives in serving as a marketing channel: Evidence from low-income regions of Sichuan province in China", *International Food and Agribusiness Management Review*, Vol. 22, No. 2, pp. 265-282. ISSN 1559-2448. DOI 10.22434/IFAMR2018.0058.

[25] Lowder, S. K., Sánchez, M. V. and Bertini, R. (2019) "Farms, family farms, farmland distribution and farm labour: What do we know today?", FAO Agricultural Development Economics Working Paper 19-08. Rome, FAO. [Online]. Available: https://openknowledge.fao.org/server/api/core/bitstreams/e6d7a673-cebd-4be6-9b77-686d50d9171c/content [Accessed: Oct. 20, 2024].

[26] Lowry, P. B. and Gaskin, J. (2014) "Partial Least Squares (PLS) Structural Equation Modeling (SEM) for Building and Testing Behavioral Causal Theory: When to Choose It and How to Use It", *IEEE Transactions On Professional Communication*, Vol. 57, No. 2, pp. 123-146. ISSN 0361-1434. DOI 10.1109/TPC.2014.2312452.

[27] Magesa, M. M., Michael, K. and Ko, J. (2014) "Access to Agricultural Market Information by Rural Farmers in Tanzania", *International Journal of Information and Communication Technology Research*, Vol. 4, No. 7, pp. 264-273. E-ISSN 2783-4425. ISSN 2251-6107.

[28] Mangan, A. and Ward, A. M. (2024) "Cooperative resilience: Toward a heuristic model of collective action in a crisis", *Financial Accountability and Management*, Vol. 40, No. 3, pp. 308-325. E-ISSN 1468-0408. DOI 10.1111/faam.12382.

[29] Markelova, H., Meinzen-Dick, R., Hellin, J. and Dohrn, S. (2009) "Collective action for smallholder market access", *Food Policy*, Vol. 34, No. 1, pp. 1-7. ISSN 0306-9192. DOI 10.1016/j.foodpol.2008.10.001.

[30] Markelova, H. and Mwangi, E. (2010) "Collective Action for Smallholder Market Access: Evidence and Implications for Africa", *Review of Policy Research*, 27, 621-640. E-ISSN 1541-1338, ISSN 1541-132X. DOI 10.1111/J.1541-1338.2010.00462.X.

[31] Martínez-López, F. J., Gázquez-Abad, J. C. and Sousa, C. M. P. (2013) "Structural equation modelling in marketing and business research: Critical issues and practical recommendations", *European Journal of Marketing*, Vol. 47, No. 1, pp. 115-152. E-ISSN 1758-7123, ISSN 0309-0566. DOI 10.1108/03090561311285484.

[32] McLeay, F., Martin, S. and Zwart, T. (1996) "Farm business marketing behavior and strategic groups in agriculture", *Agribusiness*, Vol. 12, pp. 339-351. E-ISSN 0742-4477, ISSN 1520-6297. DOI 10.1002/(SICI)1520-6297(199607/08)12:4<339::AID-AGR4>3.0.CO;2-%23.

[33] Migose, S. A., Bebe, B. O., de Boer, I. J. M. and Oosting, S. J. (2018) "Influence of distance to urban markets on smallholder dairy farming systems in Kenya", *Tropical Animal Health and Production*, Vol. 50, No. 7, pp. 1417-1426. E-ISSN 1573-7438, ISSN 0049-4747. DOI 10.1007/s11250-018-1575-x.

[34] Otekunrin, O. A., Momoh, S. and Ayinde, I. A. (2019) "Smallholder Farmers' Market Participation: Concepts and Methodological Approach from Sub-Saharan Afric", *Current Agriculture Research Journal*, Vol. 7, No. 2, pp. 139-157. E-ISSN 2321-9971, ISSN 2347-4688. DOI 10.12944/carj.7.2.02.

[35] Pachoud, C., Delay, E., Da Re, R., Ramanzin, M. and Sturaro, E. (2020) "A Relational Approach to Studying Collective Action in Dairy Cooperatives Producing Mountain Cheeses in the Alps: The Case of the Primiero Cooperative in the Eastern Italians Alps", *Sustainability*, Vol. 12, No. 11, p. 4596. ISSN 2071-1050. DOI 10.3390/su12114596.

[36] Prabhavathi, Y., Kishore, N., Siddayya and Ramachandra, C. (2023) "An Analytical Study on Managerial Competencies and Business Performance of Farmer Producer Organizations (FPOs): A Value Chain Perspective from India", *Millennial Asia*. E-ISSN 2321-7081, ISSN 0976-3996. DOI 10.1177/09763996231200180.

[37] Rhoads, T. and Shogren, J. (1999) "On Coasean bargaining with transaction costs", *Applied Economics Letters*, Vol. 6, pp. 779-783. ISSN 1350-4851. DOI 10.1080/135048599352150.

[38] Shrestha, N. (2021) "Factor Analysis as a Tool for Survey Analysis", *American Journal of Applied Mathematics and Statistics*, Vol. 9, No. 1, pp. 4-11. E-ISSN 2328-7292, ISSN 2328-7306. DOI 10.12691/ajams-9-1-2.

[39] Sitoe, T. A. and Sitole, A. (2019) "Determinants of Farmer's Participation in Farmers' Associations: Empirical Evidence from Maputo Green Belts, Mozambique", *Asian Journal of Agricultural Extension, Economics & Sociology*, Vol. 37, No. 1, pp. 1-12. ISSN 2320-7027. DOI 10.9734/ajaees/2019/v37i130259.

[40] Ssajakambwe, F., Elepu, G., Walekhwa, P. N. and Mulebeke, R. (2020) "Collective action for improved market access among smallholder maize farmers in Masindi District, Uganda", *African Journal of Marketing Management*, Vol. 12, No. 2, pp. 11-20. ISSN 2141-2421. DOI 10.5897/AJMM2020.0638.

[41] Stockbridge, M., Dorward, A. and Kydd, J. (2003) "Farmer organizations for market access", *Briefing paper presented at Stakeholders Meeting on Farmer Organisations in Malawi,* 18–19 June 2003, Kalikuti Hotel, Lilongwe, Malawi.

[42] Ton, G., Flores, L., Monasterios, R. and Yana, E. (2015) "Capabilities and Performance in Collective Marketing: the importance of learning to cope with agency dilemmas", *Innovative Institutions, Public Policies and Private Strategies for Agro-Enterprise Development,* pp. 113-149. World Scientific. DOI 10.1142/9789814596619_0005.

[43] Untari, D. W. and Vellema, S. (2022) "Are Collective Trading Organisations Necessarily Inclusive of Smallholder Farmers?: A Comparative Analysis of Farmer-led Auctions in the Javanese Chilli Market", *Journal of Agricultural and Environmental Ethics*, Vol. 35, No. 4. E-ISSN 1573-322X, ISSN 1187-7863.

[44] Wang, J. and Wang, Y. (2024) "Economic performance of rural collective-owned cooperatives: Determinants and influence mechanism", *Annals of Public and Cooperative Economics*, Vol. 95, No. 3, pp. 629-653. E-ISSN 1467-8292. ISSN 1370-4788. DOI 10.1111/apce.12454.

[45] Yamane, T. (1973) "*Statistics: An introductory Analysis*", 3rd. ed. Harper & Row. 1130 p. ISBN 9780060473136.

[46] Yanduri, P. and Siddayya (2024) "Understanding Determinants of Managerial Competencies of FPO (Farmer Producer Organization) Leadership for Driving Sustainable Growth: A Strategic Framework and Policy Implications", *Vision: The Journal of Business Perspective*. ISSN 2249-5304. DOI 10.1177/09722629241267161.

Volume XVII Number 4, 2025

Development of a Supply Chain Management Platform for Rubberwood Biomass in Southern Thailand

Sirirat Pungchompoo, Nikorn Sirivongpaisal, Sirirat Suwatcharachaitiwong, Dollaya Buakum

Faculty of Engineering, Prince of Songkla University, Thailand

Abstract

This research aims to manage biomass raw materials in line with industrial needs by developing a platform that links stakeholders in the rubberwood biomass supply chain in southern Thailand. Geographic Information System (GIS) technology was applied to build a database and estimate rubber plantation areas. The trees were grouped by age into three categories: 14–20, 21–27, and over 27 years. The platform also provides information on garden and factory locations, including sawmills, rubberwood processing plants, biomass production plants, and biomass power plants in 14 southern provinces. The system, available on Android and iOS, supports users in making decisions about transportation costs such as distance, time, and fuel. Results from the technology transfer show that the platform is practical, matches user requirements, and is used effectively. The average user satisfaction scores were 4.538 for function and 4.504 for overall use, reflecting the platform's usefulness and acceptance among stakeholders.

Keywords

Rubberwood supply chain, rubberwood biomass, rubberwood biomass platform.

Pungchompoo, S., Sirivongpaisal, N., Suwatcharachaitiwong S. and Buakum, D. (2025) "Development of a Supply Chain Management Platform for Rubberwood Biomass in Southern Thailand", *AGRIS on-line Papers in Economics and Informatics*, Vol. 17, No. 4, pp. 117-135. ISSN 1804-1930. DOI 10.7160/aol.2025.170409.

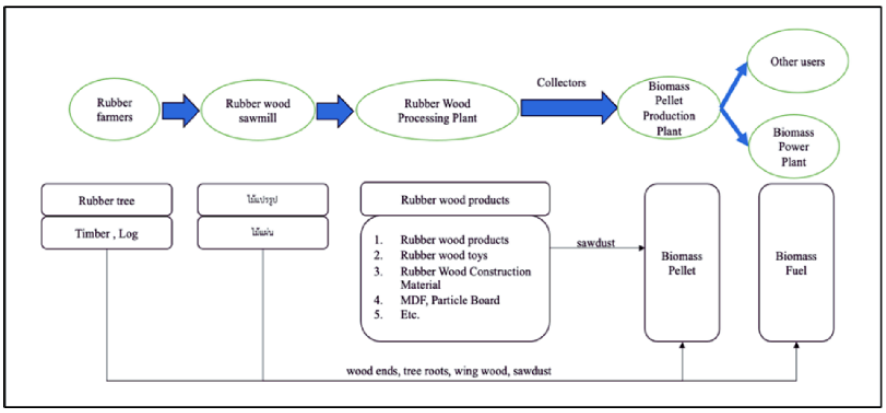
Introduction

Thailand has a policy to promote the production of electricity from renewable energy sources, including the production of electricity from biomass to maximize the use of resources in the country and enhance energy security. Support is provided for the purchase of electricity from private sector renewable energy power plants, both small power producers (SPP) and very small power producers (VSPP). This has led to an increase in the number of small and very small power plant projects selling electricity to the commercial grid, reaching a total of 226 projects in the year 2021 from 1,000 companies. These companies produce electricity from biomass, which is derived from the agricultural process, including oil palm, rice, rubberwood, and fast-growing trees planted specifically as an energy source, such as giant reed, eucalyptus, and Napier grass. Normally, these biofuels are used in various general industries for heat energy production, which may not be sufficient if they are utilized in the energy sector. Therefore, the management of biofuels and their allocation, both for the industrial sector and bioenergy power plants, is a problem that impacts the raw material management for the bioenergy production chain as per the PDP2018 plan. Regarding the electricity system's reliability, for example, in the year 2037, the estimated maximum electricity demand in the southern region is 5,264 megawatts, with an anticipated capacity to generate electricity from power producers totaling 8,662 megawatts. Of this, 29% comes from small, very small, and community power plants operated by private companies. This illustrates that the quantity of biomass available for electricity production is a crucial factor to consider (Ministry of Energy, 2020). At the same time, data on distributed biomass resources is lacking, and stakeholders cannot access timely and connected information, posing challenges in predicting biomass availability for biofuel use, both directly and indirectly, in the rubberwood industry. This includes rubber plantation farmers and rubberwood processing businesses, totaling 290 factories that use rubberwood branches, roots, and sawdust as biofuels for bioenergy power plants. In the year 2020, the total capacity was 196.8 MW (Office of Industrial Economics, 2020). However, bioenergy power producers are not aware of precise information about biomass producers, including estimating production costs, which are affected by transportation costs. Another significant issue is that bioenergy power plants are distributed across various regions and have increased in number in each region, yet there is a lack of regional management, planning, and access to digital resources. For example, the cost of biomass feedstock for small bioenergy power plants is high because they cannot directly access biomass from local producers. Transportation costs have also increased due to sourcing from distant areas. Moreover, the price of biomass feedstock is determined by large-scale biomass producers who serve as both producers, distributors, and transporters of biomass. This results in inadequate supply, price fluctuations, and instability in biomass feedstock prices.

As a result of the aforementioned problems, the objective of this research is to develop a suitable service model for stakeholders involved in community-based bioenergy power plant projects, including farmers, biomass producers, intermediaries in biomass procurement and distribution, power plant operators, and rubberwood processing factories, through a digital platform that utilizes information technology and the internet of things, enabling access to user requirements within the system. This aims to support and facilitate decision-making in the efficient management of domestically produced biomass resources in Thailand, ensuring cost-effective raw materials and efficient transportation services. Additionally, it aims to establish sustainability within the bioenergy supply chain. Therefore, this research aims to establish an intelligent information-sharing service model through a digital platform among stakeholders in the Thai rubberwood industry. It will highlight the key components of the supply chain connecting rubber plantation farmers, rubberwood processing factories, biomass producers, and bioenergy power plants. This will be achieved by implementing a geographic information system that displays location and transportation routes, facilitating calculations for rubberwood resource estimates, distances, and transportation costs concurrently. Therefore, the literature review details are as follows:

Biomass supply chain from para rubber wood

The rubber wood industry is related to farmers who wish to cut down old rubber trees for replanting or rubber trees that have been damaged or are more than 25 years old. They offer rubber wood for sale at the garden location, and buyers from farmers will cut down the old rubber trees after reaching a mutually satisfactory price agreement. In the process of cutting down one acre of rubber trees, only 50% of the logs obtained can be used. Normally, rubber wood that is cut is divided into three parts: roots or tree stumps, wood ends with a central diameter of 3 inches or less, and log sections with a central diameter of 4 inches or more. The log sections are cut to a length of 1.5 meters for delivery to sawmills and furniture factories. The remaining 50% consists of wood scraps, branches, and sawdust, which are considered waste. This waste was not used in the past but is now used in the wood industry, especially for manufacturing MDF (Medium Density Fiber Board) and particle board in the wood processing industry, such as fiberboards and particle boards. Currently, most of the processed wood panels come from the processing of rubber wood because there are not enough other types of log sources in large quantities for industrial use. Additionally, the waste from production, such as wood ends and tree root, comes from rubber tree gardens, wing wood, and sawdust from rubber wood sawmills, and tree branches, as well as toadstool waste from rubber wood furniture factories. This waste is converted into biomass fuel pellets for use in other industries or as biofuel in electricity production plants. Therefore, the above process can be illustrated in the form of a biomass supply chain from rubber wood, as shown in Figure 1 below.



Source: Authors' own work.

Figure 1: Biomass supply chain from para rubber wood.

Renewable energy and biomass potential in Thailand

From the complete report of the Ministry of Energy in the year 2022, which gathered data on energy crops with agricultural practices covering 21.75 million rai, or 49.20 percent of the total area. Key economic crops include rubber and oil palm (Department of Agricultural Economics, 2019). Rubber is of significant economic importance to the southern region, generating income for the local population and providing employment opportunities on a large scale. According to data from the year 2019, the southern region exported rubber with a total value of 130.18 billion Thai Baht, comprising 80.84 billion Thai Baht in natural rubber products and 49.34 billion Thai Baht in rubber products. Therefore, rubber is the most important export commodity in the southern region and has a higher export value than other product categories. (Bank of Thailand, 2563) However, the rubber tree cultivation area in Thailand continuously decreased during the period from the year 2014 to 2018, with a decline of 0.99 percent in the year 2018 compared to the year 2017. In the year 2018, the rubber tree cultivation area was 22,626,277 rai, a decrease of 0.99 percent from the year 2017 when it was 22,852,178 rai. Even though the southern region had the largest rubber tree cultivation area in the country, accounting for over 60 percent, in the year 2018, the rubber tree cultivation area in the southern region was 13,584,115 rai, with a decrease of 1.56 percent compared to the year 2017 when it was 13,800,212 rai. When considering the provinces in the southern region, in the year 2018, the group of provinces on the Gulf of Thailand side, consisting of 5 provinces including Chumphon, Nakhon Si Thammarat, Surat Thani, Phatthalung, and Songkhla, had the largest rubber tree cultivation area, totaling 7,550,875 rai, a decrease of 1.48 percent from the year 2017 when it was 7,664,168 rai. The next group was the provinces on the Andaman Sea side, including Ranong, Phang Nga, Phuket, Krabi, Trang, and Satun, which had a rubber tree cultivation area of 3,425,331 rai in the year 2018, a decrease of 2.41 percent from the year 2017 when it was 3,509,800 rai. The southern border provinces group, consisting of Pattani, Yala, and Narathiwat, had a rubber tree cultivation area of 2,607,912 rai in the year 2018, a decrease of 0.64 percent from the year 2017 when it was 2,624,737 rai. This decrease was due to the impact of drought and low agricultural product prices, which discouraged farmers from rubber and led some of them to cut down rubber trees to plant other crops suitable for the area. Additionally, they received support from the government to reduce cultivation area and switch to other crops. As a result, the rubber tree cultivation area decreased every year, including in the present year. Regarding the rubber tree's biomass, cutting down one rai of rubber trees produces a total biomass of 59 tons per rai, divided into 1) rubber tree logs: 30 tons of biomass, 2) tree branches: 12 tons of biomass, 3) rubber tree branches/stems: 9 tons of biomass, 4) tree roots: 5 tons of biomass, and 5) sawdust: 3 tons of biomass. This biomass from rubber trees has the potential to be used as biofuel for electricity generation, consisting of tree branches, rubber tree branches/stems, sawdust, wood scraps, and tree roots, accounting for 49.20 percent of the biomass from rubber trees per rai. (The Association of Rubber Wood Business, 2015, as cited in Chaiya Kongmani and colleagues, 2017)

Geographic Information Systems and digital platforms

Currently, Geographic Information Systems are widely used in both agriculture and industry. There is ongoing research that applies Geographic Information Systems and technology in various aspects, including production planning, site selection, transportation, and overall cost analysis. For example, a study by Akarariya, Siriwongpaisarn, Kongkaew, and Preechaveerakul (2018) investigated the establishment of collection points to create a balance between areas for oil palm cultivation and palm oil extraction factories. The study focused on logistics for inbound transportation and the current state of palm oil transportation in the southern palm oil industry. It also explored strategies for establishing collection points (known as 'Lan Te') in the southern region to achieve a balance between planting areas and production capacity of palm oil extraction factories, using mathematical models. This included determining the position and quantity for setting up representatives for sub-farmer groups growing oil palm. The study addressed the procurement and distribution of fertilizer to cultivation areas within Surat Thani province under the large-scale agriculture promotion policy. The goal was to minimize the overall cost, including procurement, setup and operation expenses, and delivery costs, within the service area of representatives for sub-farmer groups. These representatives were determined based on sub-farmer groups growing oil palm in each sub-district of Surat Thani province, totaling 121 sub-districts (Siriwongpaisarn and Kongkaew, 2018). Janarthanan et al. (2023) reviewed recent developments in GIS applications, highlighting their role in solving complex socio-economic engineering problems, which supports the integration of GIS into various aspects of agriculture and industry. Mennecke and Crossland (1996) provided an introduction to GIS and a research framework for information systems researchers, emphasizing the importance of GIS in data acquisition, interpretation, and decision support. Li at al. (2024) highlighted the use of GIS for geographic information data elements on the basis of enumerating and summarizing the spatial data infrastructure, which support the construction of digital China and economic and social development.

Furthermore, the application of information and communication technology (ICT) and Internet of Things (IoT) in the biomass supply chain for the development of Industry 4.0, there is the utilization of the ICT BIOCHAIN platform. This platform provides services, including a database and processing capabilities that can link data throughout the system, from the beginning to relevant producers. A survey of business groups in Europe found that 17% utilize the ICT BIOCHAIN for system management, while 15% still adhere to traditional management methods. The remaining percentage employs a mixed approach (Flak, 2020). Similar to the research by (Wajszczuk et al., 2008), the concept of building a biomass transportation network for communities connecting biomass markets with digital platforms for biomass power plants is introduced. The research reveals that challenges in the biomass market arise from the lack of anticipation of biomass demand due to insufficient data from cultivation areas, biomass quantities, data accessibility, and assessment of waste from wood processing or related industries. The cost of transporting biomass for biomass power production is another factor in establishing regional power plants. In this context, digital platforms serve as fundamental tools to connect producers and buyers to address the aforementioned challenges.

The role of GIS in Logistics and Environment also has a great impact. Eldrandaly (2023) explored the integration of GIS with decision support systems to model domain-specific contexts and enhance expert intelligence, which supports the development of digital platforms. Swain (2024) emphasized the role of GIS and remote sensing in revolutionizing logistics and supply chain management, which is essential for optimizing transportation routes and reducing costs. Barnett (2024) discussed the importance of GIS in environmental impact assessment, highlighting its role in providing spatial data, analysis, and visualization for informed decision-making. Atlan (2024) reviewed the significance of database management in GIS, focusing on the organized handling, storage, and retrieval of spatial data to ensure high data quality and accessibility. Goodchild (2022) explored the grand challenges in geographic information science, emphasizing the need for advanced GIS applications to address complex societal issues. Weiner et al. (2002) discussed the integration of community participation with GIS, highlighting its transformative potential in various social and environmental contexts. Sangeetha et al. (2024) reviewed the applications of remote sensing and GIS in precision agriculture, emphasizing their role in optimizing crop production and minimizing environmental impacts.

Platform development by applying GIS and mobile application display

The use of Geographic Information Systems (GIS) in platform development has become crucial in industries such as agriculture, logistics, and transportation. By integrating GIS into mobile applications, businesses can enhance decision-making, improve resource management, and optimize operations. These platforms enable users to estimate plantation areas, optimize transportation routes, and access real-time data via mobile devices.

In agriculture, GIS plays a critical role in estimating plantation areas, especially for crops like palm oil and rubber. For example, Lee et al. (2020) and Li et al. (2020) showed how GIS allows industries to monitor plantation growth in real-time, aiding in land use optimization, production planning, and scheduling. Furthermore, the ability to combine satellite imagery with spatial data makes GIS an essential tool for sustainable land management (Green et al., 2021). Brovelli and Magni (2020) explored open-source mobile GIS solutions, highlighting their versatility across various application fields, which supports the integration of GIS into mobile applications for enhanced decision-making. In addition, GIS is vital in precision agriculture, where analyzing large datasets helps improve land use and sustainability (Jones and Smith, 2019; Wang et al., 2023). Sarbazvatan and Karimi (2023) developed a mobile GIS application called LandInfo for land use and land cover field data collection, demonstrating the practical applications of GIS in real-time agricultural monitoring.

Similarly, GIS is indispensable for optimizing transportation routes in logistics. By factoring in distance, traffic, fuel consumption, and environmental impact, GIS-based platforms help businesses reduce transportation costs. Brown et al. (2019) demonstrated how GIS improves biomass transportation in rural areas, an approach particularly relevant to industries like palm oil and rubber (Kim et al., 2020). Takahashi (2021) introduced Geo-Log Mobile, a mobile GIS application based on a new database framework, which enhances real-time data access and decision-making in agriculture and logistics. Additionally, GIS's real-time capabilities allow businesses to adjust routes dynamically, enhancing delivery efficiency and reducing environmental impact (Miller et al., 2020; Parker & Van Alstyne, 2020). Nitin Liladhar Rane (2023) explores the integration of mobile GIS applications in precision agriculture to enhance crop management practices. The findings indicate that mobile GIS tools significantly improve the efficiency of monitoring and managing agricultural fields, leading to increased crop yields and better resource management.

The growing importance of mobile technology has led to the integration of GIS into mobile applications. These platforms provide real-time travel data, allowing users to optimize routes and estimate travel times and costs (Zhang et al., 2019). Prastacos (1991) discusses the integration of GIS technology in urban transportation planning to address traffic congestion and improve spatial analysis and decision-making, and highlights that GIS significantly enhances the efficiency and effectiveness of urban transportation planning. In agricultural supply chains, mobile GIS applications help transport materials like rubberwood or palm oil to processing facilities. For instance, Li et al. (2020) developed a mobile platform for real-time route optimization, benefiting smallholder farmers by reducing costs and fuel consumption. Giuffrida (2024) presents a GIS-enhanced Real-Time Spatial Delphi approach for spatial participatory planning in urban logistics, and indicate that involving logistics experts in the decision-making process through this participatory method significantly improves the location and efficiency of urban logistics facilities, leading to reduced traffic congestion and environmental impacts. Furthermore, recent advances have integrated predictive analytics into GIS applications. White et al. (2022) explored mobile platforms that predict route changes based on traffic and weather forecasts, enhancing decision-making (Smith et al., 2018). Brown and Affum (2022) developed TRAEMS, a GIS-based environmental modelling system that integrates transport planning outputs with land use information, enabling transportation planners to assess environmental impacts alongside traffic efficiency for more sustainable decision-making. These developments, along with environmental monitoring capabilities (Li et al., 2021; Wang and Zhao, 2023), continue to transform industries that require real-time operational adjustments.

Lu et al. (2024) developed an integrated approach combining Bayesian networks (BNs) and Geographic Information Systems (GIS) to assess flood disaster risks in Yinchuan, China. Their findings demonstrate that the BNs-GIS method effectively evaluates hazard, vulnerability, and exposure, offering a robust framework for comprehensive flood risk management. Chen et al. (2024) investigate the application of GIS in environmental monitoring and risk assessment, with a focus on water, soil, and atmospheric conditions. Their study reveals that integrating GIS technology with remote sensing and spatial data analysis significantly improves the monitoring of environmental changes and risk assessment. This integration leads to more informed decision-making and effective environmental management strategies. The study on GIS-based route optimization for waste management by Swadhin Das, Ankon Baral, Islam M. Rafizul, and Senta Berner found that optimizing waste collection routes using GIS reduced travel distance by 9.40% and fuel costs by 11.6%, significantly enhancing efficiency in Khulna City, Bangladesh.

Materials and methods

This research can be considered in three parts, namely Part 1: Study and data collection of biomass producers, Part 2: Data collection, preparation of data for the database, and study of the cost structure of biomass fuel transportation using GIS, and Part 3: Development of a biomass management platform with innovations using information technology and digital technology. The details are as follows:

Study and data collection of biomass producers

This study covered 14 provinces in the southern region of Thailand, including Chumphon, Ranong, Phang Nga, Phuket, Surat Thani, Nakhon Si

Thammarat, Krabi, Trang, Phatthalung, Songkhla, Pattani, Yala, Narathiwat, and Satun. The main purpose was to identify problems and define the requirements of the rubber wood biomass supply chain. This included cultivation areas, tree age, wood quantity, processing factories, middlemen or collectors, and regulations for establishing community biomass power plants. Another objective was to collect information on the demand and supply of biomass from industries and biomass power plants in the region, to support decisions on the size and location of new community power plants.

The main dataset consisted of plantation areas, rubber age, and production volume, together with information on four key stakeholder groups: sawmills, rubber wood processing factories, biomass production plants, and biomass power plants. Data for each group included their locations, type of operation, contact information, and production capacity. In addition, information on collectors, collection points, transportation methods, and structural factors such as regulations, distance from communities, road networks, and resource availability were also gathered. Secondary data were mainly taken from the Department of Industrial Works and the Ministry of Energy between 2020 and 2022, together with information from agricultural agencies, rubber associations, and satellite imagery. The preliminary verification was carried out using Google Maps, Street View, and official documents, while primary data was collected from field surveys. The survey included GPS-based location recording, photographs of sites, and interviews with stakeholders. The initial results showed that there were 450 sawmills, 30 rubber wood processing factories, 39 biomass production plants, and 82 biomass power plants across the 14 provinces.

To accuracy, the data were cross-checked with multiple sources, such as satellite images, orthophotos, Google Earth, and field observations. Records that were duplicated or incomplete were removed, and all updates were recorded in a data log. The database was developed in two forms: spatial data (points stored in GIS layers, separated by business type) and attribute data (tables including id, factory name, address, type, contact number, and capacity). All data were processed in ArcGIS, where spatial and attribute data were linked, and maps were generated with symbols and colors to distinguish business types. Finally, the database was exported as Shapefile and GeoJSON/JSON files for connection with a digital platform. Quality control was performed by random checking at least 10% of records in each group with field data and satellite images, which helped to improve data reliability for future spatial analysis and decision-making.

Data collection, preparation for database, and study of biomass fuel transportation cost models using GIS

This part entails studying and collecting topographic data in the 14 provinces of the southern region, such as cultivation areas (referenced from the Ministry of Agriculture and Cooperatives and the Thai Rubber Association), production, orchard age, the number and locations of collectors, transportation patterns, the number and locations of biomass production factories, and the number and locations of biomass power plants. Geographic Information System (GIS) is utilized as a tool to display cultivation areas, collector locations, transportation patterns, factory locations, and power plant locations. Field surveys are conducted, and samples are collected from relevant parts of the supply chain to verify the data. The methodological procedures are explained step by step in the following subsections.

Data and tools

The satellite images used in this study consisted of Landsat 8 OLI/TIRS (2013–2015) covering Path 127–130 and Row 51–56 (13 scenes), Landsat 5 TM (2005–2007) for time series analysis, high-resolution Google Earth imagery (2008–2015), and orthophoto maps at a scale of 1:4,000 (2002 and 2010). Institutional and industrial data were obtained from the Ministry of Agriculture and Cooperatives, the Thai Rubber Association, the Department of Industrial Works, and the Ministry of Energy. Data processing relied on ERDAS Imagine for image processing, ArcGIS for spatial analysis, and a high-performance computer system for large-scale data handling.

Image pre-processing

The selection of satellite imagery was carefully controlled to avoid the leaf-shedding period of rubber trees, which occurs around February in the upper South and March in the lower South. Images with the least cloud cover were prioritized, and when necessary, supplementary images from nearby years were included. For provinces that required mosaicking of multiple scenes, images from the same season or acquisition date were selected to minimize differences. Geometric correction of Landsat 5 TM data was performed using topographic maps at 1:50,000 scale (UTM/WGS84) with an image-to-map approach and evenly distributed ground control points (GCPs). Mosaicking and histogram matching were applied to normalize tonal differences across scenes. Image enhancement techniques, such as brightness–contrast adjustment and radiometric enhancement (equal percentage and Gaussian methods), were also applied to improve visual interpretability according to land characteristics.

Bands and composites

Spectral analysis employed bands 1–7 of Landsat 8. Bands 5, 6, and 7 provided the best tonal contrast for land cover differentiation. Composite images were generated to support visual interpretation and training data identification. The false color composite (5-6-3, RGB) was effective in highlighting vegetation, where young rubber appeared orange, mature rubber pink, and oil palm dark orange. The natural color composite (6-5-4, RGB) and the true color composite (4-3-2, RGB) were also applied to support discrimination of rubber plantations from other land covers.

Area of Interest (AOI) and training areas

Training areas were identified to cover at least 30 plots, with a minimum of 10 plots for each age class, and each plot larger than 100 rai to ensure homogeneity. Four categories were defined: rubber aged 7–14 years, 14–20 years, over 20 years, and other land covers such as oil palm, forests, agricultural preparation areas, wetlands, mangroves, water bodies, settlements, and cloud shadows. In provinces where small rubber plots (<100 rai) were predominant, additional training areas were selected to compensate for size limitations.

Supervised, pixel-based classification

The classification process used the Maximum Likelihood algorithm with several combinations of spectral bands and training areas tested to obtain the best results. Class separability was evaluated using Transformed Divergence (TD), ranging from 0 to 2000. A TD value above 1900 indicated clear separation, values between 1700 and 1900 showed good separation, while values below 1700 suggested overlapping classes. The results showed that rubber plantations could be clearly distinguished from other land uses. However, some overlap occurred among the three rubber age groups, especially between the 14–20 years category and adjacent classes, due to the similarity in spectral signatures of rubber leaves.

Field survey and accuracy assessment

Field surveys were conducted after the leaf-shedding period and before the rainy season. Survey points were distributed across the 14 provinces, where GPS coordinates, photos of tapping marks, and environmental context were recorded. Interviews with landowners were also carried out to verify plantation age. Classification accuracy was evaluated using an error matrix and overall accuracy was calculated based on 477 reference points. The classification was accepted only when overall accuracy reached at least 85%.

Post-processing and GIS integration

The classification results in raster format were converted into vector polygons. Small polygons (<3 rai) caused by misclassification were eliminated or merged, and boundaries were smoothed to better match real-world patterns. Polygons were further converted into points to facilitate system integration, and the data were exported into JSON format for use in digital platforms.

Facility database for four stakeholder groups

The database was also developed for four stakeholder groups: sawmills, rubber wood processing factories, biomass production plants, and biomass power plants. The process began with filtering facility lists from the Department of Industrial Works and the Ministry of Energy to remove duplicate or incomplete records. Facility status and locations were verified using Google Maps, Street View, and supporting documents. Geocoding was performed to convert addresses into geographic coordinates, and when possible, field surveys with GPS were conducted for higher precision. All locations were standardized to WGS84/UTM and imported into ArcGIS to create point layers by facility type. Attribute tables including id, facility name, address, type, telephone, and production capacity were linked to spatial data. Symbolization by color and shape was applied to represent different facility types, and summary maps were generated to show their distribution across the 14 provinces. The final database was exported as Shapefile, GeoJSON, and JSON formats for platform integration. Initial results showed 450 sawmills, 30 rubber wood processing factories, 39 biomass production plants, and 82 biomass power plants.

Quality control procedures included random checks of at least 10% of records in each group using base maps, satellite images, and field surveys. Facility types were cross-verified against official definitions, and missing information such as contact details and production capacities was updated from government sources or direct communication. All modifications were recorded in a data log to ensure transparency. As a methodological note, secondary data may be subject to changes such as facility closures, relocations, or ownership transfer; therefore, database updates are recommended at least once a year. Ethical considerations regarding the publication of commercial contact details were also taken into account.

Spatial–attribute update for 2020/2563

To update the database to 2020, spatial boundaries from 2015 were maintained, while attribute data were recalculated. The area of each polygon was expressed as a proportion of the total district area in 2015, then multiplied by the updated district-level plantation area in 2020. This method allowed estimation of new plantation areas while reducing time and cost compared with producing a completely new land cover map.

Age-class re-class

The rubber age classes derived from 2015 classification (7–14, 14–20, and >20 years) were adjusted by adding five years to match the 2020 reference year. The updated classes were therefore defined as 12–19, 19–25, and >25 years. The recalibrated results were compared with official data from the Rubber Authority of Thailand at the district level to ensure consistency and reliability before integration into the platform.

Transport cost modeling – outline

To connect the database with logistics decision-making, a transport cost framework was designed. The analysis considered road networks, travel speeds, and the relationship between supply points (plantations, collectors, sawmills, and processing factories) and demand points (biomass plants and power plants). Network analysis was applied using shortest path, least-cost path, and cost-distance approaches. Cost parameters included transport cost per kilometer per ton, vehicle type, load capacity, loading/unloading costs, and travel time constraints. The expected outputs were minimum distance, time, and cost per route, the supply areas that each plant can serve, and efficient allocation of transportation routes. Numerical results of the cost model are presented in the results section.

Development of a biomass management platform with innovations using information technology and digital technology

The biomass management platform was developed by the research team in collaboration with key stakeholders in the rubber wood biomass supply chain. Its aim was to increase business value and improve the efficiency of information exchange between producers and consumers. The platform was also designed to support the sustainable use of biomass resources through digital connection, data visualization, and system performance evaluation. The development process was divided into five steps, as described below.

Design and development of the prototype

The development began with designing a prototype using Google Maps to display biomass sources from rubber wood. The prototype played a critical role in testing the main concept, collecting user feedback, and reducing risks before creating the full version. The design emphasized simplicity and usability, with appropriate use of colors, fonts, icons, text, and responsive buttons that allowed users to interact quickly and easily.

Application development on Android and iOS

A mobile application prototype was created for both Android and iOS platforms. Input was drawn from focus group discussions with stakeholders such as rubber plantation owners, sawmills, processing factories, biomass producers, and biomass power plants. The application included several functions: system login, display of plantation data by age and size within a 5–10 km radius, summary of plantation data by province and district, connection to Google Maps for navigation, and visualization of four key stakeholder groups.

Route and transport cost analysis

The application was also designed to calculate transport routes and costs. Users could select their own starting and ending points, either from plantations or facilities, to identify the lowest-cost route for buying, selling, or combined transactions. This function was intended to improve decision-making in biomass supply chain logistics.

System testing and evaluation

The prototype was tested using the Black Box Testing method to evaluate efficiency, accuracy, and user satisfaction. A questionnaire was designed with two main sections: efficiency and functional requirements, and usability. The questionnaire was validated by five experts through content validity checks (IOC $\geq$ 0.50) and revised based on their suggestions. Reliability testing was conducted using Cronbach's alpha with a pre-test on a small group before actual field use.

Stakeholder participation

Hands-on workshops were organized to present and test the platform with stakeholders. Participants were invited to use the application directly, explore key functions (such as searching and filtering plantations and facilities, summarizing data by province and district, connecting to Google Maps, and calculating transport routes), and then provide feedback through the validated questionnaire. Their comments and suggestions were used to improve the data structure, user interface (UI), and calculation modules of the platform before its wider deployment.

Results and discussion

The preliminary results of the platform development research, conducted under the specified research methodology, are as follows. Initially, the results include the development of a biomass database from rubber trees and databases for sawmills, rubber wood processing plants, biomass production plants from rubber trees, and bioenergy power plants. Subsequently, a prototype biomass platform was designed and developed using data from rubber trees and various plant databases for bioenergy power plants. Furthermore, a satisfaction evaluation was conducted among users who accessed and used the prototype platform application on mobile or tablet devices. Therefore, all results are presented as follows.

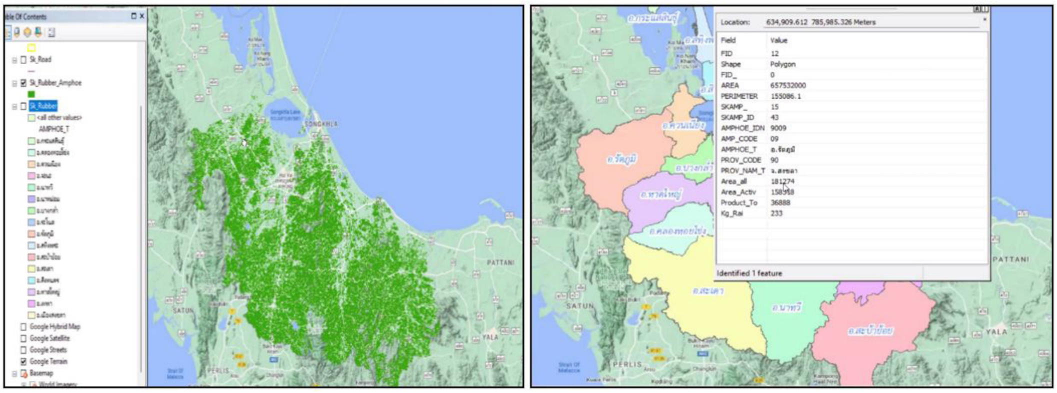
Study, collection, and data compilation for establishing a rubber plantation area database and assessing the age of rubber trees using geographic information system (GIS) have the following details:

The survey data from the Thai Rubber Authority in 2020 reveals changes in rubber plantation areas in the southern region over the past five years. Some districts have seen an increase, while others have experienced a decrease. Consequently, there is a need to update rubber plantation area data for each district to ensure its accuracy. The new area is calculated using descriptive data from the GIS database. However, the spatial data from the original rubber plantation areas in 2015 is retained due to the time and cost associated with updating it using satellite imagery. The area (in hectares) of each polygon for the 2015 rubber plantation areas is calculated as a percentage of the total planting area per district. This percentage is then applied to the total planting area for 2020, derived by multiplying the percentage from 2015 by the total planting area for each district in 2020. Subsequently, the age range of the new rubber trees is calculated by adding five years to the latest satellite imagery data in 2018. Age groups are divided into three ranges: 1) 7-14 years, 2) 14-20 years, and 3) >20 years. The age range of the new rubber trees is calculated in the GIS database as attributes, resulting in three age ranges: 1) 12-19 years, 2) 19-25 years, and 3) >25 years. The testing results are presented in Table 1, and this data is utilized to develop a database specifying rubber plantation areas and the age of rubber trees. Figures 2 to 4 depict this information, which is crucial for the development of a platform managing the rubber biomass supply chain in the next phase.

District	Data from the Survey Year: 2015	Age of Rubber Trees >27	Age of Rubber Trees 21-27	Age of Rubber Trees 14-21	Unclassified	Data from Estimation Year: 2020	RAOT Data Year: 2020
Krasae Sin	6,962.70	712.17	3,192.97	1,846.16	4,017.70	9,769.00	9,765.00
Khlong Hoi Khong	55,531.18	46,235.17	25,644.02	32,308.68	2,440.05	106,627.92	106,628.00
Khuan Niang	32,671.86	13,610.63	12,799.82	13,501.00	10,476.55	50,388.00	50,388.00
Chana	198,535.60	90,908.63	30,933.56	52,946.34	49,094.43	223,882.96	223,883.00
Thepha	213,643.85	172,441.80	39,108.15	46,065.63	6,355.39	263,970.97	263,971.00
Na Thawi	366,856.35	184,792.22	29,080.14	92,352.89	11,001.18	317,226.43	317,227.00
Na Mom	39,678.02	14,749.22	9,256.92	16,551.88	19,275.91	59,833.93	59,834.00
Bang Klam	27,030.63	16,291.11	6,781.16	9,659.55	1,976.22	34,708.04	34,798.00
Mueang Songkhla	32,208.75	15,235.41	4,134.40	5,022.95	8,048.26	32,441.02	32,441.00
Rattaphum	169,361.04	36,294.77	51,596.44	44,251.94	49,130.42	181,273.57	181,274.00
Sadao	380,919.13	279,264.37	73,229.59	74,657.36	8,809.20	435,960.52	435,961.00
Saba Yoi	269,963.30	160,603.70	39,059.55	76,851.71	22,165.19	298,680.15	298,680.00
Hat Yai	204,380.81	72,220.77	44,831.01	45,856.73	45,795.85	208,704.36	208,713.00
Ranot	-	-	-	-	-	-	30.00
Sathing Phra	-	-	-	-	-	-	172.00
Singhana-kron	-	-	-	-	-	-	491.00

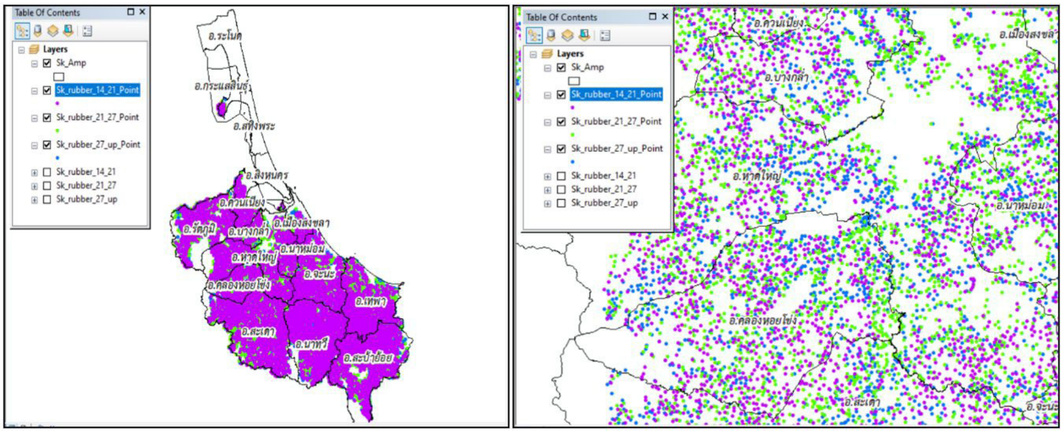
Source: Rubber Authority of Thailand (RAOT), 2020

Table 1: Estimation of rubber plantation area in 2020 based on 2015 data compared with Rubber Authority of Thailand (RAOT). Data for 2020.



Source: Authors' own work

Figure 2: Example data from the assessment of rubber plantation area and age of new rubber trees.



Source: Authors' own work

Figure 3: Spatial database in the Geographic Information System (GIS) showing rubber plantation areas in Songkhla Province.

Sk_rubber_14_21_Point

Shape	OBJECTID	NAME_T	PV_TN	AP_TN	Rai_63	Class_New	ORIG_FID	POINT_X	POINT_Y
Point	15295	ยางพารา	สงขลา	สะบ้าย้อย	53.82214	14 - 21 ปี	18	100.903918	6.353916
Point	15296	ยางพารา	สงขลา	สะบ้าย้อย	23.2373	14 - 21 ปี	19	100.89357	6.354916
Point	15297	ยางพารา	สงขลา	สะบ้าย้อย	18.69737	14 - 21 ปี	20	100.919022	6.354417
Point	15298	ยางพารา	สงขลา	สะบ้าย้อย	38.23104	14 - 21 ปี	21	100.889266	6.354899
Point	15299	ยางพารา	สงขลา	สะบ้าย้อย	26.91107	14 - 21 ปี	22	100.908081	6.355608
Point	15300	ยางพารา	สงขลา	สะบ้าย้อย	206.35801	14 - 21 ปี	23	100.916103	6.353126
Point	15301	ยางพารา	สงขลา	สะบ้าย้อย	46.83302	14 - 21 ปี	24	100.910329	6.357444
Point	15302	ยางพารา	สงขลา	สะบ้าย้อย	12.0368	14 - 21 ปี	25	100.904026	6.359826
Point	15303	ยางพารา	สงขลา	สะบ้าย้อย	119.41226	14 - 21 ปี	26	100.895938	6.358817
Point	15304	ยางพารา	สงขลา	สะบ้าย้อย	72.69871	14 - 21 ปี	27	100.923122	6.361924

Source: Authors' own work

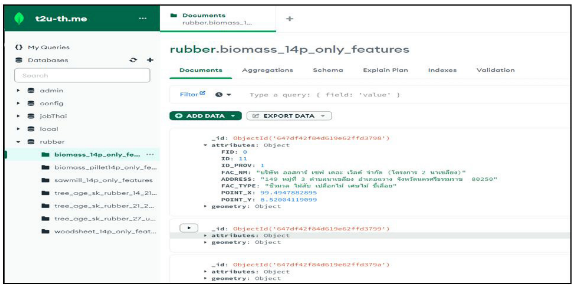
Figure 4: Descriptive database in the Geographic Information System (GIS) showing rubber plantation areas in Songkhla Province

Establishment of a database for wood sawmills, rubber wood processing plants, biomass production facilities from rubber wood, and biomass power plants

For data collection to establish a database for the four user groups in the biomass industry, namely wood sawmills, rubber wood processing plants, biomass production facilities from rubber wood, and biomass power plants, preliminary data has been gathered by obtaining geographical information from the Department of Factories and the Ministry of Energy. This data is then cross verified against business operations and locations using Google Maps. Upon examination, it was found that there is up-to-date information in the southern region across all 14 provinces. Specifically, the wood sawmill group has 450 factories, rubber wood processing plants have 30, biomass production facilities from rubber wood have 39, and biomass power plants have 82. The data is then visualized using coordinates obtained through the Global Positioning System (GPS) and imported into the ArcGIS program to create a comprehensive database. This allows the connection of attribute data for each group of factories with spatial data, and the presentation of this information is depicted on the map, as shown in Figure 5.

Design and development of a prototype for the platform

The Prototype, functioning as a platform displaying biomass data from rubber wood on Google Maps, is a crucial tool to support the development of Thailand's bio-based economy. It facilitates easier access for users to information about biomass from rubber wood, leading to more efficient and sustainable utilization of this resource. This software will be the initial version with tested functionalities, and the prototype design is a key step in the software development process. It allows the development team to test software concepts and address any shortcomings before proceeding with the development of the full version. In this process, various program components are created to build the prototype platform, showcasing the database system and data on the mobile application screen, as illustrated in Figures 6 and 7.

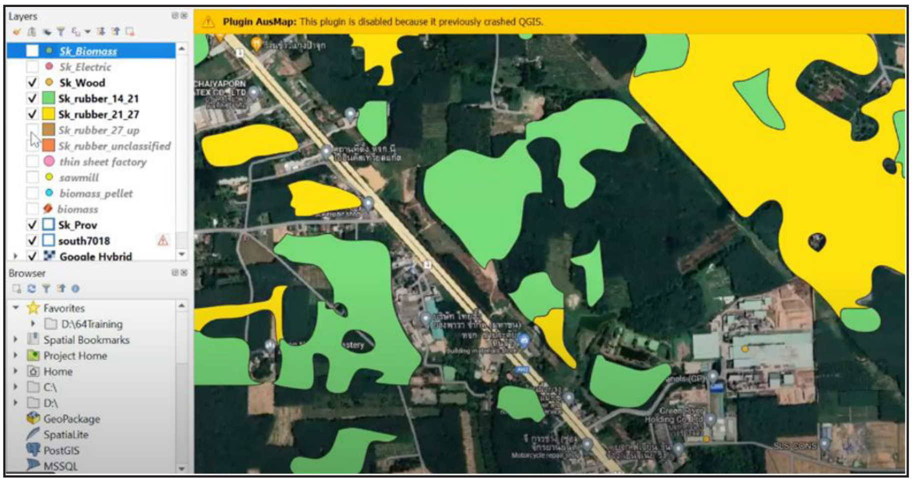


Source: Authors' own work

Figure 6: Database of the biomass platform from rubber wood.

Source: Authors' own work

Figure 7: Example of program code for the map generation to display data.



Source: Authors' own work

Figure 5: Example data and database development for wood sawmills, rubber wood processing plants, biomass production facilities from rubber wood, and biomass power plants.

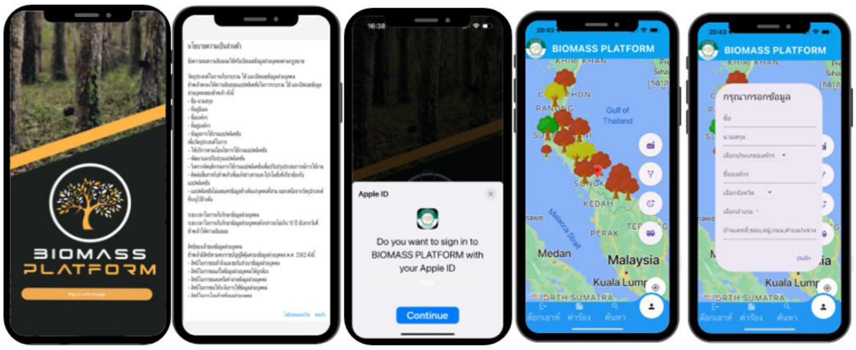
The data is displayed through the application on both Android and iOS operating systems. The operational results, as well as details regarding the application's design and functionality, are outlined as follows:

1. Initial access to the application during system. Testing Users can access the application on both platforms, as illustrated in Figure 8. Next, the initial screen displayed when users access the application on both Android and iOS platforms in Figure 9.

Source: Authors' own work

Figure 8: The logo representing the application, which is available on both Android and iOS operating systems.to display data.

For the user interface through the application, users can search for information on the types of industries they are interested in, specifically rubber plantations. These can be categorized into three sizes: 1) Small rubber plantations (rubber trees with a size of 1-50 rai), 2) Medium-sized rubber plantations (rubber trees with a size of 51-100 rai), and 3) Large rubber plantations (rubber trees with a size larger than 100 rai). Additionally, the system can categorize rubber plantations based on the age of the rubber trees, as illustrated in Figure 10. The application displays rubber plantations within predefined distances, initially set at 5 kilometers and 10 kilometers. It also estimates the transportation distances along actual transportation routes in Figure 11.

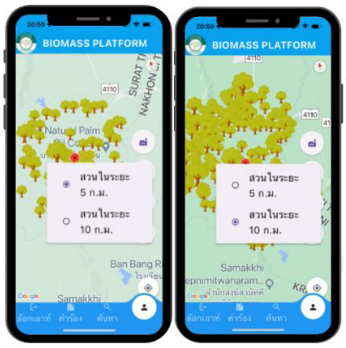


Source: Authors' own work

Figure 9: The initial screen displayed when users access the application.

Source: Authors' own work

Figure 10: Display through the application - information on the interested industry types, rubber plantations, and the age categories of rubber trees.



Source: Authors' own work

Figure 11: Display through the application - displaying rubber plantations within the defined distance and estimating transportation distances.

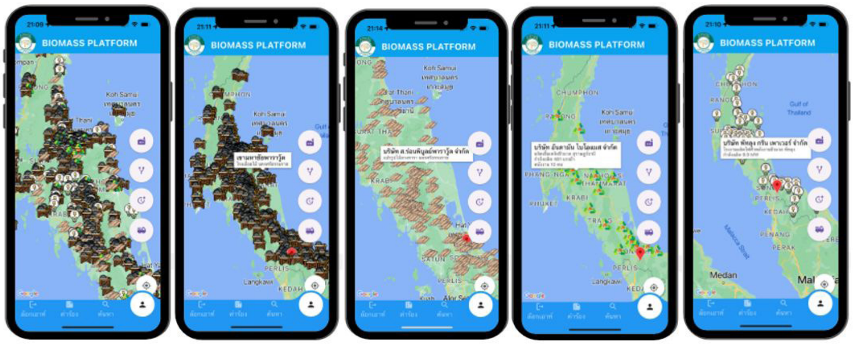
Moreover, data display of the four key business groups related to the rubberwood supply chain: This includes sawmills, rubberwood processing plants, biomass pellet production plants, and biomass power plants. The application is designed to display information specifically for sawmills and rubberwood processing plants, as illustrated in Figure 12.

Estimating transportation routes and optimal costs for managing rubberwood biomass fuel

In addition to displaying basic data based on the creation of the database, route identification, and cost estimation for transportation in accordance with the research project's objectives, the application should also allow users to estimate their own transportation routes. This feature would enable users to connect routes from rubber plantations and the four key business groups, allowing them to set their own starting and ending points without restriction. This capability can be used for purchasing, selling, or both simultaneously, with the goal of selecting the route that minimizes transportation costs for each trip. This is illustrated in Figure 13, which shows the estimated transportation routes and costs for managing rubberwood biomass fuel between rubber plantations and related four industrial groups.

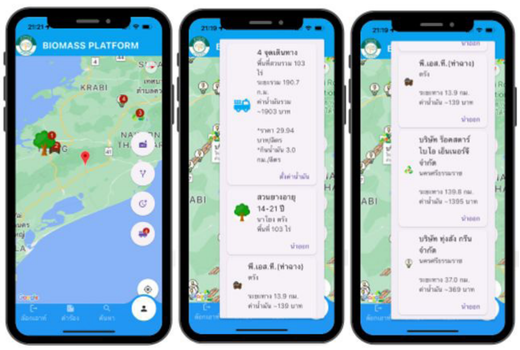
Evaluation of the prototype platform for rubberwood biomass supply chain management

In the system development phase (Construction), after the programmers completed the coding process based on the design, the next step involved testing the program to ensure its correctness according to the specified requirements. Additionally, user manuals were created, and relevant personnel were trained to understand and efficiently use the system. The most used testing method for such systems is Black Box Testing, which focuses on evaluating the input and output data without assessing the internal processing mechanisms. Before the system was launched, the research team performed Black Box Testing to evaluate the system's efficiency, functionality, and usability. Specifically, the system's performance was evaluated in three areas: 1) Efficiency, to determine how effectively the system operates; and 3) Usability, to assess user satisfaction and identify any system errors.



Source: Authors' own work

Figure 12: The data display of the four key business groups related to the rubberwood supply chain.



Source: Authors' own work

Figure 13: Estimating transportation routes and costs for managing rubberwood biomass fuel between rubber plantations and industrial groups.

The evaluation was conducted using a survey, which was designed and validated by five experts who examined the Index of Objective Congruence (IOC). Only questions with an IOC score above 0.50 were included, and the language and phrasing were reviewed to ensure the questionnaire was comprehensive. After revisions, a pre-test was conducted with a sample of 10 individuals, including rubber plantation owners, sawmill operators, rubberwood processing industry representatives, biomass factory workers, and biomass power plant personnel. The reliability coefficient (Cronbach's alpha) was 0.970, indicating that the questionnaire was sufficiently reliable for data collection.

Subsequently, a stakeholder feedback session was organized, and the technology was transferred through a workshop held on Saturday, December 16, 2023. The mobile application platform, which presents a database of stakeholders in the rubberwood biomass supply chain, was demonstrated and evaluated by 30 participants from the southern region of Thailand. The evaluation results are presented in Tables 2 and 3, and further details are as follows below.

Category	Participant	Percentage (%)
1. Gender		
- Male	20	66.67
- Female	10	33.33
2. Age Group		
- Less than 18 years	0	0
- 20–25 years	0	0
- 25–30 years	5	16.67
- 30–45 years	19	63.33
- 45–50 years	2	6.67
- 50–60 years	2	6.67
- Over 60 years	2	6.67
3. Education Level		
- No formal education	0	0
- High school education	0	0
- Primary education	0	0
- Vocational education	0	0
- Bachelor's degree	26	86.67
- Postgraduate degree	4	13.33
4. Involvement in Industry		
- Rubber plantation	2	6.67
- Middleman for rubberwood trading	2	6.67
- Rubberwood factory/sawmill	8	26.67
- Rubberwood processing plant	7	23.33
- Biomass production facility	2	6.67
- Biomass power plant	8	26.67
- Government agency/academic	1	6.67
- Other	0	0
5. Operating System Used*		
- Android	17	47.22
- iOS	19	52.78
6. Recommendation to Others		
- Would recommend	29	96.67
- Would not recommend	0	0
- Not sure	1	3.33

Note: *Participants were allowed to select more than one operating system.
Source: Authors' survey (2024)

Table 2: Demographic information of respondents (n = 30).

Factor	Average	Standard Deviation	Importance Level
1. Efficiency and Functional Requirements			
- The application has reliable security features	4.733	0.583	High
- User authentication is required every time before use	4.533	0.73	High
- The application is accurate and reliable	4.367	0.615	High
- Users can access the application anytime, anywhere	4.833	0.461	High
- Data processing is accurate and reliable	4.033	0.669	High
- The application provides detailed, comprehensive data	4.5	0.572	High
- The services meet user requirements	4.5	0.572	High
- The application is beneficial for work and analysis	4.8	0.484	High
2. Usability Test			
- The application is clear, modern, and easy to use	4.033	0.765	High
- Icons are understandable and appropriate	4.367	0.809	High
- The position, shape, and color of icons are suitable	4.167	0.699	High
- Font size and color are clear in both orientations	4.333	0.661	High
- The application is appropriate for mobile/tablet use	4.267	0.583	High
- The application responds quickly and efficiently	4.067	0.828	High
- Displayed data is detailed and complete	4.167	0.747	High
- The application is stable	4.1	0.548	High
- No errors were encountered during use	4.233	0.504	High

Source: Authors' survey (2024)

Table 3: Factors influencing the platform for rubberwood biomass supply chain management.

In evaluating the platform's performance regarding Efficiency and Functional Requirements, a survey of 30 users involved in both rubberwood biomass production and rubberwood processing indicated high overall satisfaction with the platform on both Android and iOS operating systems, with an average score of 4.538. At the factor level, the platform was rated highly for having a reliable security system for access, receiving average satisfaction scores of 4.733 and 4.533 on Android and iOS, respectively. Additionally, users expressed a high level of satisfaction with the platform's accessibility, as it can be used anytime and anywhere, achieving an average score of 4.833. The accuracy and reliability of the data and processing results were also positively evaluated, with average scores of 4.033 and 4.500. Moreover, users expressed high satisfaction with the platform's detailed and comprehensive database, which effectively met their requirements, earning an average score of 4.500. The platform was also found to be highly beneficial for facilitating users' work and analysis, with an impressive average score of 4.800.

In terms of usability, the platform's overall performance on both Android and iOS was rated very highly, with an average score of 4.504. When evaluating satisfaction at the factor level, it was found that the application was clear, easy to understand, modern, and user-friendly, earning a satisfaction score of 4.033. The icons displayed within the application were deemed easily understandable and appropriate, with an average score of 4.367. The positioning, shape, size, and color of the icons representing the data were also rated as suitable and accurate, with an average score of 4.167. The font size and color on the display were rated highly for clarity, both in horizontal and vertical orientations when adjusting the mobile device or tablet's direction, achieving an average score of 4.333. Additionally, the application was considered highly suitable for mobile or tablet use, with an average score of 4.267. Users also praised the platform's responsiveness, citing its convenience and speed, with an average score of 4.067. For data display, the information provided by the application was rated as detailed, complete, and stable, with average scores of 4.167 and 4.100, respectively. Finally, users reported no errors during the use of the application on either operating system, with an average satisfaction score of 4.233.

The development of the platform for managing the rubberwood biomass supply chain has shown promising results, especially in terms of its ability to gather and present important data for different stakeholders. The use of Geographic Information System (GIS) technology allowed for the accurate collection of data about rubber plantation areas and the age of rubber trees. This is crucial for the efficient management of biomass resources, as it helps users understand the distribution of biomass potential in the region. By accurately calculating plantation areas and tree ages, the platform supports better planning for rubberwood harvesting and biomass utilization. Moreover, the creation of databases for sawmills, rubberwood processing plants, biomass production facilities, and biomass power plants ensures that all important players in the supply chain are connected. Users can access this information through the platform, which improves communication and coordination among businesses. This is particularly useful for industries that rely on biomass resources for energy production, as it allows them to better manage their operations and make informed decisions about sourcing and transportation.

The results of the user satisfaction survey indicate that the platform is highly effective in terms of both efficiency and usability. Users found the platform to be secure, with a reliable authentication system, and were satisfied with how quickly they could access data. The ability to use the platform on both Android and iOS operating systems added to its accessibility, making it more convenient for a wide range of users. The platform's ability to display detailed and accurate data, especially concerning the biomass supply chain, was also rated highly by users. This suggests that the platform meets the needs of its users by providing them with relevant and easy-to-understand information. Additionally, the usability test showed that users found the platform to be modern, clear, and easy to navigate. The icons and user interface were well-designed, making it simple for users to find the information they needed. The application's ability to adjust to both horizontal and vertical orientations on mobile devices further improved the user experience, as it allowed for more flexible use in different settings.

Overall, the platform succeeded in addressing the main challenges faced by stakeholders in the rubberwood biomass industry. By offering a tool that integrates data from multiple sources and allows for accurate transportation route and cost estimation, the platform enhances the efficiency of biomass management. The high user satisfaction scores across various aspects of the platform —security, accuracy, and usability—highlight its effectiveness in meeting the goals of the research project.

The limitation of this research lies in its geographic scope, as the platform was developed specifically for rubberwood biomass supply chains in only 14 provinces of Southern Thailand. Consequently, the platform's findings and functionalities may not be directly applicable to other regions or nationwide without substantial adaptation. Furthermore, the data utilized such as plantation areas and tree ages were based on secondary sources, which may not accurately reflect real-time conditions or recent changes in biomass availability. Therefore, future research should focus on expanding the platform's geographic scope to cover the entire country. This would involve the collection of primary, real-time data to ensure more accurate and current information on plantation areas, biomass availability, and tree age. Moreover, the platform could be tailored to specific regions by factoring in local conditions, such as varying climate patterns, transportation infrastructure, and regional industry requirements. Such an approach would broaden the platform's utility, making it more effective for nationwide biomass management and strategic decision-making.

Conclusion

The study aimed to plan and manage biomass fuel resources in alignment with the needs of industries and biomass power plants. Initially, data collection was conducted through focus group discussions with rubber plantation owners and factory representatives to identify their requirements for a platform to manage rubberwood biomass. The discussions revealed that platform users required quantitative data, such as information on rubber plantation areas, the location of plantations, and related factories for the rubberwood biomass supply chain across 14 southern provinces. Additionally, the need for linking this information to support users in estimating biomass volumes and planning biomass collection, including managing transportation costs such as distance, time, and fuel expenses, was highlighted.

As a result, this data was used as a guideline for creating a biomass database for stakeholders in the rubberwood biomass supply chain, utilizing Geographic Information Systems (GIS) to represent plantation areas and locations. The rubber plantation database was updated using 2015 GIS data combined with secondary data from 2020. The updated plantation area and rubber tree age ranges were calculated, dividing the tree age into three groups: 1) 14-20 years, 2) 21-27 years, and 3) over 27 years. Additionally, databases were established for four main factory groups within the biomass industry: sawmills, rubberwood processing plants, biomass production facilities, and biomass power plants, covering all 14 southern provinces. The data, referenced from 2022-2023, served as the basis for developing a prototype platform capable of functioning on both Android and iOS operating systems as a mobile application.

Finally, technology transfer activities, such as workshops and focus group discussions, were conducted to verify the database and evaluate the platform's performance in terms of both efficiency and functional requirements, as well as usability on both operating systems. The evaluation results for both aspects showed a very high level of performance, with average scores of 4.538 and 4.504, respectively.

Corresponding author:
Sirirat Pungchompoo
Department of Industrial and Manufacturing Engineering, Faculty of Engineering
Prince of Songkla University, 15 Karnchanavanich Road, Hat Yai, Songkhla 90110, Thailand
E-mail: sirirat.pu@psu.ac.th

References

[1] Akarariya, K., Siriwongpaisarn, N., Kongkaew, V. and Preechaveerakul, S. (2018) "A study on the establishment of collection points to balance oil palm plantation areas and palm oil extraction factories", Paper presentation, *Annual Industrial Engineering Network Conference*, Ubon Ratchathani, Thailand.

[2] Atlan, T. (2024) "*What is Database Management in GIS | Atlan*", Atlan. [Online]. Available: https://atlan.com/know/glossary/database-managment-in-GIS/ [Accessed: June 15, 2025].

[3] Barnett, K. (2024) "*Why is GIS important in environmental impact assessment?*", Geographic FAQ Hub: Answers to Your Global Questions. [Online]. Available: https://www.ncesc.com/geographic-faq/why-is-gis-important-in-environmental-impact-assessment/ [Accessed: June 15, 2025].

[4] Brovelli, M. A., Magni, D. (2007) "*Open source mobile GIS solutions for different application fields*", Politecnico Di Milano. [Online]. Available: https://www.isprs.org/proceedings/xxxvi/5-c55/papers/magni_diego.pdf [Accessed: May 12, 2025].

[5] Brown, A. and Affum, J. (2002) "A GIS-based environmental modelling system for transportation planners", *Computers Environment and Urban Systems*, Vol. 26, No. 6, pp. 577-590. ISSN 1873-7587. DOI 10.1016/s0198-9715(01)00016-3.

[6] Brown, R., Davis, T. and White, M. (2019) "Route optimization using GIS for transportation logistics", *Journal of Spatial Information Science*, Vol. 45, No. 3, pp. 245-258. E-ISSN 1948-660X.

[7] Chen, L., Mao, Y. and Zhao, R. (2022) "GIS application in environmental monitoring and risk assessment", *2022 3rd International Conference on Geology, Mapping and Remote Sensing (ICGMRS)*, pp. 908-917. DOI 10.1109/icgmrs55602.2022.9849269.

[8] Das, S., Baral, A., Rafizul, I. M. and Berner, S. (2024) " Efficiency enhancement in waste management through GIS-based route optimization", *Cleaner Engineering and Technology*, Vol. 21, p. 100775. E-ISSN 2666-7908. DOI 10.1016/j.clet.2024.100775.

[9] Eldrandaly, K. A. (2021) "*Integrating expert systems and GIS for spatial decision making: Current practices and new trends*", Akademia. 22 p. [Online]. Available: https://www.academia.edu/53829528/Integrating_Expert_Systems_and_GIS_for_Spatial_Decision_Making_Current_Practices_and_New_Trends [Accessed: June 15, 2025].

[10] Flak, J. (2020) "Technologies for sustainable biomass supply—overview of market Offering", *Agronomy*, Vol. 10, No. 6, p. 798. ISSN 2073-4395. DOI 10.3390/agronomy10060798.

[11] Giuffrida, N., Calleo, Y., Di Zio, S., Pilla, F. and Ottomanelli, M. (2024) "Spatial participatory planning for urban logistics: A GIS-enhanced real-time spatial delphi approach", *Research in Transportation Economics*, Vol. 108, p. 101488. E-ISSN 1875-7979, ISSN 0739-8859. DOI 10.1016/j.retrec.2024.101488.

[12] Green, S., Patel, R. and Kim, H. (2021) "GIS applications in palm oil plantation management: A review", *Journal of Agricultural Engineering*, Vol. 36, No. 2, pp. 178-192. ISSN 1974-7071.

[13] Jones, A. and Smith, L. (2019) "Leveraging GIS for plantation area estimation: A case study of rubber and palm oil", *Remote Sensing Applications*, Vol. 67, No. 4, pp. 100-112. ISSN 2352-9385.

[14] Kim, J., Lee, C. and Park, S. (2020) "Using GIS to optimize route planning in the rubber industry supply chain", *Journal of Industrial Engineering*, Vol. 52, No. 5, pp. 333-348. E-ISSN 2224-7890.

[15] Kongmanee, C., Jongrungrot, V., Somboonsuke, B. and Phitthayaphinant, P. (2017) "A simple method of potential biomass feasibility: A case study of potential rubberwood biomass feasibility as feedstock for a very small biomass power plant in the three southern border provinces", *WMS Journal of Management*, Vol. 6, No. 3, pp. 93-106. E-ISSN 2286-718X.

[16] Lee, J., Park, C. and Yoon, H. (2020) "GIS and its role in rubber plantation management: A spatial analysis approach", *Journal of Agricultural Systems*, Vol. 54, No. 3, pp. 256-270. E-ISSN 1873-2267, ISSN 0308-521X.

[17] Li, H., Huang, W., Zhai, Y., Zhao, W., Zheng, X. and Liu, J. (2024) "Research on new spatial data infrastructure supports the circulation of geographic information data elements", *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. 48, pp. 303-309. ISSN 2194-9034. DOI 10.5194/isprs-archives-XLVIII-4-2024-303-2024.

[18] Li, W., Zhang, P. and Zhao, L. (2020) "Developing GIS-based mobile applications for travel estimation in agricultural logistics", *Computers and Electronics in Agriculture*, Vol. 10, No. 3, pp. 66-78. ISSN 0168-1699.

[19] Li, X. and Zhao, F. (2021) "Using GIS and mobile applications for environmental monitoring and sustainability", *Journal of Environmental Science*, Vol. 65, No. 2, pp. 145-160. E-ISSN 1878-7320, ISSN 1001-0742.

[20] Lu, Y., Zhai,G., Zhou, S. (2024) "An integrated bayesian networks and geographic information system (BNs-GIS) approach for flood disaster risk assessment: A case study of Yinchuan, China", *Ecological Indicators*, Vol. 166, p. 112322. E-ISSN 1872-7034, ISSN 1470-160X. DOI 10.1016/j.ecolind.2024.112322.

[21] Mennecke, B. and Crossland, M. (1996) "Geographic information systems: Applications and research opportunities for information systems researchers", *Proceedings of the 29th Annual Hawaii International Conference on System Sciences*, Vol. 3, pp. 537-546. DOI 10.1109/hicss.1996.493249.

[22] Miller, P., Johnson, D. and Liu, Q. (2020) "Real-time GIS applications in rural transportation planning", *Journal of Cybersecurity and Mobility*, Vol. 7, No. 2, pp. 89-104. E-ISSN 2245-4578.

[23] Ministry of Energy. (2020) "*Thailand power development plan 2018-2037*", Office of Energy Policy and Planning, Ministry of Energy. [Online]. Available: https://climate-laws.org/document/thailand-power-development-plan-2018-2037_110d [Accessed: June 15, 2025].

[24] Narayanan, K. J. S. and Manimaran, A. (2023) "Recent developments in geographic information systems across different application domains: A review", *Knowledge and Information Systems*, Vol. 66, No. 3, pp. 1523-1547. ISSN 0219-3116. DOI 10.1007/s10115-023-01969-5.

[25] Siriwongpaisarn, S. and Kongkaew, V. (2018) "Selection of location of farmers' group representatives for fertilizer management in the supply chain of oil palm industry in Surat Thani Province", *UBU Engineering Journal*, Vol. 11, No. 2, pp. 47-57. E-ISSN 2775-2674. ISSN 1906-392X.

[26] Office of Industrial Economics. (2020) "*Statistical report*", Office of Industrial Economics. [Online]. Available: http://rubber.oie.go.th/Directory.aspx [Accessed: June 20, 2025].

[27] Parker, G. and Van Alstyne, M. (2020) "GIS in mobile application development: A review of travel estimation methods", *International Journal of Mobile Computing*, Vol. 12, No. 4, pp. 103-119. ISSN 1937-9404.

[28] Prastacos, P. (1991) "Integrating GIS technology in urban transportation planning and modeling", *Transportation Research Record*, p. 1305, pp. 123-130. ISSN 2169-4052.

[29] Rane, N. L., Giduturi, M., Choudhary, S. P. and Pande, C. B. (2023) "Remote sensing (RS) and geographical information system (GIS) as a powerful tool for agriculture applications: Efficiency and capability in agricultural crop management", *International Journal of Innovative Science and Research Technology*, Vol. 8, No. 4, ISSN 2456-2165. DOI 10.5281/zenodo.7845276.

[30] [Sangeetha, C., Moond, V. M N. R. G., Damor, J. S., Pandey, S. K., Kumar, P. and Singh, B. (2024) "Remote sensing and geographic information systems for precision agriculture: A review", *International Journal of Environment and Climate Change*, Vol. 14, No. 2, pp. 287-309. E-ISSN 2581-8627. DOI 10.9734/ijecc/2024/v14i23945

[31] Sarbazvatan, A. and Karimi, N. (2023) "Development of a mobile GIS application (LandInfo) for land use and land cover field data collection", *Research Gate*. DOI 10.21203/rs.3.rs-3752659/v1.

[32] Smith, E., Williams, J. and Thompson, B. (2018) "Integrating GIS into mobile app development for forestry management", *Journal of Geographical Systems*, Vol. 62, No. 3, pp. 345-360. E-ISSN 1435-5949, ISSN 1435-5930.

[33] Swain, S. (2024) "*How has GIS and remote sensing revolutionized the logistics and supply chain industry"*, AGSRT. [Online]. Available: https://www.agsrt.com/post/how-has-gis-and-remote-sensing-revolutionized-the-logistics-and-supply-chain-industry [Accessed: April 8, 2025].

[34] Takahashi, S., Okumura, M., Tsuruta, N., Torii, M., Inakura, H., Ohno, M. and Okuno, M. (2016) "Geo-Log mobile: Development of mobile GIS application based on new geological database framework for eruptive history and informatics", *Proceedings of the International MultiConference of Engineers and Computer Scientists,* Vol. 2015, No. 2, pp. 8-10. ISSN=21895155

[35] Wajszczuk, K., Baum, R. and Wielicki, W. (2008) "A proposal of a logistics model for the use of biomass for energy for local communities within the concept of sustainable rural development", *European Association of Agricultural Economists*, *107th Seminar,* January 30-February 1, 2008 1-14. Sevilla, Spain. 13. p. DOI 10.22004/ag.econ.6454.

[36] Wang, Z. and Zhao, L. (2023) "Using GIS for real-time route mapping in biomass transportation", *IEEE Transactions on Cloud Computing*, Vol. 10, No. 2, pp. 120-135.

[37] Weiner, D., Harris, T. M. and Craig, W. J. (2002) "Community participation and geographic information systems", In Craig, W. J., Harris, T. M. Weiner (eds).*Community participation and geographical information systems,* pp. 3-16. 1st ed., CRC Press. E-ISBN 9780429203961. DOI 10.1201/9780203469484

[38] White, T. and Lee, A. (2022) "GIS-based mobile applications: enhancing real-time decision making", *Journal of User Experience Design*, Vol. 15, No. 2, pp. 109-122. ISSN 1931-3357.

[39] Zhang, H., Chen, L. and Gao, J. (2019) "Mobile GIS applications for travel estimation in agricultural logistics", *Journal of Transport Geography*, Vol. 45, No. 4, pp. 90-104. ISSN 1873-1236.

Agris on-line Papers in Economics and Informatics

Volume XVII Number 4, 2025

Comparative Advantages and Specialization Dynamics in Agri-food Trade of Argentina, Paraguay and Uruguay

Ivo Zdráhal[1], Martin Hrabálek[2], Petr Kadlec[3], Oldřich Krpec[4]

[1] Department of Regional and Business Economics, Mendel University in Brno, Czech Republic

[2] Department of Area Studies, Mendel University in Brno, Czech Republic

[3] Department of Radio Electronics, Brno University of Technology, Czech Republic

[4] Department of International Relations and European Studies, Masaryk University, Czech Republic

Abstract

The article interrogates the shape, dynamic, and fragility of revealed comparative advantages of 46 agri-food products traded by Argentina, Paraguay, and Uruguay in the period 1995-2020 using normalized revealed comparative index and summary statistics, stochastic kernels, Galtonian regression, Markov chains, and Kaplan-Meier survival analysis. The analysis reveals the agri-food flagship products and the agri-food trade of these countries has formed mainly around these flagship products. The results support the argument that changes in distribution of comparative advantages in agri-food trade underwent an increase in specialization in these countries, especially in the period from the beginning of millennia until about period slightly after the Great Recession. The results also indicate slight convergence in the change in agri-food comparative advantages in these three countries, as well as the increased complexity of agri-food comparative advantages in Argentina and Paraguay at the end of the period under scrutiny. Despite these variations and differences among countries under scrutiny the distribution of comparative advantages remains stable and persistent. Given this evidence, we conclude that these countries will continue to develop their agriculture-led growth economic model and these flagship products will play an important role in the overall agri-food export structures of these countries in the future.

Keywords

Liberalization, Markov chain model, regression analysis, revealed comparative advantage, RCA, agri-food trade.

Zdráhal, I., Hrabálek, M., Kadlec, P. and Krpec, O. (2025) "Comparative Advantages and Specialization Dynamics in Agri-food Trade of Argentina, Paraguay and Uruguay", *AGRIS on-line Papers in Economics and Informatics*, Vol. 17, No. , pp. 137-155. ISSN 1804-1930. DOI 10.7160/aol.2025.170410.

Introduction

Trade liberalization has been recognized as a vital engine of economic growth, with substantial contributions to rising living standards worldwide (Wacziarg and Welch, 2008; Gnangnon, 2018). Yet, despite these strides, numerous barriers still hinder the full realization of trade's potential benefits, both locally and globally, due to persisting trade distortions and policy constraints (WTO, 2024). The international performance of agri-food trade has significant implications for the entire external economic balance of member countries of Mercado Común del Sur (The Southern Common Market), also known as MERCOSUR, as could be seen, for example, in UNCTAD data (UNCTAD, 2023). On average, the importance of agri-food trade (SITC 0+1+22+4) in the structure of total exports is 31.0%, 51.7%, 56.6% and 55.5% in the case of Brazil, Argentina, Uruguay, and Paraguay between 1995 and 2020. Contrary to the developed countries, the share of agri-food exports in total exports increased in these countries in the past decades. MERCOSUR represents 9.5% of global agri-food exports in 2022 (UNCTAD, 2024) and is a serious competitor to the global market dominance of the world's leading agri-food exporters, the EU and the United States (Hopewell, 2016). This also puts the MERCOSUR countries in an important position with regard to the provision of food security at the global level.

In this article, we would like to examine comparative advantages and specialization dynamics in the agri-food trade of three out of four

MERCOSUR countries – Argentina, Paraguay and Uruguay. The reason behind it is that those countries, and Latin America in general, have enormous potential to increase their competitiveness in agri-food production and trade and become a global leader, as pointed for example by Borges Aguiar and Balogh (2022).

We are aware of the fact that Brazil is by far the most important agricultural producer out of the MERCOSUR countries, but given that its position in international agri-food trade was addressed several times (see Hubbart et al., 2017; Jank et al., 2019; Zdráhal et al., 2021), we focus on the other three members of MERCOSUR. The idea is that these countries have not been adequately covered so far by academic literature, although the agri-food sector is a crucial part of their economies, even more so when we speak about exports. The analysis covers a long period of years 1995-2020. This approach gives the authors the possibility to cover different stages in the period before COVID-19 and to concentrate on the trade dynamics in a larger period than in other studies.

Against this background, the aim of the article is to evaluate evolution of the overall degree of specialization (the external shape of the distribution) and the degree and the pattern of intra-distributional mobility (changes in the intra-distributional dynamics) within the agri-food trade of smaller MERCOSUR countries as Argentina, Uruguay, and Paraguay in the last almost three decades before COVID-19.

Literature review

Dynamic of trade patterns and comparative advantages

The pattern of trade refers to the nature of trade between countries and how this changes over time. The evolution of trade patterns often reflects deep structural changes in the whole economies of countries, and such patterns usually emerge over long periods, and comparative advantages may not change in the short run (Bojnec and Ferto, 2018). The theory of comparative advantage indicates that specialization according to comparative advantage is a prerequisite for reaping gains from trade (Kowalski, 2011). Generally, trade theories give different predictions regarding the specialization dynamics of a country. According to the Heckscher-Ohlin model, the specialization pattern is formed based on the relative endowment of the countries (and its change) in the production factors. A certain limiting feature of this framework is that assumptions of the model do not have to be met by the economic realities. Other streams of theoretical literature emphasize the endogeneity of technological change (Grossman and Helpman, 1991; Krugman, 1987; Lucas, 1988; Redding, 1999) or economic geography that underlines the importance of agglomeration economies (Krugman, 1991; Fujita et al., 1999). Removing tariffs, product standards and simplifying government formalities reduces the transaction costs of trade, which should lead to an increase in the degree of specialization (Aiginger, 2001). Higher specialization can lead to higher productivity and competitiveness (and vice versa). Each of these streams of theoretical research identifies some forces that lead to persistence in trade patterns and others that stimulate mobility.

However, since comparative advantage is dynamic and develops endogenously over the years, changes in trade patterns often reveal structural changes that occurred in certain countries and regions. Liberalization and integration are channels for improving productivity, scale, and export expansion and a way to improve comparative advantage. The specialization pattern of a country also develops around the comparative advantage of each product. Because of the influence of factor endowment as well as endogenous and exogenous factors on the formation of specialization shape of trade, the shape of specialization is the matter of empirical testing (its evolution over time and the intra-distributional dynamics).

Agri-food trade dynamics of Argentina, Uruguay and Paraguay

Even though the most important economy of the MERCOSUR group is Brazil, the smaller countries – Argentina, Uruguay, and Paraguay – have also a significant influence on the direction of MERCOSUR. Agri-food production and trade are important for these smaller MERCOSUR economies due to their economic development context. It is also important to mention that these countries are important regarding providing food security on the global level.

From the beginning of the new millennium, the intensive growth of agri-food trade followed the implementation of the commitments of the Uruguay Round Agreement on Agriculture. It led to a mandatory 36% cut in average bound tariffs at the end of 2000 for developed countries and 2004 for developing ones (Bureau et al., 2017). Together with the decrease in unilaterally applied tariff levels and preferential applied tariff levels, it improved market access of these countries and increased integration into the global.

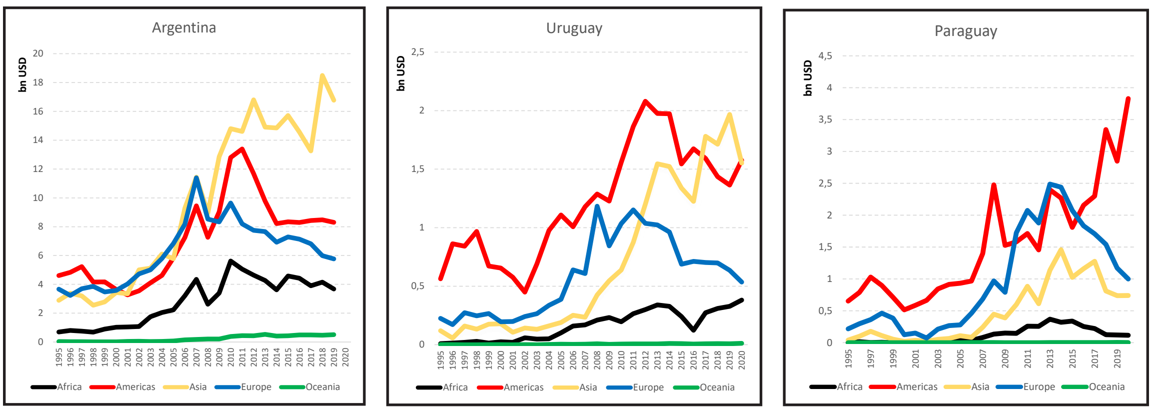
There were also economic and political structural changes in these countries that could affect the comparative advantages and specialization dynamics in the agri-food sectors in Argentina, Uruguay, and Paraguay in the last decades. The whole Latin American region has witnessed rapid increase in agro- and pastoral areas in the last two decades. More specifically, as regards the growth of fields and pastures, the region grew faster than any other region in the world, with the Southern Cone being particularly dynamic in that respect. Most of this growth was an answer to the increasing foreign demand for agri-food products (Norberg, 2020). Agricultural production rose to the forefront of national economic growth, and the agroexport model became rather typical in these countries. The changes in the total agri-food trade of Argentina, Uruguay and Paraguay are presented in the following figure (Figure 1).

The dimension of Argentina's agri-food trade is significantly higher compared to Uruguay and Paraguay. The balance of agri-food trade with its turnover shows, in general, that all three countries are net exporters of agri-food products. Especially Argentina shows a high balance of agri-foods with a high turnover ratio. All three countries have experienced rapid growth in agri-food trade since the start of the new millennium up to about 2014. Liberalization led to a significant increase in the balance-to-turnover ratio, especially in the case of Paraguay. Since 2014, growth trends have changed to slight drop and stagnation.

Comparing the change in agri-food exports of countries under review with the FAO Food price index (see Figure A1 in Appendix), it is evident that the change in agri-food exports was also influenced by the dynamics of prices on the world market.

During the period under scrutiny the territorial shape of agri-food trade has changed (see Figure A2 in the Appendix). The share of traditional export markets of the European Union and to some extent also of the United States has decreased. On the other hand, the rise of Asian (mainly Chinese) demand for many of the typical agri-food products in the production structure of Argentina, Uruguay, and Paraguay caused the change in the territorial shape of agri-food trade and served as an additional impulse forming agri-food trade of these countries. Despite the slowdown in total agri-food exports in the second half of the period under scrutiny, exports to Asia are still growing in Argentina and Uruguay. The change in territorial shape is caused also by intraregional agri-food trade. However, our data indicate a mixed intensity with respect to which each country trades agri-food products regionally. Such a dynamic suggests that there is a common group of factors that form the agri-food trade of Argentina, Uruguay, and Paraguay, and there are probably factors specifically affecting the agri-food trade of each country.

Despite the fact that most of the agri-food production is exported, the MERCOSUR countries face severe constraints as regards their integration into the world agri-food trade, as they were so far not able to successfully conclude negotiations on free trade with their largest external partners,



Note: 1) Trade balance index (TBI) compares the balance of trade (X-M) to its turnover (X+M); the value close to 1 would mean that the country only exports (is a purely net exporter). 2) Note the different scale of the y-axis when comparing Argentina to Uruguay and Paraguay.

Source: Authors' calculations based on UNCTAD data (SITC, 3-digit level)

Figure 1: Change in total agri-food trade of Argentina, Uruguay, and Paraguay; 1995-2020.

namely United States, European Union, and China. This causes certain frustrations between the MERCOSUR countries and could undermine the unity of the countries, as Uruguay was considering bilateral trade cooperation with China in the past (Urdinez et al., 2016), something that is clearly against the rules of the integrational project. However, China presents the largest export market for agricultural products from MERCOSUR countries: Two-thirds of Argentina and more than half of Uruguayan bovine meat exports end up there (Nolte, 2021).

In the case of the negotiations with the European Union, the progress is very slow. After almost 25 years, the countries have not sealed a final deal, mostly due to the resistance from the European side in the agricultural chapter, mostly about the beef quotas. This resistance is based both on the agricultural interests of particular European countries (Luciano and Borges, 2022), given by the high level of competitiveness of MERCOSUR countries and concerns about environmental impacts of the agreement (Monnerat, 2023).

Material and methods

Revealed comparative advantage

To assess the shape of trade specialization and its stability, it is critical to define the distribution of the country's comparative advantage (CA) in the products trade. The traditional approach is based on the concept of ‘revealed’ comparative advantage (RCA). Balassa (1965; 1977) developed the empirical method and it is widely used to identify a nation's most robust and weakest export sectors. The theoretical foundation and empirical distribution characteristics of the Balassa index have long been debated and criticized in the literature (Bowen, 1983; Vollrath, 1991; Hinloopen and Van Marrewijk, 2001; Sanidas and Sin, 2010; De Benedictis and Tamberi, 2004; Gnidchenko and Salnikov, 2015). Due to the shortcomings of the Balassa index, other indices have been proposed (i.e., Bowen, 1983; Vollrath, 1991; Lafay, 1992; Dalum et al., 1998; Proudman and Redding, 2000; Hoen and Oosterhaven, 2006; Yu et al., 2009; Leromain and Orefice, 2014; Gnidchenko and Salnikov, 2015; Stellian and Danna-Buitrago, 2022).

The normalized revealed comparative advantage index (NI) was proposed by Yu et al. (2008) as another alternative measure of RCA. Due to the fact that it is comparable across products, countries, and time, the index should more precisely and consistently reveal the extent of CA that a country has in a certain product, making it robust quantitative tool.

$$\mathbf{NI} = \frac{x_{ij}}{x_w} - \frac{x_j \times x_{wj}}{x_w \times x_w} \tag{1}$$

where *x* indicates exports, *i* represents a nation, *j* signifies a product and *w* represents a set of countries (world). The *NI* index ranges from -0.25 to 0.25 and the comparative neutral point is zero. The sum (and the mean value) of the *NI* scores is constant and equals zero, and a sum of positive scores equals the sum of negative scores. If *NI* is higher (lower) than 0, the country reveals comparative advantage (disadvantage) in product *j*. The higher the value, the stronger the CA and vice versa.

Statistical data were obtained from databases of the UNCTAD for the period between 1995 and 2020. The commodity structure of individual sectors (product groups) in the agri-food trade is defined according to the SITC (revision 3). The analysis is carried out at the level of a 3-digit code for 46 various agri-food products traded (SITC 0 + 1 + 22 + 4). The values of NI have been calculated for each agri-food product group (Table A1 in the Appendix) traded between analysed countries (Argentina, Uruguay, and Paraguay) and the World market for the period 1995-2020.

Scores of NI scores were examined on how weak or strong the comparative advantage of each product is. There is no general guidance in the literature for classifying NI values into classes. Following current empirical studies (eg, Kostoska and Hristoski, 2018), the data were grouped based on quartiles (relative thresholds) from positive values of RCA scores calculated for all MERCOSUR countries. Positive NI values are grouped as: revealed comparative disadvantage (CdisA) if $NI \leq 0$, weak comparative advantage (CA) if $NI > 0$ and $NI \leq 2.877$ (first and second quartile), medium CA if $NI > 2.877$ and $NI \leq 16.352$ (third quartile), and strong CA if $NI > 16.352$ (fourth quartile).

First, a static view of the comparative advantage of individual products is applied using mean values (1995-2020) of NI to assess and identify agri-food products revealing CA or CdisA and so the potential capacity to cope with competitive pressures (and to identify agri-food flagship products). Following recent empirical studies in agri-food trade (Hinloopen and Van Marrewijk, 2001; Ferto and Hubbard, 2003; Ban, 2017; Smutka et al., 2018; Kostoska and Hristoski, 2018; Hoang, 2019; Vondráček et al., 2022), a battery of empirical approaches is employed to analyse structural stability of the comparative advantage of agri-food of Argentina, Uruguay, and Paraguay. According to Hinloopen and Van Marrewijk (2001), one can distinguish at least two types of stability; first, the stability of the distribution of the indices from one period to the next (Stability I); second, the stability of the value of the indices for particular product groups from one period to the next (Stability II). The stability of the distribution of 4 product groups (CdisA, weak, medium, and strong CA) was evaluated, according to the strength of Cdis/CA. In addition, an alternative approach is the duration analysis based on the survival function methodology.

Stability I

Based on procedures suggested by Hinloopen and van Marrewijk (2001), summary statistics was employed to investigate the external shape of the distribution of RCA indices.

Next, following Ban (2017) the intra-distributional dynamic was analysed applying stochastic kernels. It is a non-parametric approach useful to estimate the density function. Because the empirical distributions are highly leptokurtic, in order to minimize the impact of extreme values, the NI scores were transformed:

$$\mathbf{TNI} = \frac{\mathbf{NI}}{\mathbf{abs(NI)+1}} \tag{2}$$

An accurate estimation depends on the bandwidth (h) correctly determined. The rule suggested by Silverman (1986) was applied:

$$\mathbf{h = 0.9 \times A \times N^{-1/5}}, \text{ where}$$

$$\mathbf{A = min(\sigma, IQR/1.349)}, \tag{3}$$

N is the number of observations, and *IQR* represents the interquartile range and is calculated as the difference between the third and first quartile. Estimated probability densities are presented for two-year averages identified by summery statistics as time points when there could be potential change in dynamic in the distribution of CA.

Next, Galtonian regression (Galton, 1889) was calculated and applied as a quantitative technique. This methodology was successfully deployed in previous studies (e.g., Hart, 1995; Cantwell, 1989; Dalum et al., 1998; Laursen, 1998; Ban, 2016). The regression is used as an indicator of structural stability that indicates changes in trade specialization patterns. The Galtonian method is linear simple regression with respect to the two cross sections of two different time periods of RCA indices scores (Sanidas and Shin, 2010). The corresponding regression equation for a given country to test the changes in the trade specialization pattern is as follows.

$$\mathbf{RCA_{ij}^{t2} = \alpha_i + \beta_i \times RCA_{ij}^{t1} + e_{ij}} \tag{4}$$

where *t*1 indicates the initial point in time and *t*2 the last point in time, *ai* and *ßi*, are standard regression coefficients and eij is an error term.

The fundamental concept of the Galtonian regression is basically to assess the similarity or dissimilarity between the distributions of revealed comparative advantages for two different points in time. This provides an indication of the stability, or convergence/divergence of the trade specialization patterns. To interpret the corresponding results, the overall changes can be expressed in terms of the regression effect, indicated by *ß*, and the mobility effect, indicated by *R*, where:

- If $ß > 1$, there is a divergence of trade specialization.
- If $R < ß < 1$ divergence of trade specialization
- If $ß < R < 1$ convergence of trade specialization
- If $ß = R$ there was no convergence or divergence, the trade specialization pattern remained more or less stable then there are no changes in the hierarchy of the industries (the ones that are competitive remain competitive and the ones that are non-competitive remain non-competitive) or in the relative position versus the other industries (the products that have an advantage do not have an advantage).

Divergence means that the pattern of trade specialization has strengthened due to industries with an initial revealed comparative advantage that exhibit an increase in this score, while the score for those exhibiting an initial revealed comparative disadvantage decreased. Convergence means that the opposite dynamics is taking place. That means that the pattern of trade specialization can be considered to have weakened.

Another indicator of stability is the relative importance (in the export structure) of product groups that reveal CA in the period *t*, but reveal CdisA in the period *t* + 1 or vice versa (Kostoska and Hristoski, 2018). Because of the year-by-year fluctuation, the comparison is made between average value of indices at the beginning (avg. 1995-1996) and at the end (2019-2020) of the observed period.

Stability II

To assess structural changes in the overall structure, as well as at the sectoral level, stability is analysed in terms of the distribution of NI scores of specific products in 4 groups (CdisA, weak, medium, and strong CA) from period to period. Following Quah (1996), Proudman and Redding (2000), Brasili et al. (2000), Hinloopen and Van Marrewijk (2001) and Zaghini (2005), the changes in distribution of product groups among the particular classes were analysed using the Markov chain model. The evolution of the NI distribution over time may be modelled formally to measure the probability that a product group moves from one class to another. Thus, represent the NI by the measure x and its distribution across sectors at time t by $Ft(x)$. Similarly to Ft, it can be define a probability measure λt, where $\lambda t((\lambda(t - 1), x)) = Ft\ (x)$. The evolution of the distribution of RCA over time is then modelled in terms of a stochastic difference equation:

$$\lambda_t = P(\lambda(t - 1), u_t) \tag{5}$$

where u_t is the error term and P is an operator that measures if an element, initially part of the Ft - 1 distribution, will end in Ft. If the operator P is time invariant and the disturbances are equal to zero, by iterating the relation above, we could obtain:

$$\lambda(t + s) = P \times \lambda(t + s - 1) = ... = \mathrm{P}^s \times \lambda_t \tag{6}$$

Allocating the NI into the classes, the operator P becomes a transition matrix. An element of i_t, p_{ij}, represents the probability that a value that at the beginning of the period was in the state i, will be, after s years, in the state j. If the large values are situated on the main diagonal of the transition matrix, then the mobility inside the distribution is rather small and vice versa. The general degree of mobility can be assessed using the trace and the determinant of transition matrix, as follows:

$$M1 = (n - tr(P))/(n - 1), \text{ respectively}$$

$$M2 = 1 - |det(P)| \tag{7}$$

where n is the number of classes used in the product mapping; $tr(P)$ is the trace of matrix P; $det(P)$ is the determinant of matrix P. $M1$ captures the importance of diagonal and off-diagonal terms. In the case of total persistence, the value of $M1$ would be zero. In the case of total mobility, $M1$ would be 1. The $M2$ gives a similar explanation. When the values on the main diagonal are close to 1 and those on the off-diagonal small (high persistence), the matrix determinant takes a value close to 1 and the mobility indicator is 0. The assessment of persistence, resp. mobility was done throughout the reporting period. Furthermore, the year-on-year $M1$ and $M2$ indices were calculated to assess their changes from 1995 to 2020. This allows one to assess whether the distribution of product groups among product mapping classes is already formed or is still undergoing changes.

Fragility / duration analysis

Following the latest empirical trade literature (Besedeš and Prusa, 2006; Bojnec and Fertö, 2016; Kostoska and Hristoski, 2018) we employ the duration analysis to interrogate the fragility defined as the length of duration of the revealed comparative advantage at product level. Duration is defined as the time (measured in years) that a product maintained CA without any interruption. The duration can be modelled as a sequence of conditional probabilities that the comparative advantage revealed by the product continues after t periods, as it has already survived for t periods. Specifically, let T be a random variable that denotes the length of a spell with uninterrupted $NI > 1$. In discrete time, the survival function $S(T)$ is defined as:

$$S(T) = Pr(T \geq t) \tag{8}$$

Specifically, the duration of CAs of products consisting agri-food trade of Argentina, Uruguay, and Paraguay is estimated by the Kaplan-Meier product limit estimator. The Kaplan-Meier survival analysis (Kaplan and Meier, 1958) is non-parametric statistics used to estimate the probability of survival past given time points (describes and quantifies time-to-event data). The event is defined as failure ($NI > 0$ resp. $LFI > 0$ shifts to $NI \leq 0$ resp. $LFI \leq 0$). Survival time specifies the time-to-event duration. (Time for an event to occur). The longer $NI > 0$ (resp. $LFI > 0$) a product has, the longer its survival time and the higher the survival rate. We applied the Kaplan-Meier method to estimate the cumulative survival function for the product groups as a single unit. The survivor function is the share of spells that survive at time t, but this time is cumulative of all preceding time intervals. That is, if all spells survive and the ratio is one, the survivor function is flat at this interval; otherwise, the function is gradually declining. Formally, the Kaplan-Meier estimator of the survival function is as follows:

$$S(t) = \prod_{t(i)} \frac{n_j - d_j}{n_j} \qquad (9)$$

where *NI* denotes the number of subjects at risk of failure in $t(i)$ and d_j denotes the number of observed failures. Estimated survival functions were analysed using the traditional logarithmic rank test to assess their statistical similarity.

Results and discussion

Measurement revealed comparative advantages in agri-food production

The comparative advantages revealed of 46 products (see Table A1 in the appendix) of the agri-food trade of Argentina, Paraguay, and Uruguay were evaluated using the NI index (Table 1).

On average, from 1995 to 2020, Argentina revealed CA for 9 products, Uruguay 11 products, and Paraguay 7 products. The flagship products (revealing medium or strong CA) are colored blue and green in the table. Argentina reveals six flagship products (S011, S041, S044, S081, S222, S421) and both Uruguay (S011, S042) and Paraguay (S011, S222) only two.

There is a link between NI scores and the shares of specific products in export and import structures. Products revealing CA are export sectors and account, on average 1995-2020, for 71.9% (flagship products account for 70.1%) of agri-food exports in Argentina, 81.8% (flagship products: 42.8%) of agri-food exports in Uruguay, and 91.2% (flagship products: 62.8%) of agri-food exports in Paraguay. The products revealing CdisA are the import sectors and account, on average, for 76.8% of the agri-food imports in Argentina, 87.3% of the agri-food imports in Uruguay and 86.1% of the agri-food imports in Paraguay. These results suggest that the specialization patterns of the agri-food trade of Argentina, Paraguay, and Uruguay develop around comparative advantage of each agri-food sector.

S	Argentina		Uruguay		Paraguay		S	Argentina		Uruguay		Paraguay	
	avg.	rank	avg.	rank	avg.	rank.		avg.	rank	avg.	rank	avg.	rank
001	-4.47	34.	0.28	7.	-0.36	25.	057	-8,75	43.	-1.15	42.	-1.98	46.
011	2.91	6.	7.94	1.	3.80	2.	058	-1.32	22.	-0.42	31.	-0.40	28.
012	-10.26	44.	-0.39	29.	-1.32	41.	059	-0.19	13.	-0.31	27.	-0.33	24.
016	-1.00	19.	-0.09	18.	-0.11	12.	061	-4.21	33.	-0.56	35.	-0.38	27.
017	-1.25	21.	-0.04	13.	-0.38	26.	062	-1.45	25.	-0.25	25.	-0.26	22.
022	-3.28	28.	1.33	4.	-0.90	39.	071	-6.93	40.	-0.78	41.	-0.76	37.
023	-1.34	23.	0.22	8.	-0.18	21.	072	-3.67	31.	-0.38	28.	-0.41	29.
024	-4.80	35.	0.43	5.	-0.68	36.	073	-3.61	29.	-0.51	34.	-0:54	31.
025	-0.90	18.	-0.11	19.	-0.12	13.	074	-0.44	15.	-0.18	23.	-0.15	17.
034	-6.00	37.	0.02	11.	-1.33	42.	075	-1.46	26.	-0.17	22.	-0.17	20.
035	-1.11	20.	-0.13	21.	-0.14	16.	081	50.29	1.	-1.29	44.	2.55	3.
036	-0.12	11.	-0.64	37.	-0.80	38.	091	0.00	10.	0.12	10.	-0.12	14.
037	-5.20	36.	-0.50	33.	-0.61	32.	098	-10.41	45.	-1.26	43.	-1.35	43.
041	9.57	5.	-0.11	20.	-0.16	18.	111	-3.62	30.	-0.44	32.	-0.44	30.
042	-2.33	27.	3.06	2.	0.04	7.	112	-11.31	46.	-1.65	46.	-1.74	45.
043	0.50	8.	-0.07	16.	-0.16	19.	121	-0.15	12.	-0.25	24.	-0.12	15.
044	20.06	3.	-0.58	36.	1.14	5.	122	-7.47	41.	-0.40	30.	-0.62	33.
045	0.70	7.	-0.07	17.	-0.08	10.	222	16.19	4.	2.71	3.	12.34	1.
046	0.47	9.	-0.05	14.	-0.09	11.	223	-0.48	16.	-0.06	15.	0.07	6.
047	-0.22	14.	-0.03	12.	-0.03	8.	411	-0.87	17.	0.20	9.	-0.05	9.
048	-6.23	38.	0.34	6.	-1.01	40.	421	33.74	2.	-0.75	40.	1.29	4.
054	-7.88	42.	-1.37	45.	-1.37	44.	422	-6.38	30.	-0.74	39.	-0.66	34.
056	-3.96	32.	-0.67	38.	-0.67	35.	431	-1.36	24.	-0.26	26.	-0.26	23.

Note: Green - strong CA, blue - medium CA, yellow - weak CA

Source: Authors' calculations based on UNCTAD data (SITC, 3-digit level)

Table 1: Mean (1995-2020) and ranking of specific sectors according to NI values.

The year-over-year change indicates fluctuation in NI scores. Products such as S036, S043 (Argentina), S091 (Uruguay) and S042, S223 (Paraguay) gained CA and products as S017 (Argentina), S017, S34 (Uruguay) lost CA compared to the beginning and the end of the period under scrutiny.

Stability I (external shape of the distribution of RCA)

The number of products that reveal CA has changed slightly and the high coefficients of variation of some agri-food products (not presented here) indicate dispersion in the variables between 1995 and 2020. This could signal a relative instability of some products revealing comparative (dis) advantages and a possible ongoing structural change of the agri-food trade pattern. The summary statistics (mean, median, range, skewness, i.e.; see Table A2 in the Appendix), the regression analysis (Table 2), and kernel density estimation (Figure A3 in the Appendix) of NI were investigated to assess the evolution of the external shape of the distribution of RCA and thus the overall change in the degree of specialization.

In the case of Argentina, the summary statistics of the NI scores show an increasing range of quartiles and of standard deviation from 1995 to about the Great recession. The number of products that reveal CA decreased slightly during this period. This is in line with the results of the regression analysis. The regression analysis indicates divergence from the beginning of the millennia until the Great Recession. This means that products that reveal CA increased the strength of CA and the products that reveal comparative disadvantage have weakened their position compared to the previous period of time. This suggests that the liberalization and commodity boom in the first decade of the century enhanced the specialization pattern of Argentina's agri-food trade. After the Great recession, results suggest convergence (also, range and standard deviation started to decrease), and thus changes were leading to slightly (the pattern diverge again from about 2013) new specialization pattern.

Similarly to Argentina, regression analysis of NI scores for Uruguay and Paraguay indicates divergence from the beginning of the millennium until the Great Recession. Summary statistics suggest (e.g., stagnation of standard deviation, decrease in range) that since about 2013, the strengthening of specialization of agrarian trade has stopped in Uruguay and Paraguay. Regression analysis between 2013-2014 and 2021-2022 indicates convergence. In other words, it means that the pattern of trade specialization can be considered to have weakened. There is a question whether this convergence will lead to a new specialization pattern or its only fluctuation (e.g., linked to the impact of the COVID pandemic on international trade). Median values of the indices fluctuate, but there is a visible tendency to decrease. Overall, these trends in both countries led to a change in the number of products that revealed comparative advantage. The number of products revealing CA has slightly decreased in Uruguay and slightly increased in Paraguay.

The evolution of the external shape of the distribution of RCA and thus the overall change in the degree of specialization was interrogated using kernel density estimation (Figure A3 in the Appendix) of the NI. A result of kernel density estimation confirms conclusions based on summary statistics and regression analysis. Visualization

Country	period	1999-2000				2007-2008				2013-2014			
		β	R	β/R		β	R	β/R		β	R	β/R	
Argentina	07-08	0.984	0.948	1.038	DIV								
	13-14	0.683	0.726	0.940	CON	0.703	0.777	0.906	CON				
	21-22	0.917	0.867	1.057	DIV	0.910	0.894	1.018	DIV	0.900	0.800	1.125	DIV
Uruguay	07-08	0.920	0.848	1.085	DIV								
	13-14	0.874	0.683	1.280	DIV	1.067	0.904	1.180	DIV				
	21-22	0.841	0.734	1.146	DIV	0.972	0.920	1.056	DIV	0.835	0.933	0.895	CON
Paraguay	07-08	1.161	0.964	1.204	DIV								
	13-14	1.283	0.941	1.363	DIV	1.104	0.976	1.132	DIV				
	21-22	1.187	0.916	1.295	DIV	1.021	0.950	1.076	DIV	0.941	0.990	0.950	CON

Note: DIV – divergence, CON - convergence
Source: Authors' calculations based on UNCTAD data (SITC, 3-digit level)

Table 2: Regression analysis of NI scores (selected years).

of the distribution of comparative advantages in agri-food highlights the important position of flagship products in the distribution structure.

Another indicator of (un)stability and of an increased specialization pattern in agri-food trade is the relative importance of products (in the export and import flows) that reveal a CA in the period *t*, but a CdisA in the period *t* + 1 (Table 3) or vice versa (Ballance et al., 1987).

In the case of Argentina, the product groups that revealed CA in 1995-1996 but CdisA in 2019-2020 represented 3.65% of total agri-food export at the beginning of the period and 0.27% at the end. In the case of Uruguay, these product groups represented 11.30% of the total agri-food export at the beginning and 2.75% at the end. In the case of Paraguay, these product groups accounted for 2.70% of the total agri-food export at the beginning and 0% at the end.

The products showing opposite (revealing CdisA at the beginning of the period and CA at the end) accounted for 3.10% of total exports at the beginning of the period and 4.76% at the end. In the case of Uruguay, the change was from 1.03% to 18.88% of the total agri-food exports. In the case of Paraguay, these products represented 0.27% to 6.59% of total agri-food export.

Stability II (intra-distributional dynamics)

Using results of the previous analysis of the overall specialization pattern, it is possible to gather only some information about the shape of the overall distribution of the NI indices, but not much can be said as regards the changes in the relative position of any single product. The mobility of products within the distribution was analysed to investigate intra/distributional dynamics and transitions among the subsequent classes: CdisA (class a), weak CA (class b), medium CA (class c), and strong CA (class d). The scores in Markov transition probability matrices are presented in Table 4.

Products revealing CdisA (class a) are highly persistent in time. This means that products without an initial comparative advantage appear to remain uncompetitive in all three countries during the period under scrutiny. Using NI scores, 29 out of 46, 27 out of 46 and 35 out of 46 products analysed consisting of agri-food trade of Argentina, Uruguay, and Uruguay resp. Paraguay never revealed CA from 1995 to 2020.

	CA → CdisA			CdisA → CA		
	No.	RCAt	RCAt+1	No.	RCAt	RCAt+1
		1995-1996	1919-2020		1995-1996	1919-2020
		%	%		%	%
ARG	1	3.65	0.27	2	3.10	4.76
URG	2	11.30	2.75	2	1.03	18.88
PRG	1	2.70	0.00	3	0.27	6.59

Source: Authors' calculations based on UNCTAD data (SITC, 3-digit level)

Table 3: Stability of revealed comparative advantage - relative importance of products.

Argentina (NI)

Pi,j (-)	i (-)	a	b	c	d
j (-)					
a		0.97	0.03	0.00	0.00
b		0.26	0.64	0.11	0.00
c		0.07	0.18	0.56	0.18
d		0.00	0.00	0.11	0.89

Uruguay (NI)

Pi,j (-)	i (-)	a	b	c	d
j (-)					
a		0.98	0.02	0.00	0.00
b		0.09	0.89	0.02	0.00
c		0.00	0.11	0.89	0.02
d		0.00	0.00	0.00	1.00

Paraguay (NI)

Pi,j (-)	i (-)	a	b	c	d
j (-)					
a		0.99	0.01	0.00	0.00
b		0.09	0.88	0.03	0.00
c		0.00	0.05	0.91	0.05
d		0.00	0.00	0.67	0.33

Source: Authors' calculations based on UNCTAD data (SITC, 3-digit level)

Table 4: Markov transition probability matrices for the NI indices.

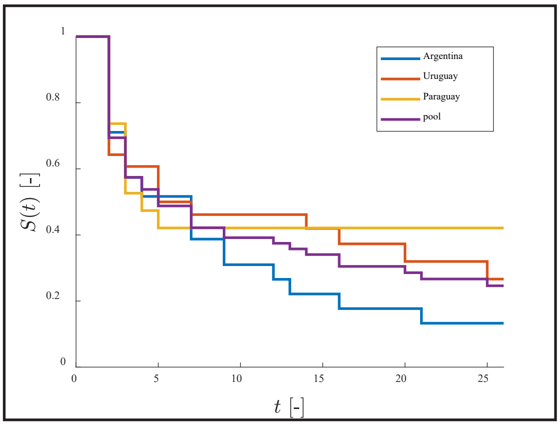
Generally, the products in class a (CdisA) and in class d (flagship products; strong CA) are more persistent than the products revealing weak and medium CA (classes b and c), except of Paraguay. Paraguay revealed a strong CA in export of oil seeds (soya) only in 3 out of 26 years, however, the product revealed most of the time medium CA.

Despite variation when comparing Argentina, Uruguay, and Paraguay among each other, another result indicates that products initially revealing weak CA (class b) rather change its status to CdisA that to change its status to medium or strong CA (groups c and d). Also, the probability of closer shifts is higher than the probabilities of longer moves between classes. The change in the status from weak CA to CdisA can be considered as a channel through which the number of products that reveal CA has decreased and the structure of comparative advantages has weakened in the countries under scrutiny. In the case of Uruguay, another channel through which the structure of comparative advantages has weakened is the relatively low persistent CA in group c (medium CA) because more products have changed its status from medium CA to weak CA that change its status from medium CA to strong CA.

The general degree of mobility was assessed using traces and the determinants of transition matrixes for the whole period, as well as in sub-periods (Table 5). In the case of total persistence, the value of M1 would be zero. In the case of total mobility, M2 would be 1. The M1 and M2 scores indicate high persistence (low mobility), especially in the case of Uruguay, and medium persistence (modest mobility), especially in the case of Argentina. This means that elements on the main diagonal are close to 1. There is no rallying any trend when comparing M1 and M2 scores in sub-periods.

Fragility / duration analysis of the comparative advantage

The Kaplan-Meier survival function was deployed to evaluate the duration of the NI indices during the period under review. We developed a survival function for Argentina, Uruguay, and Paraguay as well as a survival function for pool data. The results of the estimate of survival functions show that survival times are not persistent over the period under scrutiny and there are some differences when comparing Argentina, Uruguay and Paraguay among each other (Figure 2).



Source: Authors' calculations based on UNCTAD data (SITC, 3-digit level)
Figure 2: Plot of survival functions S(t) for Argentina, Uruguay, and Paraguay; 1995-2020.

In the case of pooled NI scores, the chances of comparative advantages of a particular product to survive (revealing CA) at least in four years continuously (not moving from CA to CdisA) are about 50%. Taking into account individual country data, the 50% chance of survival is between 3 to 7 years. The steep slope of the survival function in the first part of the period and then the flat slope over the remaining years indicate that new products turning from CdisA to CA are much more likely to fail than products already revealing CA. The countries under scrutiny differ in this aspect.

NI	Mobility index	Period					
		95-20	95-00	00-05	05-10	10-15	15-20
M1	ARG	0.29	0.34	0.35	0.32	0.26	0.20
	URG	0.07	0.04	0.12	0.07	0.06	0.07
	PRG	0.09	0.20	0.00	0.04	0.18	0.04
M2	ARG	0.67	0.81	0.69	0.65	0.68	0.51
	URG	0.20	0.11	0.32	0.20	0.17	0.19
	PRG	0.25	0.56	0.01	0.11	0.46	0.11

Source: Authors' calculations based on UNCTAD data (SITC, 3-digit level)
Table 5: Mobility indices of the NI.

In the case of Paraguay, if the product survives (reveals CA) more than 5-6 years, there is fairly small risk of failure. In the case of Uruguay and especially Argentina, the product has to survive more years to create stable CA and to increase the probability of the CA to survive. After 26 years (our analysed period), the probability of CA of a product to survive is 13.3%, 26.6% and 42.1% in the case of Argentina, Uruguay, and Paraguay.

Summary of empirical findings and discussion

We examined comparative advantages, specialization dynamics, and its fragility in the agri-food trade of these three countries. We applied the NI index to ask for the comparative advantage of each agri-food product/sector.

The results suggest that the agri-food trade of Argentina, Paraguay, and Uruguay have developed around comparative advantage of each agri-food sectors. Argentina reveals more flagship products (six products: bovine meat, wheat, maize, animal feed, oil seeds, vegetable oils) compared to Uruguay (two products: bovine meat, rice) and Paraguay (two products: bovine meat, oil seeds). These products revealed a medium or strong comparative advantage in the agri-food trade. The observed persistence and gradual evolution of comparative advantages are consistent with Bojnec and Ferto (2018), who emphasize that structural changes in agricultural competitiveness often unfold slowly over long periods due to entrenched factor endowments and institutional constraints. This finding reinforces the notion that agri-food specialization is not highly volatile but path-dependent.

When comparing the change in the number of products revealing comparative advantage, NI indicates that Argentina has increased the number in the second half of the period under scrutiny. A similar increase is seen in the case of Paraguay. The number of products that reveal comparative advantages remains about the same in the case of Uruguay. This suggests a slight improvement in the complexity and export portfolio of the agri-food trade of Argentina and Paraguay.

We evaluated the overall change in the shape of the specialization pattern. The NI index suggests that competitive products became more competitive and products revealing comparative disadvantage have weakened their position in the period from the beginning of the millennia until about 2013 when the agri-food exports dynamic slowed down /stalled in subsequent period. At the end of the covered period, the results indicate a change in trend and an opposite process in Uruguay and Paraguay.

The divergence of trade specialization in the period from the beginning of the millennia until about 2013 is in line with Aiginger (2001) and his conclusion that removing tariffs product standards and simplifying government formalities reduces the transaction costs of trade, which should lead to an increase in the degree of specialization, and higher specialization can lead to higher productivity and competitiveness. The observed divergence of comparative advantages during the early 2000s also aligns with arguments by Lederman, Olarreaga, and Perry (2009), who emphasized how the commodity boom fostered specialization in primary exports across Latin America. Our results are also in line with Fernandez and Curado (2019) concluding that although historically it has been the developed countries that have been responsible for the dominance of commodities in Argentina's export pattern, in recent decades it has been the developing countries of Asia.

Using results of the previous analysis of the overall specialization pattern, it is possible to gather only some information about the shape of the overall distribution of the NI indices, but not much can be said as regards the changes in the relative position of any single product. The mobility of products within the distribution was analysed applying Markov transition probability matrices to investigate intra-distributional dynamics. Products without an initial comparative advantage remained uncompetitive in all three countries during the period under scrutiny. Similarly, the class of flagship products is also stable and persistent. This suggests that the agri-food trade of these three countries was and probably would also be formed around these flagship products in the future. The analysis of fragility/durability of comparative advantage indicates that the 50% chance of survival (product revealing comparative advantage on year-on-year basis) is between 3 and 7 years. In the case of Paraguay, if the product survives more than 5-6 years, there is fairly little risk of failure and loss of its comparative advantage. In the case of Uruguay and especially Argentina, the product must survive more years to create a stable trade partnership and increase the probability of the comparative advantage to survive. After 26 years (our analysed period), the probability of a product's comparative advantage to survive is 13.3%, 26.6% and 42.1% in the case of Argentina, Uruguay, and Paraguay. However, the steep slope of the survival function in the first part of the period and then the flat slope over the remaining years indicate that new products turning from comparative disadvantage to comparative advantage are much more likely to fail than products already revealing comparative advantage. This result implicates that policy makers must focus on supportive policies and improve the enabling environment to ensure the survival of comparative advantages.

The specialization of exports from the agri-food sectors has created a specific channel that transmits impulses/incentives from the global economy to the national economies of these countries. The agri-food led economic growth has support from the political representation. The impact of it on general economic development and domestic food security should become the subject of further research and policy interventions as the recent study of Carrington et al. (2023) concludes that institutional and structural characteristics can bolster resilience in the face of such shocks.

Conclusion

We ex-post examined comparative advantages, specialization dynamics, and its fragility in the agri-food trade of Argentina, Uruguay and Paraguay. These countries have potential to increase its competitiveness and further enhance its position at the World markets in agri-food production and the agri-food led economic growth has support from the political representation. The results suggest that the agri-food trade of Argentina, Paraguay, and Uruguay have developed around comparative advantage of specific agri-food sectors and similar trend can be probably expected in the future. This study identifies these flagship sectors and shows that Argentina reveals more flagship products (bovine meat, wheat, maize, animal feed, oil seeds, vegetable oils) compared to Uruguay (bovine meat, rice) and Paraguay (bovine meat, oil seeds). Results suggest that the multilateral liberalization and commodity boom in the first decade of the century enhanced the specialization pattern of agri-food trade of countries under scrutiny. However, after the Great recession, results suggest slight convergence in specialization patterns, and thus changes that could lead to slightly new specialization pattern and further research is needed. Our results implicate that policy makers should focus on supportive policies and improve the enabling environment to ensure the survival of comparative advantages in these countries under scrutiny (and especially in Argentina).

Despite these minor variations and differences among countries under scrutiny, the distribution of comparative advantages remains mostly stable and persistent (especially in the case of flagship products and products revealing comparative disadvantage). Potential further liberalization steps (removing tariffs and other transaction costs of trade) with EU and China could lead to the further increase in the degree of specialization in specific agri-food sectors and thus in improvement of productivity and competitiveness (ceteris paribus). These findings are important to inform agriculture policymakers and to improve or design structures and mechanisms to stimulate the international performance of agri-food trade in these countries.

The development of global agricultural markets after 2020 has been marked by several serious shocks, e.g. the COVID-19 pandemic, Russian aggression against Ukraine, or now the escalation of tariffs with the accession of the Trump administration in the USA. The methodology used in this article provides the framework to further analyze the competitive position and its stability of MERCOSUR or other countries.

We conclude the three countries under scrutiny have managed to establish very competitive parts of the agri-food sector within global agricultural markets. This is in line with Norberg (2020) concluding that the whole Latin American region has witnessed rapid increase in agro- and pastoral areas as a consequence of the increasing foreign demand for agri-food products globally in the last two decades. Agriculture will most likely remain an important part of its development in the coming years, and the role of the agri-food sector in the export mix of those countries might still grow as a result of further specialization of the markets.

Policy implication

In the context of the EU-MERCOSUR free trade agreement, our analysis has shown strong competitiveness of the three smaller MERCOSUR countries in bovine meat, a product that has been difficult to negotiate in the EU-MERCOSUR trade deal. The fate of the deal remains uncertain today, but if it is successfully ratified and functioning in the near future, it could have further impact on trade specialization of those countries, even though the agreement includes limits on imports. The role of beef meat within the negotiations with the EU is still significant. It has to be noted though that EU is not the most important market for Argentinian, Uruguayan, and Paraguayan bovine meat exports, as most of those exports end up in China, whose role as a destination was increasing in the last decades. The impact of a potential trade liberalization with EU might be limited by that fact, while to Chinese-MERCOSUR trade deal, that is being talked about in 2023 (with negotiations not being started yet) might in fact be far more significant for all countries involved.

Limitations and perspectives of future research

This study has certain limitations and highlights avenues for future research. First, the study relies on aggregated trade data. Despite providing useful understandings already now, a more detailed product classification beyond the SITC 3-digit level could provide further insights into specific product trade dynamics e.g. value-add variances, product quality differentiation, intra-industry trade or microeconomic firm-level complexity and heterogeneity. Additional research is warranted on the role of intermediaries, value-added flows and complex export arrangements within global agri-food value chains between the MERCOSUR countries and major food markets in the world.

Additionally, the use of Kernel densities, Markov chains, and Kaplan–Meier survival functions provides valuable descriptive insights but does not fully uncover causal mechanisms. This methodology does not fully support robust causal conclusions, as we were unable to fully isolate the specific effects determining agri-food trade dynamic of analysed countries as well to fully isolate the country-specific factors on the observed trade differences. Future research could employ econometric approaches to analyse causal relationships more rigorously, specifically examining how global, regional and national factors influence trade flows over time.

Corresponding author:
Ing. Ivo Zdráhal, Ph.D
Department of Regional and Business Economics, Faculty of Regional Development and International Studies
Mendel University in Brno, Zemědělská 1665/1, 613 00 Brno, Czech Republic
E-mail: ivo.zdrahal@mendelu.cz

References

[1] Aiginger, K. (2001) "Specialization of European manufacturing", *Austrian Economic Quarterly*, Vol. 5, No. 2, pp. 81-92. ISSN 1936-4806.

[2] Balassa, B. (1965) "Trade Liberalisation and Revealed Comparative Advantage", *The Manchester School*, Vol. 33, No. 2, pp. 99-123. ISSN 1467-9957. DOI 10.1111/j.1467-9957.1965.tb00050.x.

[3] Balassa, B. (1977) "Revealed' Comparative Advantage Revisited: An Analysis of Relative Export Shares of the Industrial Countries, 1953-1971", *The Manchester School*, Vol. 45, No. 4, pp. 327-344. ISSN 1467-9957. DOI 10.1111/j.1467-9957.1977.tb00701.x.

[4] Ballance, R., Fostner, H. and Murray, T. (1987) "Consistency tests of alternative measures of comparative advantage", *The Review of Economics and Statistics*, Vol. 69, No. 1, pp. 157-161. ISSN 00346535. DOI 10.2307/1937915.

[5] Ban, I.M. (2017) "Measuring trade specialization dynamics: the case of Romania and Bulgaria", *Empirica*, Vol. 44, pp. 229-248. ISSN 1466-4526. DOI 10.1007/s10663-016-9317-7.

[6] Besedes, T. and Prusa, T. (2006) "Product differentiation and duration of US import trade", *Journal of International Economics*, Vol. 70, No. 2, pp. 339-358. ISSN 0022-1996. DOI 10.1016/j.jinteco.2005.12.005.

[7] Bojnec Š. and Ferto I. (2016) "Export competitiveness of the European Union in fruit and vegetable products in the global markets", *Agric. Econ. - Czech*, Vol. 62, No. 7, pp. 299-310. ISSN 1805-9295. DOI 10.17221/156/2015-AGRICECON.

[8] Bojnec, Š. and Ferto I. (2018) "Economic Crisis and the Fragility of Comparative Advantage in European Agriculture", *German Journal of Agricultural Economics*, Vol. 67, No. 3, pp. 147-159. ISSN 2191-4028.

[9] Borges Aguiar, G.M. and Balogh, J.M. (2022) "Analysis of the Competitiveness in the Agri-food sector: The case of Latin America and the Caribbean Region", Competition, Vol. 21, No. 1-2, pp. 92-117. ISSN 2939-7324. DOI 10.21845/comp/2022/1-2/2.

[10] Bowen, H. (1983) "On the Theoretical Interpretation of Indices of Trade Intensity and Revealed Comparative Advantage", *Review of World Economics*, Vol. 119, No. 3, pp. 464-472. ISSN 1610-2886.

[11] Brasili, A., Epifani, P. and Helg, R. (2000) "On the dynamics of trade patterns", *De Economist*, Vol. 148, No. 2, pp. 233-258. ISSN 1572-9982. DOI 10.1023/A:1004065229330.

[12] Bureau, J., Guimbard, H. and Jean, S. (2019) "Agriculture Trade Liberalisation in the 21st century: Has It Done the Business?", *Journal of Agricultural Economics*, Vol. 70, No. 1, pp. 3-25. ISSN 1477-9552. DOI 10.1111/1477-9552.12281.

[13] Carrington, S. J., Olarte, S. H., Urbina, G. (2023) "Commodity cycle management in Latin America: The importance of resilience in face of vulnerability", *Resources Policy*, Vol. 81, p. 103316. ISSN 1873-7641. DOI 10.1016/j.resourpol.2023.103316.

[14] Dalum, B., Laursen, K. and Villumsen, G. (1998) "Structural change in OECD export specialization patterns: de-specialization and stickiness", *International Review of Applied Economics*, Vol. 12, No. 3, pp. 423-443. ISSN 1465-3486. DOI 10.1080/02692179800000017.

[15] De Benedictis, L. and Tamberi, M. (2004) "Overall Specialization Empirics: Techniques and Applications", *Open Economies Review*, Vol. 15, No. 4, pp. 323-346. ISSN 1573-708X. DOI 10.1023/B:OPEN.0000048522.97418.99.

[16] Fernández, V. L. and Curado, M. L. (2019) "Argentina's competitiveness matrix: The natural resource controversy and the country's evolving trade position", *Revista CEPAL*, Vol. 2019, No. 127, pp. 67-89. ISSN 1684-0348. DOI 10.18356/ebf22171-en.

[17] Ferto, I. and Hubbard, L. J. (2003) "Revealed Comparative Advantage and Competitiveness in Hungarian Agri-Food Sectors", *The World Economy*, Vol. 26, No. 2, pp. 247-259. ISSN 1467-9701. DOI 10.1111/1467-9701.00520.

[18] Fujita, M., Krugman, P. and Venables, A. (1999) "*The Spatial Economy: Cities, Regions, and International Trade*", Cambridge, MA: MIT Press. ISBN-13 978-0262561471.

[19] Gnangnon, S. K. (2018) "Multilateral Trade Liberalization and Economic Growth", *The Journal of Economic Integration*, Vol. 33, No. 2, pp. 1261-1301. ISSN 1976-5525. DOI 10.11130/jei.2018.33.2.1261.

[20] Grossman, G. and Helpman, E. (1991) "*Innovation and Growth in the Global Economy*", Cambridge, MA: MIT Press. ISBN 978-0262570978.

[21] Hart, P. E. (1976) "The dynamics of earnings, 1963-1973", *The Economic Journal*, Vol. 86, No. 3, pp. 551-565. ISSN 0013-0133. DOI 10.2307/2230799.

[22] Hinloopen, J. and Marrewijk, C. (2001) "On the Empirical Distribution of the Balassa Index", *Review of World Economics*, Vol. 137, No. 1, pp. 1-35. ISSN 1610-2886. DOI 10.1007/BF02707598.

[23] Hoang, V. (2019) "Investigating the evolution of agricultural trade specialization in transition economies: A case study from Vietnam", *The International Trade Journal*, Vol. 33, No. 4, pp. 361-378. ISSN 1521-0545. DOI 10.1080/08853908.2018.1543622.

[24] Hoen, A. and Oosterhaven, J. (2006) "On the measurement of competetive Advantage", *The Annals of Regional Science*, Vol. 40, No. 3, pp. 677-691. ISSN 1432-0592. DOI 10.1007/s00168-006-0076-4.

[25] Hubbard, C., Alvim, A. M., Mattos, E. J. and Hubbard, L. (2017) "Agri-food Trade Between Brazil and the EU", *EuroChoices*, Vol. 16, No. 1, pp. 11-16. ISSN 1746-962X. DOI 10.1111/1746-692X.12144.

[26] Jank, M. S., Zerbini, A. N. and Cleaver, I. (2019) "Global competitiveness of the Brazilian agri-food sector", In: Buainain, A. M., Lanna, R., Navarro, Z. (eds): "*Agricultural Development in Brazil*". London, Routledge. DOI 10.4324/9781351029742.

[27] Kostoska, O. and Hristoski, I. (2018) "Trade dynamics, revealed comparative advantage, and international competitiveness: Evidence from Macedonia", *Economic Annals*, Vol. 63, No. 218, pp. 23-58. ISSN 0013-3264. DOI 10.2298/EKA1818023K.

[28] Kowalski, P. (2011) "Comparative Advantage and Trade Performance: Policy Implications", *OECD Trade Policy Papers*, No. 121, Paris, OECD Publishing. DOI 10.1787/5kg3vwb8g0hl-en.

[29] Krugman, P. (1987) "The narrow mowing band, the Dutch disease, and the competitive consequences of Mrs. Thatcher: notes on trade in the presence of dynamic scale economies", *Journal of Development Economics*, Vol. 27, No. 1-2, pp. 41-55. ISSN 0304-3878. DOI 10.1016/0304-3878(87)90005-8.

[30] Krugman, P. (1991) "*Geography and Trade*", Cambridge, MA: MIT Press. ISBN 9780262610865.

[31] Lafay, G. (1992) "The measurement of revealed comparative advantages" In: Dagenais, M. G. and Muet, P.-A. (eds.) "*International Trade Modelling*", London: Chapman & Hall. p. 209-234. ISBN 0-412-45000-3.-1992.

[32] Laursen K. (2015) "Revealed Comparative Advantage and the Alternatives as Measures of International Specialisation", Eurasian Business Review, Vol. 5, pp. 99-115. ISSN 2147-4281. DOI 10.1007/s40821-015-0017-1.

[33] Lederman D, Olarreaga M, Perry G (2009) "China's and India's Challenge to Latin America", Washington, DC: World Bank. ISBN 978-0-8213-7308-8. [Online]. Available: https://1url.cz/TJfwW [Accessed: April 14, 2025].

[34] Leromain, E. and Orefice, G. (2014) "New revealed comparative advantage index: Dataset and empirical distribution", *International Economics*, Vol. 139, pp. 48-70. ISSN 2110-7017. DOI 10.1016/j.inteco.2014.03.003.

[35] Lucas, R. (1988) "On the mechanics of economic development", *Journal of Monetary Economics*, Vol. 22, No. 1, pp. 3-42. ISSN 0304-3932. DOI 10.1016/0304-3932(88)90168-7.

[36] Luciano, B. and Borges, C. (2022) "Beyond parliamentary ratification: the role of national and subnational parliaments in EU-Mercosur trade negotiations", *Journal of European Integration*, Vol. 45, pp. 665-682. ISSN 1477-2280. DOI 10.1080/07036337.2022.2129630.

[37] Monnerat, J. (2023) "*Benevolent Protectionism? Environmental Regulation in the Trade Policy of the European Union and its Implications for the EU-MERCOSUR Trade Agreement*", Rio de Janeiro, Direito Rio, p. 8. [Online]. Available: https://1url.cz/PJUx1 [Accessed: April 14, 2025].

[38] Nolte, D. (2021) "*The EU's Beef with Mercosur: Geo-economics versus Climate Diplomacy*", Washington, Woodrow Wilson International Center for Scholars, p. 29. [Online]. Available: https://1url.cz/Aus7x [Accessed: April 14, 2025].

[39] Norberg, M. B. (2020) "*The Political Economy of Agrarian Change in Latin America: Argentina, Paraguay and Uruguay*", Cham, Palgrave Macmillan, 404 p. ISBN 978-3-030-24585-6.

[40] Proudman, J. and Redding S. (2000) "Evolving patterns of international trade", *Review of International Economics*, Vol. 8, No. 3, pp. 373-396. ISSN 1467-9396. DOI 10.1111/1467-9396.00229.

[41] Quah, D. T. (1996) "Empirics for economic growth and convergence", *European Economic Review*, Vol. 90, No. 6, pp. 1353-1375. ISSN 0014-2921. DOI 10.1016/0014-2921(95)00051-8.

[42] Redding, S. (1999) "Dynamic Comparative Advantage and the Welfare Effects of Trade", *Oxford Economic Papers*, Vol. 51, No. 1, pp. 15-39. ISSN 1464-3812.

[43] Silverman, B. W. (1986) "*Density estimation for statistics and data analysis*", London, Chapman and Hall, p. 175. ISBN 0-412-24620-1.

[44] Smutka, L., Maitah, M. and Svatoš, M. (2018) "Changes in the Czech agrarian foreign trade competitiveness - different groups of partners' specifics", *Agric. Econ. - Czech*, Vol. 64, No. 9, pp. 399-411. ISSN 1805-9295. DOI 10.17221/399/2016-AGRICECON.

[45] Stellian, R. and Danna-Buitrago, J.P. (2022) "Which revealed comparative advantage index to choose? Theoretical and empirical considerations", *Revista CEPAL*, Vol. 138, No. 45-65. ISSN 1684-0348.

[46] Stellian, R., Ojeda-Joya, J. N., Danna-Buitrago, J. P. (2024) "Time stationarity, shape and ordinal ranking bias of RCA indexes: a new set of measures", *Review of the World Economics*, Vol. 160, pp. 675-711. ISSN 1610-2886. DOI 10.1007/s10290-023-00512-6.

[47] Swinnen, J. F. M. (2007) "*Global Supply Chains, Standards and the Poor: How the Globalization of Food Systems and Standards Affects Rural Development and Poverty*", Cambridge, MA: CABI, p. 322. ISBN 978-1-84593-185-8.

[48] UNCTAD (2023) "*World statistical database*". [Online]. Available: https://bit.ly/21GbfKX. [Accessed: Apr. 7, 2024].

[49] Urdinez, F., Burian, C. and de Oliveira, A. (2016) "MERCOSUR and the Brazilian Leadership Challenge in the Era of Chinese Growth: A Uruguayan Foreign Policy Perspective", *New Global Studies*, Vol. 10, No. 1, pp. 1-25. ISSN 1940-0004. DOI 10.1515/ngs-2015-0015.

[50] Vondráček, M., Smutka, L., Vacek, T., Pulkrábek, J., Timoshenkova, I. and Maitah, K. (2022) "Distribuce komparativních výhod na trhu s cukrem v zemích EU a specifické postavení českých exportů " (Distribution of Comparative Advantages on Sugar Market in EU Countries and Specific Position of Czech Exports"), *Listy Cukrovarnické a Řepařské*, Vol. 138, No. 11, pp. 374-384. ISSN 1805-970. (In Czech).

[51] Vollrath, T. (1991) "A Theoretical Evaluation of Alternative Trade Intensity Measures of Revealed Comparative Advantage", *Weltwirtschaftliches Archiv*, Vol. 127, pp. 265-289. ISSN 10043-2636. DOI 10.1007/BF02707986.

[52] Wacziarg, R. and Welch, K. H. (2008) "Trade liberalization and Growth: New Evidence", *The World Bank Economic Review*, Vol. 22, No. 2, pp. 187-231. ISSN 0258-6770. DOI 10.1093/wber/lhn007.

[53] WTO. World Trade Organization (2024) "*World Trade Report 2024 Trade and inclusiveness*" [Online]. Available: https://1url.cz/k1VQP [Accessed: Mar. 12, 2024].

[54] Young, A. (1991) "Learning-by-doing and dynamic effects of international trade", *The Quartterly Journal of Economics,* Vol. 106, No. 2, pp. 396-406. ISSN 0033-5533. DOI 10.2307/2937942.

[55] Yu, R., Cai, J. and Leung, P. (2009) "The Normalized Revealed Comparative Advantage Index", *The Annals of Regional Science,* Vol. 43, No. 1, pp. 267-282. ISSN 1432-0592. DOI 10.1007/s00168-008-0213-3.

[56] Zaghini, A. (2003) "Trade advantages and specialization dynamics in acceding countries", Working paper № 249. Frankfurt am Main, Germany: European Central Bank, pp. 4-15.

[57] Zaghini, A. (2005) "Evolution of trade patterns in the new EU member states", *Economics of Transition*", Vol. 13, No. 4, pp. 629-658. ISSN 2577-6983. DOI 10.1111/j.0967-0750.2005.00235.x.

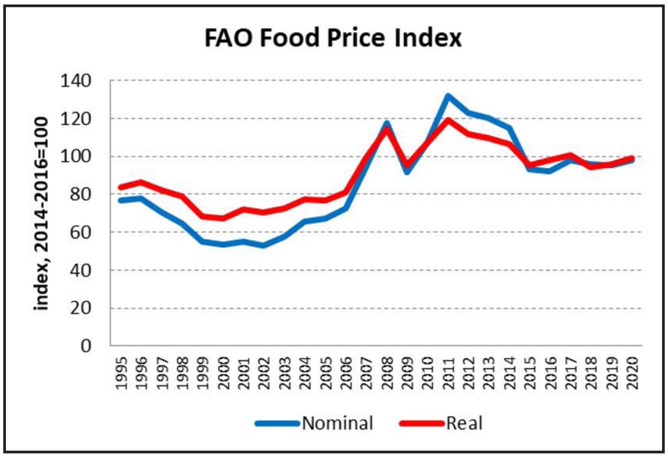
[58] Zdráhal, I., Hrabálek, M., Kadlec, P. and Krpec, O. (2021) "Brazil's Comparative Advantages and Specialization Dynamics in Agri-food Trade", *AGRIS on-line Papers in Economics and Informatics*, Vol. 13, No. 2, pp. 121-139. ISSN 1804-1930. DOI 10.7160/aol.2021.130210.

Appendix

001	Live animals	057	Fruit, nuts excl. oil nuts
011	Bovine meat	058	Fruit, preserved, prepared
012	Other meat, other offal	059	Fruit, vegetable juices
016	Meat, ed. offl., dry, slt, smk	061	Sugars, molasses, honey
017	Meat, offl. Prdd, nes	062	Sugar, confectionery
022	Milk and cream	071	Coffee, coffee substitutes
023	Butter, other fat of milk	072	Cocoa
024	Cheese and curd	073	Chocolate, oth. cocoa prep.
025	Eggs, birds, yolks, albumin	074	Tea and mate
034	Fish, fresh, chilled, frozn	075	Spices
035	Fish, dried, salted, smoked	081	Animal feed stuff
036	Crustaceans, Molluscs	091	Margarine and shorten
037	Fish etc. prepd, prsvd. nes	098	Edible prod. prepetns, nes
041	Wheat, Meslin, Unmilled	111	Non-alcohol. beverage
042	Rice	112	Alcoholic Beverages
043	Barley, unmilled	121	Tobacco, unmanufactured
044	Maize unmilled	122	Tobacco, manufactured
045	Other cereals, unmilled	222	Oil seeds and oleaginous fruits (excl. flour)
046	Meal, Flour of wheat, msln	223	Oil seeds, oleaginous fruits (incl. flour, n.e.s.)
047	Other cereal meal, flours	411	Animal oils and fats
048	Cereal preparations	421	Fixed veg. fat, oils, soft
054	Vegetables	422	Fixed veg. fat, oils, other
056	Vegetables, prpd, prsvd, nes	431	Animal, veg. Fats, oils, nes.

Source: SITC rev. 3

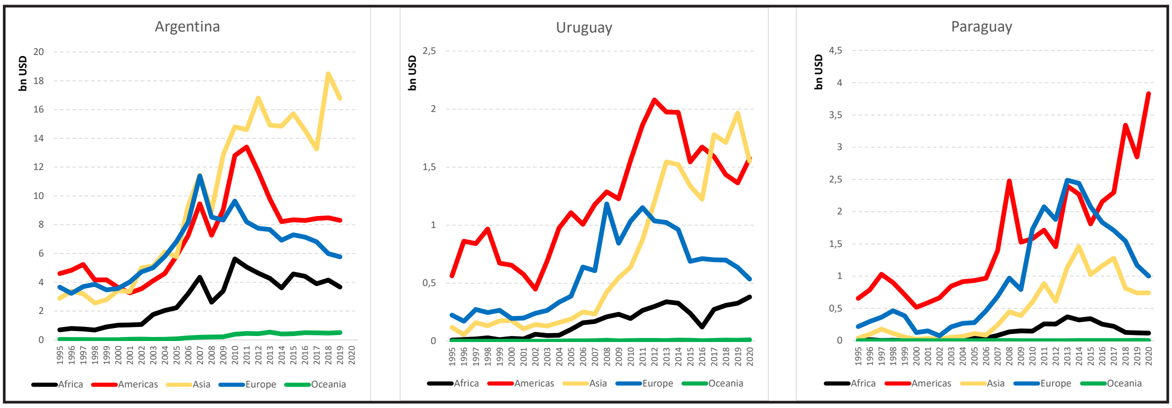
Table A1: Sectors and their numeric designations (SITC rev.3, 3-digit code.



Note: FAO Food price index records the development of world market prices and consists of average of 5 commodity group (meat price index, dairy price index, cereals price index, vegetable oil price index, and sugar price index) price indices weighted with the average export shares of each of the groups for 2014-2016. The base period price consists of the averages for the years 2014-2016.

Source: Authors' calculations based on FAO data (FAO, 2024)

Figure A1: FAO food price index, 2014-2016 = 100, 1995-2020.



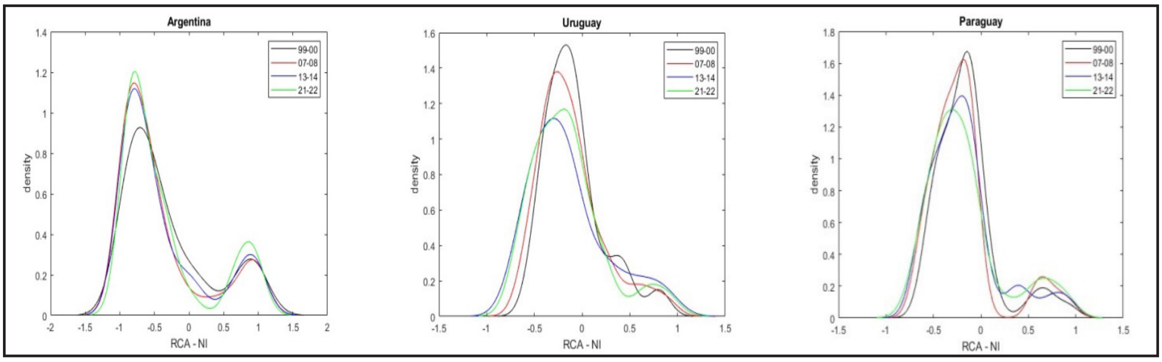
Source: Authors' calculations based on UNCTAD data (SITC, 3-digit level)

Figure A2: Change in total agri-food export of Argentina, Uruguay and Paraguay to particular continents; 1995-2020.

	1995	2000	2005	2010	2015	2020	avg.95-20
	Argentina						
stand dev.	7.85	10.27	11.29	12.91	11.86	9.54	11.14
median	-1.10	-1,55	-1.53	-1.45	-1.78	-1.62	-1.49
kurtosis	11.11	8.20	8.18	12.26	19.00	9.16	12.39
skewness	2.80	2.63	2,72	3.31	3.97	2.78	3.19
min.	-11.72	-14.48	-13.55	-11.34	-12.00	-13.12	-12.91
max.	37.79	44.06	46.52	61.97	64.39	40.87	52.25
range	49.51	58.55	60.07	73.32	76.39	53.99	65.15
	Uruguay						
stand dev.	0.898	1.329	1.614	1.688	1.872	1.599	1.53
median	-0.189	-0.172	-0.252	-0.253	-0.263	-0.220	-0.23
kurtosis	10.286	23.167	31.621	14.867	15.480	20.365	19.05
skewness	2.975	4.424	5.241	3.404	3.675	4.028	3.96
min.	-1.069	-1.518	-1.772	-1.969	-1.896	-1.893	-1.73
max.	3.945	7.535	9.917	8.620	9.329	8.850	8.07
range	5.014	9.053	11.690	10.589	11.225	10.743	9.80
	Paraguay						
stand dev.	1.549	1.268	1.660	2.510	2.216	2.330	2.12
median	-0.188	-0.156	-0.197	-0.374	-0.335	-0.373	-0.30
kurtosis	31.292	32.883	32.809	23.507	10.905	21.301	27.08
skewness	5.234	5.407	5.403	4.545	3.083	4.211	4.81
min.	-1.200	-1.069	-1.424	-2.315	-2.788	-2.771	-2.00
max.	9.497	7.863	10.282	14.233	10.154	12.954	12.34
range	10.697	8.932	11.706	16.549	12.942	15.726	14.34

Source: Authors' calculations based on UNCTAD data (SITC, 3-digit level)

Table A2: Summary statistics of NI scores (selected years).



Source: Authors' calculations based on UNCTAD data (SITC, 3-digit level)

Figure A3: Estimated Kernel densities of transformed scores of RCA indices (avg. 1999-2000, avg. 2007-2008, avg. 2013-2014, avg. 2021-2022).

Agris on-line Papers in Economics and Informatics

Volume XVII Number 4, 2025

Agris on-line Papers in Economics and Informatics

The international reviewed scientific journal issued by the Faculty of Economics and Management of the Czech University of Life Sciences Prague.

http://online.agris.cz

ISSN 1804-1930