Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 59 Issue: 07:2026
DOI: 10.5281/zenodo.21718225

ADAPTIVE CLIMATE STRESS INDEX (ACSI): AN EXPLAINABLE MULTIMODAL SENSOR FUSION FRAMEWORK FOR SOIL HEALTH ASSESSMENT AND CLIMATE-INDUCED STRESS PREDICTION AND RISK CLASSIFICATION

V. RAJAGOPALAN*
Assistant Professor, Department of Civil Engineering, University College of Engineering, Anna University, Tiruchirappalli, Tamil Nadu, India. *Corresponding Author Email: drvrgaut@gmail.com

S. SENTHILKUMAR
Assistant Professor, Department of Computer Science Engineering, University College of Engineering, Anna University, Tiruchirappalli, Tamil Nadu, India.

S. MOHAMED RIYAS
Project Assistant, Department of Civil Engineering, University College of Engineering, Anna University, Tiruchirappalli, Tamil Nadu, India.

Abstract

Most precision-agriculture systems that rely on soil and climate sensor data feed those raw readings straight into a machine learning model, leaving the model to work out on its own how temperature, moisture and nutrient levels combine into something resembling environmental stress. We take a different approach here. Working with the standard multimodal inputs used in soil health monitoring — nitrogen, phosphorus, potassium, temperature, humidity, pH and rainfall — we introduce an engineered feature we call the Adaptive Climate Stress Index (ACSI), computed before any classification takes place. ACSI folds four climate-sensitive variables together with an NPK-balance term into a single stress score bounded between 0 and 1, using a deviation-from-ideal normalization scheme. That score, combined with the raw sensor features, feeds into a continuous 0–100 Soil Health Score and a three-tier Climate Risk Level, both of which we predict using Random Forest and Decision Tree classifiers and then interpret with SHAP (SHapley Additive exPlanations). Tested on a public crop-recommendation dataset of 2,200 samples, the pipeline reaches up to 97.95% accuracy for Soil Health classification and 97.73% for Climate Risk classification, and SHAP analysis confirms that ACSI is, by a wide margin, the most influential feature driving the climate-risk predictions. What this shows is that a fairly simple, domain-motivated index can turn climate stress into something explicit, learnable and explainable — without needing any sensors beyond what standard agricultural datasets already provide.

Index Terms: Soil Health Monitoring, Climate Stress Index, Precision Agriculture, Explainable AI, SHAP, Random Forest, Multimodal Sensors, Feature Engineering.

1. INTRODUCTION

Modern agricultural decision-support systems draw on a wide mix of sensor data — soil chemistry readings for nitrogen, phosphorus and potassium, atmospheric measurements of temperature, humidity and rainfall, and soil pH. In most pipelines built on this kind of data, the raw variables are simply handed to a machine learning model as-is, and the model is left to figure out for itself how a particular combination of heat, moisture and rainfall adds up to agricultural stress. Our view is that this leaves something on the table:

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 59 Issue: 07:2026
DOI: 10.5281/zenodo.21718225

an explicit, interpretable intermediate representation of climate stress would make soil health predictions both more transparent and more diagnostically useful.

To that end, we propose the Adaptive Climate Stress Index (ACSI). It measures how far a given sensor reading strays from an agronomically comfortable range, and aggregates that across temperature, rainfall, humidity, pH and nutrient balance into one number. Rather than treating soil health as a black box that a classifier either gets right or wrong, we build it up deliberately — combining ACSI with nutrient sufficiency into a continuous, human-readable Soil Health Score, and separately deriving a Climate Risk Level that captures how exposed a given plot is to climate stress, tempered by how much buffering capacity its soil provides.

This work makes four contributions. First, the ACSI formulation itself — a bounded, weighted, deviation-based index that needs nothing beyond the sensors already present in standard agricultural datasets. Second, a Soil Health Score and Climate Risk Level scheme that is derived transparently from ACSI and the raw features, rather than learned as an opaque target. Third, a comparative evaluation of Random Forest and Decision Tree classifiers on both derived targets. And fourth, a SHAP-based explainability analysis that actually verifies ACSI's contribution to the model's decisions instead of just asserting that it matters.

2. LITERATURE SURVEY

Much of the existing work on crop recommendation and soil analysis feeds raw soil and climate features straight into a classifier. This tends to produce strong accuracy numbers including on the same public crop-recommendation dataset used in this study — but it offers little insight into why the model reached a particular recommendation. Explainable AI methods, SHAP in particular, have found their way into agricultural machine learning as a way of surfacing feature importance after the fact, but they are rarely used to validate a feature that was purpose-built in the first place. Composite environmental stress indices, meanwhile, are old news in climatology and agronomy — heat stress indices and drought indices being familiar examples — yet they are seldom folded into a supervised learning pipeline as a first-class engineered feature for soil health classification specifically. What sets this paper apart is that the stress index itself becomes the object of study: we are not simply classifying soil health, we are using SHAP to show that the engineered index is the actual mechanism the classifier relies on when reasoning about climate risk.

3. PROPOSED METHODOLOGY

The pipeline unfolds in five stages. Raw multimodal sensor features are collected first. The Adaptive Climate Stress Index is then computed from those readings. ACSI and the raw features together feed a Random Forest classifier, which we compare against a Decision Tree baseline. SHAP is applied to explain the resulting predictions. Finally, the system reports a Soil Health Score, a Soil Health Class, a Climate Risk Level, and the feature importance behind each individual prediction.

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 59 Issue: 07:2026
DOI: 10.5281/zenodo.21718225

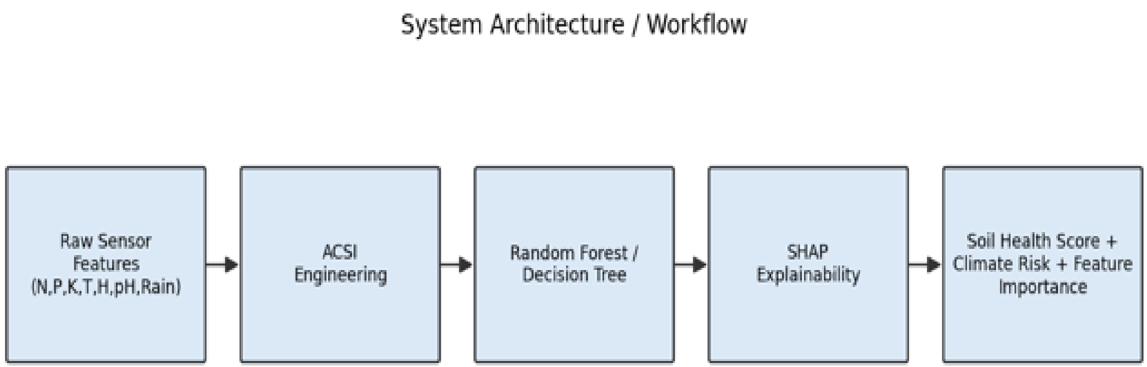

Figure1: Proposed System Architecture / Workflow

4. ADAPTIVE CLIMATE STRESS INDEX (ACSI)

ACSI draws on four climate-sensitive variables — temperature, humidity, rainfall and pH — along with an NPK-balance term, each of which is converted into a deviation-from-ideal score bounded between 0 and 1. Given a variable x with an agronomically ideal range [low, high], the deviation score works out as:

$$d(x) = 0, \text{ if } low \leq x \leq high$$

$$d(x) = (low - x) / (low - min\_obs), \text{ if } x < low$$

$$d(x) = (x - high) / (max\_obs - high), \text{ if } x > high$$

Here min_obs and max_obs refer to the observed minimum and maximum of that variable across the dataset. We settled on the following ideal bands after cross-checking agronomic norms against the observed data distribution: temperature 20–30 °C, humidity 60–90%, rainfall 60–150 mm, and pH 6.0–7.5.

The NPK-balance term is the coefficient of variation across nitrogen, phosphorus and potassium for each sample, min-max normalized across the dataset — a high value points to uneven nutrient distribution, which we treat as a mild amplifier of stress.

ACSI is then a weighted sum of these five deviation scores:

$$ACSI = wT \cdot d(T) + wR \cdot d(R) + wH \cdot d(H) + wpH \cdot d(pH) + wN \cdot d(NPK)$$

We set wT = 0.30, wR = 0.25, wH = 0.20, wpH = 0.15 and wN = 0.10, which sum to 1.0 and keep ACSI within [0,1]. Temperature and rainfall carry the most weight, since they are the dominant drivers of agricultural climate stress, while pH and NPK balance play a supporting role.

These weights were fixed ahead of time based on agronomic reasoning rather than fitted to the data — a deliberate choice that keeps ACSI an interpretable, domain-motivated feature rather than something closer to a learned embedding.

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 59 Issue: 07:2026
DOI: 10.5281/zenodo.21718225

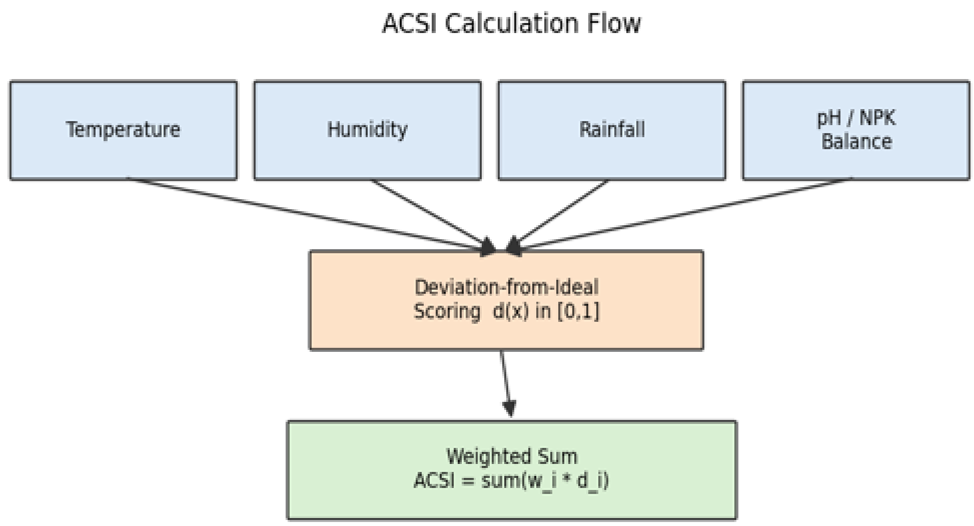

Figure 2: ACSI Calculation Flow

5. SOIL HEALTH SCORE AND CLIMATE RISK LEVEL

5.1 Soil Health Score

The Soil Health Score, on a 0–100 scale, brings together nutrient sufficiency, pH suitability and the climate resilience implicit in ACSI:

$$SHS = 0.5 \cdot NutrientScore + 0.3 \cdot pHScore + 0.2 \cdot (1 - ACSI) \cdot 100$$

NutrientScore is simply the mean of min-max-normalized N, P and K, rescaled to 0–100. pHScore sits at 100 within the ideal pH band and falls off with deviation using the same logic used for ACSI's deviation scores. We classify the resulting score as Healthy at 70 or above, Moderate between 40 and 69, and Poor below 40.

5.2 Climate Risk Level

Climate Risk Level brings ACSI, standing in for environmental stress, together with the Soil Health Score, standing in for the soil's buffering capacity. A plot is flagged High Risk if ACSI exceeds 0.30 or SHS drops below 40; it qualifies as Low Risk only if ACSI is 0.15 or under and SHS is 60 or above; everything else falls into Moderate Risk. Built this way, a plot with genuinely strong soil health can partly offset climate stress, and the reverse holds too — which mirrors, in a rough quantitative way, how agronomists already reason about vulnerability.

6. EXPERIMENTAL SETUP

We used the public Crop Recommendation dataset, which contains 2,200 samples across 22 crop labels, each carrying nitrogen, phosphorus, potassium, temperature, humidity, and pH and rainfall readings. Since the data had no missing values and no duplicate rows to begin with, preprocessing amounted to label-encoding the derived target classes and

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 59 Issue: 07:2026
DOI: 10.5281/zenodo.21718225

standardizing all eight input features — the original seven plus ACSI. We split the data 80:20 with stratification on the target class, which keeps the relative frequency of minority classes, Poor soil health being the obvious example, consistent across both partitions. For each target we trained two classifiers, a Random Forest with 200 estimators and a Decision Tree capped at a max depth of 10, and evaluated both using accuracy, weighted precision, weighted recall, weighted F1-score, and confusion matrices.

7. RESULTS

Table 1: Classification Performance on the Held-Out Test Set (N = 440)

Target	Model	Accuracy	Precision	Recall	F1
SoilHealthClass	Random Forest	97.95%	97.98%	97.95%	97.81%
SoilHealthClass	Decision Tree	97.95%	97.98%	97.95%	97.96%
ClimateRiskLevel	Random Forest	97.50%	97.55%	97.50%	97.50%
ClimateRiskLevel	Decision Tree	97.73%	97.73%	97.73%	97.73%

Both classifiers land in roughly the same high-performance territory on both targets, which is exactly what we'd expect and worth spelling out clearly: SoilHealthClass and ClimateRiskLevel are deterministic functions of ACSI, nutrient levels and pH, not independently observed ground truth. So near-ceiling accuracy here really just shows that both models successfully recovered the engineered threshold structure, rather than having cracked some open-ended prediction problem against noisy real-world labels. It also explains why the Decision Tree matches or even edges out Random Forest on ClimateRiskLevel — the underlying labels come from simple threshold rules on ACSI and SHS, a boundary that a single shallow tree can already represent well, leaving Random Forest's usual advantage of variance reduction through ensembling with little room to show up. That said, we would still recommend Random Forest for deployment, given its likely robustness to sensor noise that simply isn't present in this clean benchmark dataset.

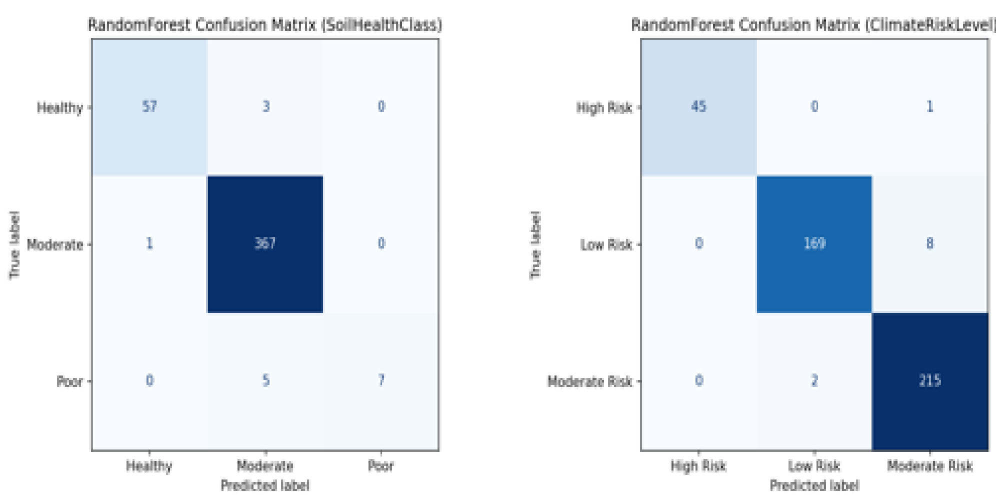

Figure 3: Confusion Matrices — Random Forest, Soil Health Class (Left) and Climate Risk Level (Right)

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 59 Issue: 07:2026
DOI: 10.5281/zenodo.21718225

8. EXPLAINABLE AI ANALYSIS

We applied SHAP, using TreeExplainer, to both Random Forest models to check what was actually driving their predictions. For ClimateRiskLevel, ACSI comes out on top by a clear margin (mean |SHAP| = 0.163), nearly three times the value of the next feature, Nitrogen, at 0.114 — solid confirmation that the engineered index is the main channel through which the model reasons about climate risk.

For SoilHealthClass, it's phosphorus and potassium that dominate instead, which lines up with their combined 50% weight in the Soil Health Score formula; ACSI comes in fifth here but still makes a meaningful contribution.

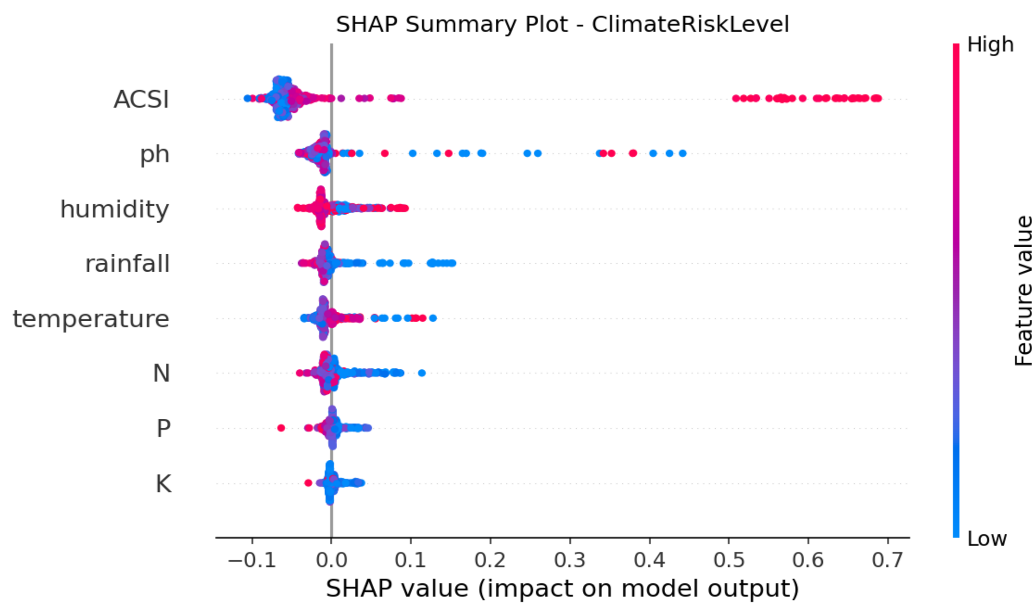

Figure 4: SHAP Summary Plot For Climate Risk Level — ACSI Is the Top-Ranked Feature

9. DISCUSSION

The main finding here is fairly simple to state: a straightforward, agronomically motivated engineered feature like ACSI can be verified, not just claimed, as the primary driver behind climate-risk predictions, and SHAP is what lets us make that case quantitatively rather than by assertion. That closes an important loop between feature engineering and explainability — instead of proposing ACSI and hoping it turns out to matter, we can actually show that it does.

One limitation worth being upfront about is that both derived targets are built from the same features used to predict them, so the accuracies reported here validate the internal consistency of our engineered scoring scheme rather than agreement with some external, independently labelled ground truth. The natural next step is to validate ACSI and the resulting classes against agronomist assessments or actual observed crop yields.

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 59 Issue: 07:2026
DOI: 10.5281/zenodo.21718225

10. CONCLUSION AND FUTURE SCOPE

In this paper we introduced the Adaptive Climate Stress Index, a bounded and interpretable engineered feature that compresses multimodal climate and soil sensor data into a single stress score, and built on it to produce a transparent Soil Health Score and Climate Risk Level. Both Random Forest and Decision Tree classifiers recover the engineered class structure with high accuracy, up to 97.95%, and SHAP analysis backs up ACSI's role as the leading predictor of climate risk. Looking ahead, the priorities are validating ACSI against real agronomist assessments or yield data, extending the index to cover additional sensor modalities such as soil moisture or solar radiation, and deploying the full pipeline for real-time, on-field monitoring.

Acknowledgement

The authors gratefully acknowledge the financial support provided by the Chief Minister's Research Grant (CMRG), Department of Higher Education, Government of Tamil Nadu, under Proceedings. No. CMRG / 3005 / H3 / 2024 / 046, Dated 16.07.2025; Project ID: CMRG2400486, Titled "Climate-Resilient Agriculture: AI-Powered Soil Health Monitoring with Advanced Multimodal Sensors." The authors also thank Anna University and University College of Engineering BIT Campus, Tiruchirappalli, for providing the necessary facilities and institutional support to carry out this research.

References

1) Kaggle, "Crop Recommendation Dataset," Accessed: Jul. 30, 2026. [Online]. Available: https://www.kaggle.com/datasets/atharvaingle/crop-recommendation-dataset

2) L. Breiman, "Random Forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.

3) S. M. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," in Proc. 31st Conf. Neural Information Processing Systems (NeurIPS), Long Beach, CA, USA, 2017, pp. 4765–4774.

4) F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.

5) J. R. Quinlan, "Induction of Decision Trees," Machine Learning, vol. 1, no. 1, pp. 81–106, 1986, doi: 10.1007/BF00116251.

6) O. Elijah, T. A. Rahman, I. Orikumhi, C. Y. Leow, and M. N. Hindi, "An Overview of Internet of Things (IoT) and Data Analytics in Agriculture: Benefits and Challenges," Journal of King Saud University – Computer and Information Sciences, vol. 34, no. 6, pp. 281–302, 2022, doi: 10.1016/j.jksuci.2020.03.002.

7) M. Javaid, A. Haleem, I. H. Khan, and R. Suman, "Understanding the Potential Applications of Artificial Intelligence in Agriculture Sector," Advanced Agrochem, vol. 2, no. 1, pp. 15–30, 2023, doi: 10.1016/j.aac.2023.04.001.

8) Food and Agriculture Organization of the United Nations (FAO), Climate-Smart Agriculture Sourcebook, 3rd ed. Rome, Italy: FAO, 2023.

9) M. N. Mowla, N. Mowla, A. F. M. S. Shah, K. M. Rabie, and T. Shongwe, "Internet of Things and Wireless Sensor Networks for Smart Agriculture Applications: A Survey," IEEE Access, vol. 11, pp. 145813–145852, 2023, doi: 10.1109/ACCESS.2023.3344578.

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 59 Issue: 07:2026
DOI: 10.5281/zenodo.21718225

10) S. P. Singh, G. Dhiman, S. Juneja, W. Viriyasitavat, G. Singal, N. Kumar, and P. Johri, “A New QoS Optimization in IoT-Smart Agriculture Using Rapid Adaption Based Nature-Inspired Approach,” IEEE Internet of Things Journal, early access, 2023, doi: 10.1109/JIOT.2023.3306353.

11) F. K. Shaikh, S. Karim, S. Zeadally, and J. Nebhen, “Recent Trends in Internet-of-Things-Enabled Sensor Technologies for Smart Agriculture,” IEEE Internet of Things Journal, vol. 9, no. 23, pp. 23583–23598, Dec. 2022, doi: 10.1109/JIOT.2022.3197850.

12) A. Morchid, R. El Alami, A. A. Raezah, and Y. Sabbar, “Applications of Internet of Things (IoT) and Sensors Technology to Increase Food Security and Agricultural Sustainability: Benefits and Challenges,” Ain Shams Engineering Journal, vol. 15, no. 3, Art. No. 102633, Mar. 2024, doi: 10.1016/j.asej.2023.102633.

13) S. Rudrakar and P. Rughani, “IoT-Based Agriculture (Ag-IoT): A Detailed Study on Architecture, Security and Forensics,” Information Processing in Agriculture, vol. 11, no. 4, pp. 524–541, Dec. 2024, doi: 10.1016/j.inpa.2023.09.002.