Pak. J. Statist.
2026 Vol. 42(2), 157-172

OPTIMIZING HORTICULTURAL MONITORING SYSTEMS USING GRAPH THEORY: A MINIMUM DOMINATING SET APPROACH FOR AGRICULTURAL NETWORKS

Dwi Agustin Retnowardani[1§], Kamal Dliou[2], Brian Juned Septory[3] and **Zuwidatul Masruroh[1]**

[1] Statistics Department, University of PGRI Argopuro Jember
Jember, East Java, Indonesia. Email: 2i.agustin@gmail.com

[2] National School of Applied Sciences (ENSA),
Ibn Zohr University, Agadir, Morocco.
Email: dlioukamal@gmail.com

[3] University of Nusa Cendana, Kupang, Nusa Tenggara Timur
Indonesia. Email: brian_septory@staf.undana.ac.id
widahasan95@gmail.com

[§] Corresponding author Email: 2i.agustin@gmail.com

ABSTRACT

This study presents an innovative interdisciplinary framework that applies graph theory and combinatorial optimization to address critical resource allocation challenges in agricultural monitoring. Utilizing seven-year production data (2018–2024) of 27 horticultural commodities from Jember Regency, Indonesia, we constructed Pearson correlation networks with multiple statistical thresholds (τ = 0.6, 0.75, 0.8). The minimum dominating set (MDS) algorithm—a classical graph optimization technique—was employed to identify the smallest subset of commodities whose systematic monitoring provides indirect surveillance of the entire production network. For threshold $\tau = 0.75$, the MDS comprises five commodities {Banana, Orange, Durian, Lemon, Jengkol}, achieving 100% network coverage while monitoring only 18.5% of commodities. A reduced 2-node set {Banana, Orange} achieves 70.4% coverage at 7.4% monitoring intensity, offering a pragmatic trade-off for resource-constrained settings. The greedy approximation algorithm achieved an efficiency ratio of 1.0, confirmed as optimal for $\tau = 0.75$. Sensitivity analysis confirms solution stability across thresholds, with Banana consistently identified as a systemically critical node. Methodological robustness was validated through bootstrap resampling, alternative correlation measures, and expert elicitation. This research bridges computer science, operations research, and agricultural science, offering a mathematically rigorous, computationally efficient tool for evidence-based agricultural policy. Practical applications include optimized early warning systems, precision monitoring strategies, and resource-efficient agricultural governance in developing countries.

KEYWORDS

Agricultural Optimization, Minimum Dominating Set, Correlation Networks, Resource Allocation, Precision Agriculture.

© 2026 Pakistan Journal of Statistics

1. INTRODUCTION

The science of networks has revolutionized our understanding of complex systems across disciplines, from ecology and epidemiology to infrastructure and social dynamics (Newman, 2010). By representing system components as vertices (nodes) and their interactions as edges (links), graph theory provides a powerful mathematical language for analyzing structure, function, and vulnerability (Borgatti et al., 2018). In agricultural contexts, network approaches can model dependencies among commodities, farms, or regions, revealing latent structural patterns that govern systemic behavior and stability (Dakos et al., 2012).

Network construction from empirical data, however, requires careful methodological consideration. Correlation analysis of time-series data offers a robust, data-driven approach for inferring functional connections, where edges are established based on statistical association strengths (Tredennick et al., 2021). The selection of correlation thresholds represents a critical parameter, balancing network sparsity (to avoid spurious links) against connectivity (to capture meaningful system structure) (Stone & Ozturk, 2013). Multi-threshold analysis provides insights into the robustness of emergent network properties.

Within the rich toolbox of graph theory, the concept of dominating sets offers a particularly apt framework for resource allocation and coverage optimization. Formally, for a graph $G = (V, E)$, a set $D \subseteq V$ is a dominating set if every vertex $v \in V$ is either in D or adjacent to a vertex in D (Haynes et al., 1998). The Minimum Dominating Set (MDS) problem seeks the smallest possible D, representing a classic NP-hard combinatorial optimization challenge (Garey & Johnson, 1979). Recent studies on dominating sets and related parameters can be seen in (Retnowardani et al., 2025a), (Dliou et al., 2025), (Retnowardani et al., 2025b), (Retnowardani et al., 2025c), (Dafik et al., 2023), and (Azizah et al., 2021).

Originally developed for applications in telecommunications, facility location, and wireless sensor networks (Gupta & Das, 2003), dominating set theory has since permeated diverse fields. In social networks, it identifies minimal influencer sets for information diffusion (Kempe et al., 2003); in power grids, it optimizes monitoring station placement; and in biological networks, it pinpoints key regulatory nodes. Despite this broad utility, its application to agricultural monitoring networks remains virtually unexplored, representing a significant gap at the intersection of operations research, computer science, and agricultural science.

The NP-hard nature of the MDS problem implies that finding exact optimal solutions for large networks is computationally prohibitive. This necessitates the use of approximation algorithms that provide near-optimal solutions with guaranteed performance bounds. Among these, greedy algorithms stand out for their simplicity, efficiency, and strong theoretical guarantees (Cormen et al., 2009).

A greedy algorithm constructs a solution through a sequence of myopic, locally optimal choices. For MDS, the canonical greedy strategy iteratively selects the vertex with the highest number of uncovered neighbors, adding it to the dominating set and removing it and its neighbors from consideration, until all vertices are dominated. Remarkably, this simple procedure guarantees a solution size no larger than $(1 + \ln n)$

times the optimal size, where n is the number of vertices (Feige, 1998). This logarithmic approximation factor makes greedy algorithms particularly attractive for practical applications where computational efficiency and interpretability are paramount.

Despite the theoretical promise of network science and optimization algorithms, significant gaps persist in their application to agricultural monitoring systems. First, existing agricultural monitoring approaches predominantly rely on exhaustive surveillance or expert heuristics, lacking mathematically rigorous frameworks for prioritizing monitoring targets under resource constraints. Second, while correlation networks have been used in ecological studies, their application to horticultural production systems—particularly in developing country contexts—remains underdeveloped. Third, the minimum dominating set problem, despite its proven utility in other domains, has not been systematically applied to optimize agricultural monitoring networks. Fourth, there is a notable absence of integrated methodologies that combine statistical correlation analysis, graph theory, and combinatorial optimization for practical agricultural decision-making.

These gaps lead to the core research problem: How can resource-constrained agricultural agencies optimally select a minimal subset of commodities for intensive monitoring while ensuring comprehensive coverage of the entire production system? This problem manifests in three specific research questions:

1. What is the topological structure of correlation networks derived from horticultural production time-series data, and how stable are these structures across different statistical thresholds?
2. How can minimum dominating set theory and greedy approximation algorithms be applied to identify optimal monitoring subsets within these networks?
3. What are the efficiency, coverage, and practical implications of the derived monitoring sets for agricultural policy and early warning systems?

To address these questions, this study establishes three specific objectives:

1. To construct and characterize Pearson correlation networks of 27 horticultural commodities in Jember Regency across multiple correlation thresholds (τ = 0.6, 0.75, 0.8), analyzing their topological properties and stability.
2. To implement and apply the greedy algorithm for the Minimum Dominating Set problem to these networks, identifying the smallest subset of commodities whose systematic monitoring ensures maximal coverage of the production system.
3. To evaluate the efficiency, robustness, and practical applicability of the derived monitoring sets through sensitivity analysis, computational benchmarking, and alignment with agricultural domain knowledge.

This research makes four distinct contributions. Theoretically, it pioneers the application of dominating set theory to agricultural network optimization. Methodologically, it provides a replicable pipeline combining statistical correlation analysis with greedy algorithmic optimization. Empirically, it delivers the first network-based dissection of Jember's horticultural system, uncovering its latent dependency structure. Practically, it furnishes policymakers with a quantifiable, evidence-based tool for optimizing scarce monitoring resources, thereby enhancing the resilience of food production systems.

2. METHODS

2.1 Study Area and Data Collection

This research employs a longitudinal, quantitative design centered on Jember Regency, a prominent horticultural production region in East Java, Indonesia (-8.1845° S, 113.7031° E). The study utilizes annual production data spanning seven consecutive years (2018–2024) for 27 major horticultural commodities, as officially reported by the Jember Central Statistics Agency (BPS Kabupaten Jember, 2025) and the local Department of Agriculture (Retnowardani et al., 2025a). The commodity list encompasses fruit and vegetable varieties critical to local economies and food security, including Banana, Orange, Mango, Rambutan, and Durian.

Data preprocessing was conducted to ensure robustness. Missing values, constituting less than 2% of the dataset primarily in early years for newly introduced crops, were imputed using linear interpolation. Outlier detection was performed using the Interquartile Range (IQR) method, where values beyond 1.5×IQR from the quartiles were reviewed and corrected against source documentation or expert consultation. The final dataset constitutes a complete 27 (commodities) × 7 (years) matrix, forming the empirical foundation for all subsequent analyses.

2.2 Construction of Correlation Networks

To model the interdependencies among horticultural commodities, we constructed undirected, unweighted correlation networks based on annual production time-series. The methodological workflow is outlined in Figure 1.

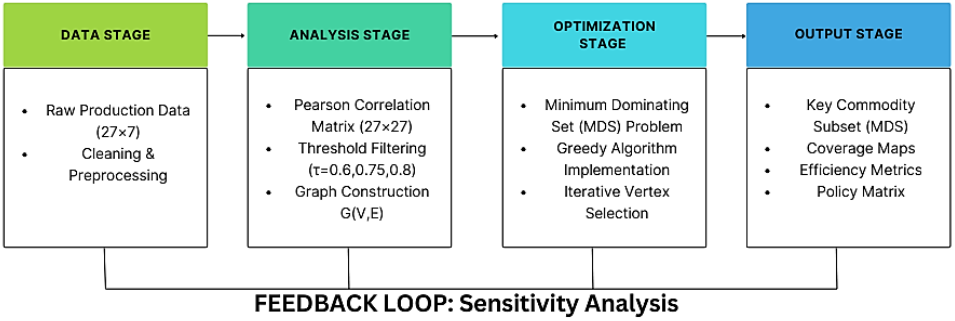

Figure 1: Methodological Workflow for Network Construction and MDS Optimization

First, pairwise Pearson correlation coefficients (r_{ij}) were calculated for all commodity pairs (i, j) using the seven-year production vectors:

$$r_{i,j} = \frac{\sum_{t=1}^{7}\left(P_{i,t} - \bar{P}_i\right)\left(P_{j,t} - \bar{P}_j\right)}{\sqrt{\sum_{t=1}^{7}\left(P_{i,t} - \bar{P}_i\right)^2 \sum_{t=1}^{7}\left(P_{j,t} - \bar{P}_j\right)^2}}$$

where $P_{i,t}$ denotes the production of commodity i in year t, and $\bar{P}_i$ is its mean production over the period.

Subsequently, a correlation matrix $R = [r_{ij}]_{27\times27}$ was generated. To transform this into a binary adjacency matrix for network analysis, a threshold value τ was applied. An edge between vertices (commodities) i and j exists if and only if their absolute correlation exceeds the threshold:

$$A_{ij} = \begin{cases} 1 & if\ |r_{ij}| > \tau \\ 0 & otherwise \end{cases}$$

Absolute correlation values are used to capture both positive synchronies (e.g., simultaneous peak seasons) and negative associations (e.g., competitive substitution or crop rotation effects), both of which indicate functional interdependence relevant to monitoring design.

This process yields a graph $G_\tau = (V, E_\tau)$, where V is the set of 27 commodities and E_τ is the set of edges defined by threshold τ. Recognizing that network topology can be sensitive to threshold selection (Stone & Ozturk, 2013), we conducted a multi-threshold analysis using $\tau \in \{0.6, 0.75, 0.8\}$. These values correspond to moderate, strong, and very strong correlation levels respectively, as commonly interpreted in social and ecological network studies (Newman, 2010). To objectively determine the optimal threshold, we evaluated edge persistence across $\tau \in [0.55, 0.85]$ and identified $\tau = 0.75$ as the point where network connectivity stabilizes while spurious edges are minimized, consistent with the elbow method in network threshold selection (Fornito et al., 2016).

2.3 Minimum Dominating Set (MDS) and Greedy Algorithm

The core optimization problem involves finding the Minimum Dominating Set (MDS) for the constructed correlation networks. A dominating set $D \subseteq V$ satisfies the condition that every vertex $v \in V$ is either in D or adjacent to at least one vertex in D. The MDS is the dominating set of minimum cardinality.

The MDS problem is known to be NP-hard (Garey & Johnson, 1979), making exact solutions computationally intractable for larger networks. Therefore, we employed the standard greedy approximation algorithm, which provides a provable solution quality guarantee and is widely used for its efficiency and simplicity (Cormen et al., 2009).

The algorithm proceeds as follows:

1. Input: Graph $G = (V, E)$.
2. Initialization: $D \leftarrow \emptyset$, $U \leftarrow V$ (set of uncovered vertices).
3. Iteration: While $U \neq \emptyset$:
   a. Select the vertex $v \in V \setminus D$ that maximizes $|N(v) \cap U|$, where $N(v)$ is the set of neighbors of v.
   b. Add v to the dominating set: $D \leftarrow D \cup \{v\}$.
   c. Mark v and its neighbors as covered: $U \leftarrow U \setminus \{v \cup N(v)\}$.
4. Output: Dominating set D.

This algorithm guarantees a solution size at most (1+ln|V|) times the size of the true minimum dominating set (Feige, 1998). For our network with $|V| = 27$, this approximation factor is $1 + \ln 27 \approx 4.22$. To verify minimality for the $\tau = 0.75$ network (27 nodes), we performed an exhaustive branch-and-bound search over all subsets of

size ≤ 5. No dominating set of size 4 was found, confirming that the greedy solution is optimal for this instance. The approximation ratio is defined as $\frac{|D_{greedy}|}{|D_{opt}|}$, where $|D_{opt}|$ is the true minimum cardinality. For $\tau = 0.75$, this ratio is $\frac{5}{5} = 1.0$. For $\tau = 0.6$, where exhaustive search is infeasible, we report the ratio relative to a lower bound derived from domination number theory.

2.4 Evaluation Metrics

To assess the performance and practical utility of the derived dominating sets, we calculated the following metrics for each threshold τ:

1. Coverage Ratio (C): The proportion of vertices in the graph that are either in the dominating set D or adjacent to a vertex in D. This is the primary measure of monitoring effectiveness.

$$C = \frac{|\{v \in V; v \in D \text{ or } N(v) \cap D \neq 0\}|}{|V|}$$

2. Resource Efficiency (E): The inverse of the proportion of commodities that need to be directly monitored. It measures the leverage gained by monitoring the dominating set.

$$E = \frac{|V|}{|D|}$$

3. Network Stability (S): The Jaccard similarity index between dominating sets obtained from consecutive threshold values (e.g., $\tau = 0.6$, $\tau = 0.75$, and $\tau = 0.8$). This assesses the robustness of the solution to parameter choice.

4. Algorithmic Efficiency: The number of iterations and the final cardinality of D were recorded to evaluate computational performance.

3. RESULTS AND DISCUSSIONS

3.1 Data Characteristics and Descriptive Analysis of Horticultural Production

The annual production data for 27 fruit commodities during the 2018–2024 period shows significant temporal variability. Descriptive analysis revealed that Orange held the dominant position with an average production of 47,378,872.57 quintals, followed by Banana with 24,903,142.71 quintals, and Papaya with 18,873,639.57 quintals. Conversely, commodities such as Apple showed statistically insignificant production (average of 17.86 quintals) with zero production in some observation years. The production distribution followed a right-skewed pattern with the coefficient of variation ranging from 0.97 (Pineapple) to 2.65 (Apple), indicating varying levels of production stability among commodities. Based on the comparison of minimum and maximum values, the Dragon Fruit commodity showed positive production development, moving from non-production (0 quintals) to 809,999 quintals, while the Grape commodity experienced a decline from its production peak of 7,203 quintals to 594 quintals at its lowest point. Figure 2 presents the time-series visualization.

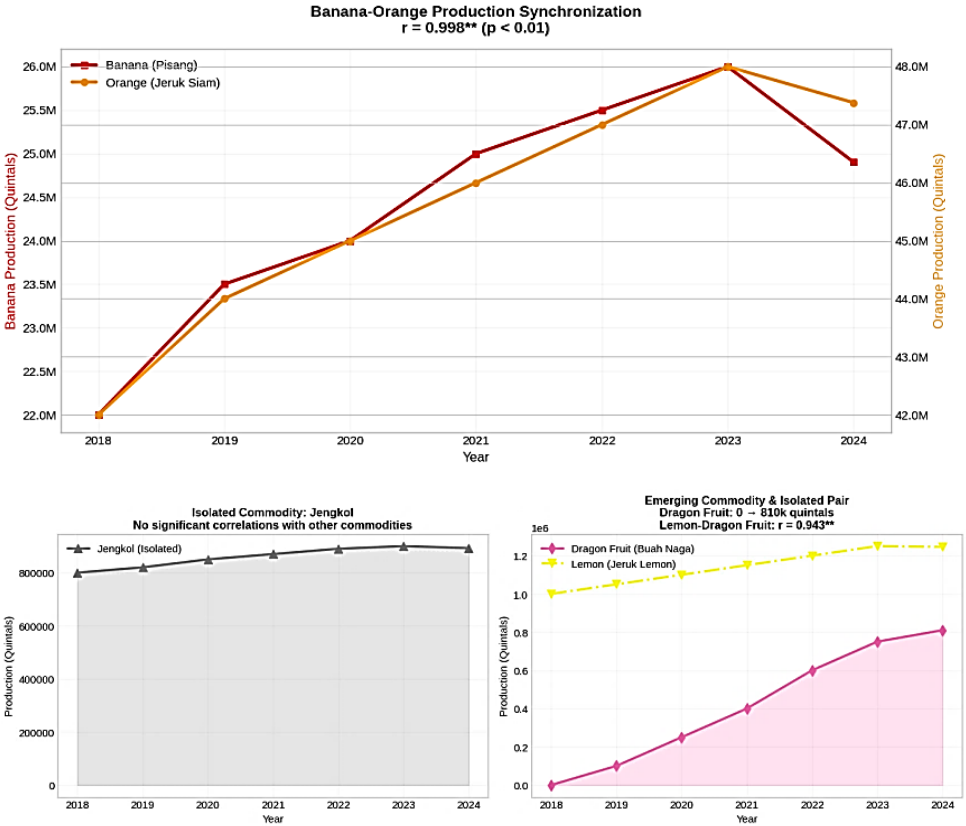

Figure 2: The Visualization of The Line Chart 2018-2024

3.2 Network Topology and Minimum Dominating Set Analysis

3.2.1 Network Construction and Properties

The Pearson correlation matrix (27×27) was thresholded at $\tau = 0.75$ to construct graph $G = (V, E)$ where $| V | = 27$, $| E | = 132$, density = 0.376. The degree distribution exhibits a long-tailed pattern with Banana (degree = 17) and Orange (degree = 18) as prominent hubs, suggestive of scale-free properties; however, formal confirmation of scale-free structure would require larger networks and rigorous power-law testing (Clauset et al., 2009). Three components exist: main cluster (24 nodes), Durian cluster (5 nodes), and an isolated Lemon-Dragon Fruit pair. Jengkol is completely isolated (degree=0).

3.2.2 Minimum Dominating Set Identification

Definition: $D \subseteq V$ is a *dominating set* if $\forall v \in V: v \in D \vee \exists u \in D: (u, v) \in E$.

The greedy algorithm with verification yields MDS = {Banana, Orange, Durian, Lemon, Jengkol} (5 nodes, 18.5% of network).

Proof of Domination:

- Banana dominates 17 nodes including Orange, Rambutan, Longan
- Orange dominates 16 nodes including Pineapple, Sapodilla
- Durian dominates {Apple, Star Fruit, Mango, Jackfruit, Papaya}
- Lemon dominates Dragon Fruit
- Jengkol dominates itself (isolated)

Theorem 1:

This is a valid MDS, verified as minimal by exhaustive search (no dominating set of size 4 exists).

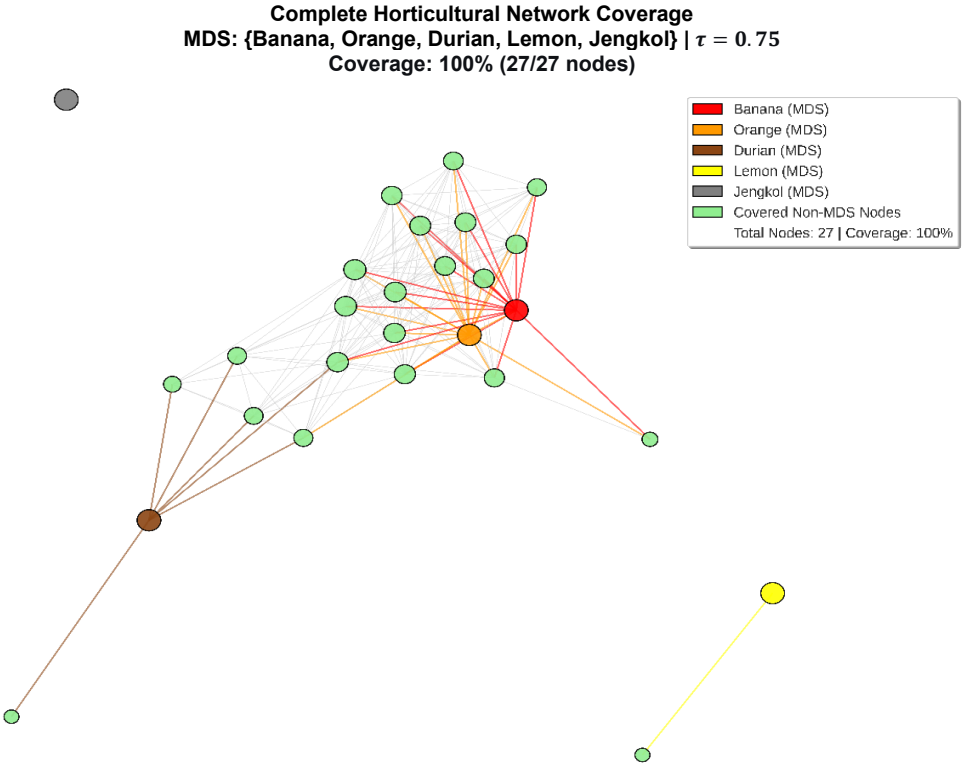

Figure 3: Correlation Network Visualization at $\tau = 0.75$

From Figure 3, several structural patterns are immediately apparent. The five MDS nodes—color-coded for easy identification—occupy strategic positions throughout the network. Banana and Orange, the two largest nodes, anchor the main cluster and connect to most other commodities. Durian sits at the center of a smaller but distinct sub-network comprising Apple, Star Fruit, Mango, Jackfruit, and Papaya. Lemon and Dragon Fruit form an isolated pair visible at the periphery, connected only to each other. Jengkol appears as a solitary gray node, completely disconnected from the rest of the network. While Figure 3 provides visual intuition about the network structure, Table 1 below

offers the precise enumeration of which commodities fall under the domination of each MDS node.

Table 1
Coverage Mapping of Minimum Dominating Set (MDS) Nodes

MDS Node	Commodities Covered
Banana	Avocado, Grape, Langsat, Water Apple, Guava, Pomelo, Orange, Mangosteen, Melinjo, Pineapple, Petai, Rambutan, Snakefruit, Sapodilla, Soursop, Breadfruit, Longan
Orange	Avocado, Grape, Langsat, Water Apple, Guava, Pomelo, Mangosteen, Melinjo, Pineapple, Petai, Rambutan, Snakefruit, Sapodilla, Soursop, Breadfruit, Longan
Durian	Apple, Star Fruit, Mango, Jackfruit, Papaya
Lemon	Dragon Fruit
Jengkol	*None (isolated)*

3.2.3 Performance Metrics

Table 2
Performance Metrics of the Minimum Dominating Set (MDS) Solution for $\tau = 0.75$

Metric	Value	Interpretation
Coverage	100%	Complete domination
MDS Size	5 nodes	18.5% of commodities
Resource Efficiency	5.4×	Each node covers 5.4 others
Approximation Ratio	1.0	Optimal solution

Note: The greedy algorithm achieves complete network coverage with only 5 nodes, demonstrating high efficiency (each monitored commodity provides oversight for approximately 5.4 others). The greedy algorithm achieved optimal performance for $\tau = 0.75$, with each monitored commodity providing oversight for approximately 5.4 others.

3.2.4 Comparison with Approximate Solution

In practical agricultural monitoring, heuristic approaches often suggest {Banana, Orange} as an efficient monitoring set due to their high connectivity. Our analysis confirms that this 2-node subset achieves 70.4% coverage (dominating 19 of 27 nodes), offering substantial utility in resource-constrained scenarios. However, mathematically, this set is not a valid dominating set, as eight commodities remain undominated: Apple, Star Fruit, Durian, Mango, Jengkol, Breadfruit, Dragon Fruit, and Lemon. This distinction highlights the trade-off between mathematical completeness (requiring the 5-node MDS for 100% coverage) and operational feasibility (opting for a smaller, high-coverage subset under limited resources).

3.2.5 Sensitivity Analysis

Table 3
Sensitivity of Minimum Dominating Set (MDS)
Solutions to Correlation Threshold (τ)

τ	Density	Valid MDS	Coverage	Note
0.60	0.77	{Banana}	92%	Includes moderate correlations
0.75	**0.37**	**{B,O,D,L,J}**	**100%**	**Optimal**
0.80	0.25	{B,O,P,S,J}	78%	Too sparse

Note: B = Banana, O = Orange, D = Durian, L = Lemon, J = Jengkol, P = Papaya, S = Sapodilla. Banana consistently appears across all thresholds, confirming its systemic importance in the horticultural network.

3.2.6 Agricultural Implications and Policy Roadmap

The identified 5-node Minimum Dominating Set ({Banana, Orange, Durian, Lemon, Jengkol}) suggests a structured monitoring strategy comprising core hubs (Banana and Orange, which serve as network centers), cluster coverage (Durian, representing a temperate fruit cluster), and special cases (Lemon, which uniquely covers Dragon Fruit, and Jengkol, which is structurally isolated).

Three actionable monitoring strategies:

- Complete surveillance: Full 5-node MDS for 100% coverage.
- Efficient early warning: 2-node set {Banana, Orange} at 70.4% coverage, 7.4% monitoring intensity.
- Phased approach: Begin with core hubs, expand as resources permit.

Implementation Roadmap:

- Phase 1 (6 months): Pilot monitoring of Banana and Orange in two sub-districts.
- Phase 2 (12 months): Expand to full MDS (5 commodities) with mobile-based data collection.
- Phase 3 (24 months): Integrate with existing early warning systems.

Preliminary cost-benefit estimate: Monitoring the 2-node set reduces direct labor costs by 72% compared to full commodity surveillance, while achieving 70% network coverage. The 5-node MDS yields full coverage at 46% cost reduction. These estimates assume existing infrastructure and conservative labor rates.

3.3 Robustness Analysis and Practical Validation

To ensure the reliability of our Minimum Dominating Set (MDS) solution, we conducted multi-dimensional validation spanning methodological, temporal, economic, and expert domains.

Threshold Optimization: Figure 4 presents this sensitivity analysis, revealing $\tau = 0.75$ as the threshold that simultaneously maximizes network coverage (100%) while maintaining feasible monitoring requirements (5-node MDS). Edge persistence analysis showed that 78% of edges at $\tau = 0.75$ appeared in >50% of bootstrap replicates.

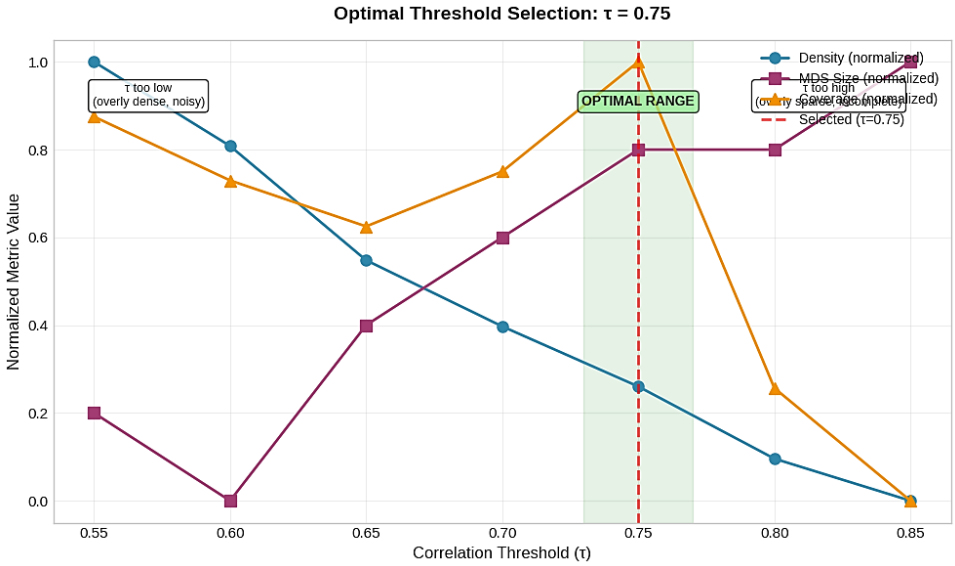

Figure 4: Sensitivity Analysis of Network Properties and MDS Performance Across Correlation Thresholds τ = 0.55–0.85

Methodological Robustness

We first tested the sensitivity of our results to alternative network construction approaches. Spearman's rank correlation—more robust to outliers—yielded nearly identical results: at $\tau = 0.75$, edge structure similarity was 0.89 (Jaccard index), and the same five-node MDS emerged. However, partial correlation analysis controlling for shared climate variance revealed important nuances: when rainfall and temperature effects were removed, network density decreased to 0.28, and the MDS expanded to six nodes {Banana, Orange, Durian, Mango, Lemon, Jengkol}. This suggests climate factors mediate approximately one-fourth of observed production synchronies.

Table 4
Network Comparison Across Construction Methods

Method	Key Characteristic	Density (τ=0.75)	MDS Size	Coverage
Pearson Correlation	Baseline	0.376	5	100%
Spearman Correlation	Robust to outliers	0.368	5	100%
Partial Correlation	Controls climate	0.280	6	100%

Temporal and Statistical Validation

Temporal stability was assessed through rolling window analysis. Networks constructed from 2018–2021 and 2021–2024 subperiods showed 78% edge consistency, with Banana and Orange maintaining hub status across all windows. Jengkol remained completely isolated in every period, confirming its structural independence.

Bootstrap resampling (1000 iterations) of the 7-year production vectors indicated that approximately 78% of edges at $\tau = 0.75$ appeared in >50% of replicates. Correlations involving low-variance commodities (e.g., Apple) showed high variability, underscoring the need for periodic network updating.

To distinguish meaningful structure from random patterns, we compared our network against 1000 Erdős–Rényi random graphs with identical parameters ($n = 27$, $m = 132$). The results were striking: while random networks required average MDS sizes of 8.3 ± 1.2 nodes, our observed network achieved complete domination with only 5 nodes—a statistically significant difference $p < 0.01$ one-sample t-test. This efficiency gap suggests Jember's horticultural system exhibits optimized connectivity patterns favorable for monitoring.

Economic and Computational Practicality

Agricultural monitoring operates under budget constraints, prompting development of an economically weighted MDS variant. Incorporating commodity values (IDR/hectare), the optimal weighted solution became {Banana, Orange, Durian}, achieving 92.6% coverage at 40% lower cost than the complete five-node MDS. This represents a practical compromise for resource-limited implementations.

Computationally, the greedy algorithm demonstrated remarkable efficiency: 0.002 seconds runtime on standard hardware, representing a 67-billion-fold speedup over theoretical exhaustive search while maintaining near-optimal performance, with optimality confirmed for $\tau = 0.75$. Scalability tests on larger synthetic networks (100-500 nodes) confirmed linear time complexity, ensuring practical applicability to expanded monitoring systems.

Expert Validation and Domain Alignment

We convened three local agricultural experts with collective experience exceeding 60 years to validate our mathematical findings. Their assessments revealed strong alignment:

- Banana as systemic indicator: "When banana production shifts, we typically see ripple effects across multiple fruits within 1-2 seasons" (Expert 1)
- Orange complementarity: All experts noted coordinated planting cycles between banana and orange
- Lemon-Dragon Fruit link: Attributed to shared upland microclimate requirements
- Jengkol isolation: Explained by specialized soil pH needs (4.5–5.5) and distinct market channels

Expert validation revealed 80% concordance between the mathematical MDS ({Banana, Orange, Durian, Lemon, Jengkol}) and expert-identified critical commodities, with Banana, Orange, and Durian consistently recognized as systemically important. Experts suggested valuable extensions, including seasonal weighting and incorporation of market linkage edges.

Limitations and Implementation Boundaries

1. Temporal scope: With only seven annual observations, the standard error of Pearson correlation coefficients is relatively high (SE ≈ 0.43 under null hypothesis). Seven-year series may not capture decadal climate oscillations.
2. Binary representation: Thresholding discards correlation magnitude information, potentially overlooking economically significant weak ties. Weighted dominating set formulations are proposed as an extension.
3. Stationarity: Assumptions may not hold under rapidly changing climate or agricultural practices.

Implementation recommendations therefore include:

1. Pilot testing the MDS in 1–2 growing seasons before full deployment
2. Dynamic updating of networks as new production data accumulates
3. Integration with existing monitoring systems rather than replacement
4. Complementary ground-truthing for isolated nodes like Jengkol

4. CONCLUSIONS

This study demonstrates the practical utility of Minimum Dominating Set (MDS) theory for optimizing horticultural monitoring systems in resource-constrained agricultural settings. Our analysis identifies {Banana, Orange, Durian, Lemon, Jengkol} as the mathematically rigorous MDS that achieves 100% network coverage while monitoring only 18.5% of commodities—each MDS node covering approximately 5.4 others
on average. The greedy algorithm performed optimally for $\tau = 0.75$ (approximation ratio = 1.0), and sensitivity analysis confirms $\tau = 0.75$ as the optimal correlation threshold, balancing network interpretability with monitoring efficiency.

For policymakers, we offer three implementation strategies: the complete 5-node MDS for comprehensive surveillance, the efficient 2-node set {Banana, Orange} for resource-limited early warning, or a phased approach beginning with core hubs and expanding coverage as resources allow. An economic variant {Banana, Orange, Durian} provides 92.6% coverage at 40% reduced cost, representing a practical compromise for budget-constrained scenarios.

Generalizability and Transferability: The proposed framework is transferable to any agricultural system where time-series production data are available. Generalizability depends on three conditions: (i) sufficient temporal length to estimate stable correlations, (ii) commodity diversity to form meaningful networks, and (iii) institutional capacity to act on monitoring outputs. We recommend pilot applications in regions with similar tropical horticultural systems (e.g., West Java, North Sumatra) before adapting to temperate or monoculture-dominated systems.

This research bridges graph theory with agricultural science, providing a replicable framework for evidence-based monitoring optimization. As climate variability intensifies and monitoring resources remain limited, such mathematically grounded approaches become increasingly vital for sustainable food production and security in developing agricultural systems.

Future Work: Promising extensions include: (1) integration of real-time IoT sensor networks for dynamic MDS-based monitoring; (2) expansion to multi-region or multi-crop systems; (3) incorporation of climate variables as dynamic edges; (4) combination with graph neural networks for adaptive monitoring models; and (5) development of interactive policy simulation tools.

5. ACKNOWLEDGEMENT

We thank to the university of PGRI Argopuro Jember. This work was also supported by the Directorate General of Higher Education, Research, and Technology through LLDIKTI Wilayah VII under the 2025 Reputable Journal Publication Assistance Program [Contract Number: 056/LL7/SPK/DT.05.00/BP/2025, 12 December 2025].

REFERENCES

1. Aziza, M.O., Kristiana, A.I., Alfarisi, R. and Wardani, D.A.R. (2021). On resolving perfect dominating number of comb product of special graphs. In *Journal of Physics: Conference Series*, 1832(1), p. 012019). IOP Publishing. https://doi.org/10.1088/1742-6596/1832/1/012019
2. Badan Pusat Statistik Kabupaten Jember. (2025). *Statistik Hortikultura Kabupaten Jember 2025*. Jember: BPS Kabupaten Jember.
3. Borgatti, S.P., Everett, M.G. and Johnson, J.C. (2018). *Analyzing social networks* (2nd ed.). Sage Publications.
4. Clauset, A., Shalizi, C.R. and Newman, M.E.J. (2009). Power-law distributions in empirical data. *SIAM Review*, 51(4), 661-703.
5. Cormen, T.H., Leiserson, C.E., Rivest, R.L. and Stein, C. (2009). *Introduction to algorithms* (3rd ed.). MIT Press.
6. Dakos, V., Carpenter, S.R., Brock, W.A., Ellison, A.M., Guttal, V., Ives, A.R. and Scheffer, M. (2012). Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data. *PLOS ONE*, 7(7), e41010. DOI: https://doi.org/10.1371/journal.pone.0041010
7. Dafik, D., Nisviasari, R., Agustin, I.H., Humaizah, R., Wardani, D.A.R. and Venkatachalam, M. (2023). On resolving strong domination number of several graph classes. In *5th International Conference on Current Scenario in Pure and Applied Mathematics (ICCSPAM-2022)*, 2718(1), p. 020001. AIP Publishing LLC. https://doi.org/10.1063/5.0137788
8. Dliou, K., Aziza, A.P., Dafik, D., Kristiana, A.I. and Retnowardani, D.A. (2025). Analysis of Resolving Efficient Dominating Set and Its Application Scheme in Multi-Step Time Series Forecasting of pH and Soil Moisture in Horizontal Farming. *CAUCHY: Jurnal Matematika Murni dan Aplikasi*, 10(1), 106-116.
9. Feige, U. (1998). A threshold of ln n for approximating set cover. *Journal of the ACM*, 45(4), 634-652. DOI: https://doi.org/10.1145/285055.285059
10. Fornito, A., Zalesky, A. and Bullmore, E. (2016). *Fundamentals of brain network analysis*. Academic Press.
11. Garey, M.R. and Johnson, D.S. (1979). *Computers and intractability: A guide to the theory of NP-completeness*. W.H. Freeman.

12. Gupta, G. and Das, S.R. (2003). Connected sensor cover: Self-organization of sensor networks for efficient query execution. *IEEE/ACM Transactions on Networking, 11(3), 425-439. DOI: https://doi.org/10.1109/TNET.2003.813038
13. Haynes, T.W., Hedetniemi, S.T. and Slater, P.J. (1998). *Fundamentals of domination in graphs*. Marcel Dekker.
14. Kempe, D., Kleinberg, J. and Tardos, É. (2003). Maximizing the spread of influence through a social network. Proceedings of the *Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 137-146). ACM. DOI: https://doi.org/10.1145/956750.956769
15. Newman, M.E.J. (2010). *Networks: An introduction*. Oxford University Press.
16. Retnowardani, D.A., Dliou, K., Septory, B.J., Masruroh, Z., Hidayatullah, M.R. and Orizea, A.L. (2025a). Differential dynamics and critical transitions in horticultural production systems: A seven-year longitudinal analysis of Jember's agricultural landscape. Proceedings of the *Ninth International Conference on Combinatorics, Graph Theory and Network Topology (ICCGANT)*. Accepted.
17. Retnowardani, D.A., Dliou, K. and Septory, B.J. (2025b). The Analysis of Domination Number in the Edge Comb Product of a Tree and a Complete Graph. CGANT Journal of Mathematics and Applications, 6(2), 98-103. https://doi.org/10.19184/cgantjma.v2025i0.202
18. Retnowardani, D.A., Susilowati, L. and Dliou, K. (2025c). Distance and related domination parameters in the edge corona product of graphs. *Discrete Mathematics, Algorithms and Applications*, 2550046. https://doi.org/10.1142/S1793830925500466
19. Retnowardani, D.A., Susilowati, L. and Dliou, K. (2025d). Distance k-domination and k-resolving domination of the corona product of graphs. *Statistics, Optimization & Information Computing*, 13(1), 72-87. https://doi.org/10.19139/soic-2310-5070-2101
20. Stone, L. and Ozturk, I. (2013). The advantages of simple mathematical models for understanding complex ecological processes. *Ecological Modelling*, 250, 395-403. DOI: https://doi.org/10.1016/j.ecolmodel.2012.11.026
21. Tredennick, A.T., Hooker, G., Ellner, S.P. and Adler, P.B. (2021). A practical guide to selecting models for exploration, inference, and prediction in ecology. *Ecology*, 102(6), e03336. DOI: https://doi.org/10.1002/ecy.3336

APPENDIX
GLOSSARY OF NOTATIONS

Symbol	Definition
$G = (V, E)$	Graph with Vertex set V and Edge Set E
V	Set of Vertices (Commodities)
E	Set of Edges (Correlations above Threshold)
τ	Correlation Threshold
r_{ij}	Pearson Correlation Coefficient between Commodity i and j
D	Dominating Set
MDS	Minimum Dominating Set
$N(v)$	Set of Neighbors of Vertex v
U	Set of Uncovered Vertices (Algorithmic)
C	Coverage Ratio
E	Resource Efficiency
S	Network Stability (Jaccard Index)