



**Journal on Innovations in Teaching and Learning**

**Vol: 4(3), September 2025**

**REST Publisher; ISSN: 2583 6188**

**Website: <http://restpublisher.com/journals/jilt/>**

**DOI: <https://doi.org/10.46632/jitl/4/3/2>**



## **Smart Agri Manager (S.A.M.): A Software-Driven Approach for Crop Recommendation, Irrigation Advisory, and Farm Management**

**S. Prabhakaran, M. Renuka devi, A. Murugavel, Dharshan Annaurai, Doguparthi Aakash, Gokula krishna R**

*V.S.B. Engineering College, Karur, Tamilnadan, India.*

\*Corresponding author: [mokipraba@gmail.com](mailto:mokipraba@gmail.com)

**Abstract:** *The present paper introduces Smart Agri Manager (S.A.M.), a software-based infrastructure that can assist sustainable and data-driven agricultural activities via providing crop recommendations, irrigation timing, farm advisory, and protracted digital agronomy services. In contrast with numerous digital agriculture frameworks which are often overly dependent on expensive hardware of the Internet of Things, S.A.M. focuses on a software-centered model that uses publicly available soil, crop, weather, and market data to deliver practical information. Other modules that are incorporated in its long-term vision include AI-based disease detection, yield prediction, a farmer-to-consumer marketplace, real-time multilingual alerts, knowledge support through chatbots. The framework integrates traditional machine-based learning pipelines, contemporary inference processes and rule-based engines, which are flexible to different regions and scales of farms and reduce the need of infrastructure. This makes S.A.M. not only a technical innovation but also a complete digital agriculture ecosystem. We scan the international practices in digital agriculture to point to the opportunities of emerging economies to gain through affordable and scalable solutions. S.A.M. can resolve budget limitations, the accessibility barrier, and deployment ease due to decoupling of physical sensors and hardware, which makes it specifically appropriate to mini-projects of students, academic research, and low-resource field pilots. The designed framework will be designed in a way that allows modular extensions in the future to allow it to be integrated with IoT devices or mobile applications where possible, with the software-only baseline remaining fully operational. We capture the data pre-processing plans, model structures, performance, and implementation decisions, showing how the system is able to provide input-efficient and region-aware suggestions. The paper highlights the need to implement cost-effective digital agriculture solutions, and how S.A.M. can be used as a viable entry point to introduce educational programs, farmer advisory services, and experimental field confirmations [1]– [3].*

**Keywords:** *Farming, Crop recommender, Scheduling irrigation, Weather analytics, Artificial intelligence, Explainable AI, Software-only deployment, Low-cost advisory, Sustainable farming, Decision support systems.*

### **1. INTRODUCTION**

Food security and rural livelihoods around the world are dependent on agriculture as it is the primary production activity in most economies worldwide, particularly the developing ones, where smallholder farmers are the main producers. However, such farmers do not always have timely, localized, and affordable information on making some of the most important decisions like what to plant, the nutrients to apply and when to irrigate. Venerable precision farming tools are powerful but generally hardware-demanding, with large sensor setups, drones, and the IoT infrastructure investments that are beyond the financial means of resource-starved farmers. This digital divide has increased productivity, efficiency and resilience differences between industrial farms of large scale and smallholder operations. To fill this gap, we present Smart Agri Manager (S.A.M.), which is a software-based framework that uses the least amount of physical equipment, but utilizes openly available soil, crop, and meteorological data to the

maximum. S.A.M. can provide farmers and researchers with localized advice at a low cost using statistical learning, machine learning pipelines, and lightweight rule-based inference [1], [2]. The system focuses much on ensuring usability even in low-connectivity and low-budget situations, unlike hardware-intensive methods of precision agriculture which focus heavily on historic climate normals, soil properties and crop demands. Meanwhile, the framework is upgrade-compatible and is built in such a way that it can be easily integrated with IoT sensors, satellite imagery, or mobile applications in case these resources are accessible [3], [4]. S.A.M. is a modular architecture that resembles the current data science pipelines: data ingestion, cleaning, feature engineering, model training, inference services, and user interfaces. The design enables the teams to implement components in small steps and use the framework to apply to various magnitudes of deployment, such as classroom demonstrations to field pilots. Also, the modularity means that additional extensions and integrations may be made without significant redesigns in the future [5].

This paper has fourfold contributions:

- a software-based reference architecture of agricultural decision support,
- a reproducible crop prescription and irrigation pipeline based on publicly available datasets,
- a literature review of datasets, policies, and international practices concerning digital agriculture, and
- relative knowledge on software-only and IoT-enhanced solutions [6], [7].

We lastly put this work in the perspective of a student-friendly mini-project, which is educative and practical at the same time. When we capture the patterns of coding, assessment measures and mode of deployment, we facilitate effortless replication of coursework, hackathons, and field tests dividing a learning theory and applied agricultural innovation [8].

## **2. DATASET DESCRIPTION**

S.A.M. is mostly based on publicly available soil, crop, and meteorological data as the basis of creating recommendations. In contrast to more sensor-intensive precision farming systems, the framework is inherently less reliant on hardware, drawing on both global and national archives including FAOSTAT, the United States Department of Agriculture (USDA), the Indian Council of Agricultural Research (ICAR) and Copernicus Earth observation services. Such repositories can provide good coverage in areas where field-level sensors are not available to ensure the advisory system is scalable, low cost, and accessible [1], [2], [10], [19]. In order to ensure interoperability, S.A.M. performs schema harmonization among heterogeneous sources of data with the help of geospatial keys (e.g., latitude-longitude, district boundaries) and standard units (e.g., rainfall in mm, soil nutrient in ppm). It allows regular downstream processing and integration of datasets of varying structure and resolution. Missing or non-consistent records are handled by plausibility checks (outlier detection, range validation) and statistical imputation to enhance the robustness in the case of a sparse coverage [12]. Since there is a trade-off between granularity and reliability, the data is temporarily aggregated at a weekly pace, and spatially aggregated at the district level. This design strikes a balance between necessary noise reduction and computational efficiency on the one hand, and the necessity of keeping the local relevance to advisory purposes on the other hand [13]. As an example, weekly aggregation fits well into irregular agricultural decision-making (e.g. sowing windows, irrigation schedules), whereas district-level aggregation is compatible with a majority of publicly available agricultural data. A key design concept of S.A.M. is the explicit maintenance of metadata of uncertainty. Rather than causing discarding of variance or confidence intervals, the framework propagates uncertainty estimates across the pipeline avoiding overconfident predictions. It is especially important where the risk-sensitive advisory is of interest, e.g. irrigation timing and sowing, when conservative defaulting is desirable, rather than an aggressive approach, which might damage the yield. Lastly, the data layer is designed taking into account privacy and compliance. Since S.A.M. does not collect any personally identifiable information (PII) as the primary source of data is open data. In the cases of incorporation of farmer records or survey data to check them, the data is anonymized and handled based on the government guidelines and data protection norms [14]. This makes sure that the system is technically valid, but it is also ethically compliant with the new norms in the field of digital agriculture regulation

### **3. SOIL AND CROP DATA (SOFTWARE-FIRST VIEW)**

The current form of S.A.M. is set as a software-only pipeline, which can be made to work without the presence of the IoT devices, remote sensors, or field instruments. Crop and soil characteristics are then inferred based on the community repositories and reference tables, and they do not rely on physical data retrieval.

#### **Soil Features**

Surveys surveys and government records that are publicly available provide the soil descriptors of texture, pH, EC, organic carbon, cation exchange capacity (CEC) and macro-nutrients. These soil properties are then calculated to compute derived proxies of water holding capacity in place of in-field moisture sensors. The method illustrates that heuristics and statistical models may be used in lieu of actual sensor measurements during initial implementations [6], [10].

#### **Crop Requirement Profiles**

Crop behavior is modeled by use of requirement tables (sowing windows, maturity length, base temperature, heat units, stage-wise Kc, and nutrient demand) that were taken out of agronomic literature and extension manuals [16], [18]. These tables are software-encoded agronomy regulations, serving the purpose of the IoT systems to do in real-time with observations.

#### **Features of Interaction and Rotation**

Interactions between the features (e.g. pH x tolerance, texture x root depth) are intended to simulate domain knowledge in a data-driven pipeline. Rotation features represent agronomic continuity obtained through encoding the influence of previous crops on soil fertility, pest dynamics, and nitrogen requirements [17]. They would be supplemented in hardware-augmented deployments with field sensors and pest surveillance, although currently they are rule- and dataset-driven.

#### **Target Variables**

Goals are determined by suitability indices, yield levels or profit proxies, based on the coverage of data and priorities of farmers. When used in a software-only environment, they are based on historical yield tables and prices statistics. These can be replaced by field information in real-time in future versions as they will be provided after the integration of IoT modules.

### **4. WEATHER DATA**

The weather is a decisive factor of crop development, water equilibrium and the stability of yield. Where there is no high-density network of stations, or sensor data available at the farm level, S.A.M. uses publicly available gridded data, historical climate normals, and forecast products to provide spatially sensitive advice.

#### **Input Variables**

The main climatic inputs are the hottest and hoollest temperature (Tmax/Tmin), precipitation, relative humidity, wind velocity, and sunshine or cloud cover index. In regions where full sets of variables to estimate evapotranspiration (ET 0 ) are not available, an approximation form of the Penman Monteith equation [18] called Hargreaves equation is applied. This guarantees a reliable way of scheduling irrigation even in low-data situations.

#### **Temporal Anchors**

Historical climate normals (19912020) are long-term averages of temperature and rainfall that the risk assessment of sowing dates and irrigation demands. In the case of operational advisories, short-term predictions (710 days) are consumed in order to optimize near-term sowing, fertilizer, and irrigation timed choices [11], [19]. This two-tier strategy balances stability in the long-run and responsiveness in the short-run.

## Derived Features

Along with the raw climatic variables, derived indicators are also calculated:

- Dry-spell sporadicity and rainfall onset logic to help denote the sowing windows, crop establishment, and irrigation breaks.
- Crop phenology: Heat unit accumulation (growing degree days)
- Crop phenology: Accumulation of heat units (growing degree days).
- Indicators of extreme events (e.g., >40C days, dry days) to indicate stresses on crops.

## Gridded Products and cover.

In regions with low station coverage, S.A.M. uses gridded reanalysis data (e.g., Copernicus ERA5, IMD gridded rainfall). Over-smoothing or systematic underestimation are artifacts that are very much reduced by bias correction, smoothing, and spatial interpolation [20]. This will guarantee that localised advice is still believable even when there is information scarcity.

## Model Stability Standardization.

All weather features are refined to the location in order to make machine learning models generalizable to heterogeneous geographies. This normalization provides stability in training, and also increased cross-regional transferability, so that any models trained in one district can be transferred to another with only slight retraining.

## 5. SYSTEM DESIGN

S.A.M. is a modular and layered architecture focusing on loose coupling, reusability and the ability to upgrade. The design divides issues into data, feature, model, service and presentation layers, so that the modules can develop independently and that the overall pipeline is not broken.

### Layered Architecture

- Data Layer: Processes the ingestion and pre-processing of soil, crop, and weather data, combining a variety of data with harmonized schema.
- Feature Layer: Codes agronomic features, interactions and transformations needed to train and infer the model.
- Model Layer: Trains machine learning models to recommend crops and schedule irrigation, as well as to calibrate measurements of uncertainty.
- Service Layer: Provides APIs that are callable models of the callable service-CropRec (crop recommendation) and Irrigate (irrigation advisory). Predicted outputs are accompanied by the explanations and uncertainty flags, which increases the quality of trust and interpretability of each service.
- Presentation Layer: This offers a low-weight, user-friendly interface to which advisory results can be made accessible to extension workers, students, or farmers (e.g. web/mobile dashboards).

### Traceability and Configurability.

The behavior of the system is defined by configuration files with an indication of regions, crops, thresholds, and advisory rules. Versioned datasets, trained artifacts, and intermediate snapshots can be reproducible and traceable to allow the teams to recreate the experiment and verify the findings, to roll back to the previous baseline.

S.A.M. modular architecture has been expanded to include the innovative services which are farmer-centric:

- Crop Recommendation and Irrigation Advisory: Core services are offered using soil, climate and agronomy data.
- Predicting Diseases in Trees: Relies on computer vision (TensorFlow/Scikit-learn) to detect plant diseases on the basis of leaf images, which may be used to intervene early on.

- Yield Prediction: Machine learning algorithms that integrate past yield tables, weather, and soil properties in the process of predicting output and resource utilization.
- Farmer-to-Consumer (F2C) Marketplace: A digital platform under S.A.M. that will allow farmers to bypass middlemen and increase the profit margin by selling products directly to consumers.
- Knowledge Hub + Chatbot: An artificial intelligence-powered natural language interface that gives immediate replies to agronomic practices, government schemes, and best practices.
- Real-time Weather & Pest Alerts: Connection to both weather APIs (e.g., OpenWeatherMap) and pest surveillance information to deliver a multilingual notification using mobile and SMS.
- Offline Activity Tracker: It is used to ensure that the farmers are able to record and get the advisories even when the internet is not available, but when connectivity is regained, the tracker synchronizes.

### **Explainability and Transparency**

One of the design requirements of S.A.M. is explainability. The advisory workflow is integrated with such techniques as permutation importance, SHAP values, or feature attribution methods [21]. Every recommendation is supported by rationale and boundaries of uncertainty so that the extension workers can know the source of system outputs. This builds trust, accountability, and adoption of both the practitioner and farmers.

### **Policies and Policy Alignment Engine**

A rules engine imposes hard agronomic and policy regulations to avoid impractical or unsafe advice. As an illustration, irrigation advice observes water-usage policies and crop recommendations are in accordance with government subsidy or procurement programs. This makes sure that the outputs are not only technically good but also policy-compliant and context-aware.

## **6. ARTIFICIAL INTELLIGENCE MODELS**

To create actionable recommendations, S.A.M. predictive core is an integration of classical machine learning pipelines and domain heuristics. The models are selectively selected in order to ensure accuracy, interpretability and robustness to make outputs reliable and understandable by the extension workers and end users.

### **Crop Recommendation Models**

The formulation of crop recommendation can either be in the form of a ranking task or a multi-class classification challenge based on the data that is available. Input characteristics are soil property, climatic factors and crop need curves, engineered interdependence (e.g. pH × tolerance, rainfall × sowing window) increasing the agronomic significance.

- Ranking method: The method generates a list of the crops that are suitable in a certain area in an ordered fashion and this method generates a suitability score.
- Classification method: Predicts the most viable crop class in a direct manner based on inputs of soil and climate.
- We prefer Random Forests (RF) and monotonic Gradient Boosted Machines (GBMs) more since they can capture non-linear interactions and they are explainable. The analysis of the importance of features (permutation importance, SHAP values) is included to justify recommendations.

### **Irrigation Advisory Models**

The crop water needs are used to schedule irrigation and these are estimated using the crop evapotranspiration equation ( $ET_c = ET_o \times K_c$ ) with  $ET_o$  being reference evapotranspiration (estimated using Hargreaves or PenmanMonteith where necessary) and  $K_c$  being a stage-dependent crop coefficient [18].

- Stage-aware modeling: The application of the different coefficients depends on the sowing, vegetative, flowering and maturity stages.
- Uncertainty bands: ETc estimates propagate the uncertainty intervals and the advisories are accompanied by a high and a low limit whereby the water management is advised to be conservative.

### **Calibration, Safety and Robustness**

To prevent biased probability scores, the calibration of outputs is performed with the help of the isotonic regression that aligns the outcomes to the predicted ones. Failover strategies maintain resilience in the presence of absent or noisy inputs e.g. in the event of untrustworthy rainfall data, fall back to conservative irrigation volumes. Such a design is skewed to risk-averse approaches, which minimize the probability of crop stress or waste of resources.

### **Implementation of Model Artifacts**

In order to ensure the consistency between training and deployment, model artifacts are serialized together with their preprocessors. This avoids training serving skew, meaning that feature scaling, encoding and imputation used in training are the same as those used in inference. The integration of the models as APIs into the service layer allows them to be integrated into advisory workflows with minimum overhead

## **7. IMPLEMENTATION**

It is designed to have a lightweight and modular software stack that means that the system can be run on relatively modest hardware and at the same time allows further future extensions to other IoT or sensor additions.

### **Technology Stack:**

- Core: Python
- Data processing: pandas Data storage: SQLite/Parquet Data and artifact storage Data processing: pandas
- Modeling: scikit-learn (training and also evaluation), ONNX/joblib (model export and model portability).
- API: FastAPI inference endpoints, both desktop and web consumers.

### **Automation Scripts:**

#### **Custom Python scripts process:**

1. Data downloading (or consumption of agricultural datasets)
2. Preprocessing (missing values, normalization, categorical encoding)
3. Model validation and training.
4. Training models to ONNX/joblib and packaging.

#### **Testing and Validation:**

Unit tests are incorporated to ensure important processing steps (e.g., normalization, feature extraction) and model results (e.g., prediction ranges).

#### **Inference Layer:**

- An interface with a command line lets offline inferencing be done on low-resource machines, so the solution can run in a rural or offline setting.
- Both the desktop and web UIs share a common inference module to ensure similar results.

### **Telemetry and Monitoring:**

- The system captures input range, prediction confidence, and concept drift measures.
- Training or dataset refresh can occur upon automated alarms that indicate drift in specific thresholds.

### **Core Stack:**

Python (FastAPI/Django REST) to scale up APIs, SQLite/PostgreSQL to maintain structured information, and Redis to cache data.

### **AI/ML Models:**

Image-based classification (disease detection), crop recommendation, yield prediction regression ensembles, TensorFlow and Scikit-learn.

### **Frontend:**

Flutter/React Native Mobile-first applications with multiple languages.

### **Marketplace Integration:**

REST APIs on product listing, price and secure payment channels.

### **Notification System:**

Push notifications and SMS (via twilio) of real-time alerts through Firebase

## **8. RESULTS AND DISCUSSION**

Simulations were retrospectively applied to representative districts of choice to assess the performance of the system. The results show that there is high accuracy between @k in ranking crop recommendations, and the consistency of the results is high across geographical zones. Water management modules proved to have quantifiable possibilities in the optimization of irrigation intervals, which indicated a possibility of real water consumption savings. A sensitivity analysis in terms of soil pH, perturbations of rainfall, and temperature showed that the recommendations are consistent in case of small environmental changes. This strength is significant in terms of operational reliability in the actual world where small deviations are the norm. Simulation of cases also noted gains in proxy yield estimations and effective irrigation scheduling. These results reinforce the thesis that the software can help in the process of sustainable agricultural practices. The usefulness of explanatory notes included with each recommendation, and uncertainty indicators were also highlighted in trial user feedback. Such capabilities made users more incentive and farmers received practical insights but noted model limitations. However, it was found to have certain limitations because of the use of rough, publicly available datasets. These kinds of datasets are not as fine-grained as they need to be to capture micro-variability in farms. Recommendations can be improved in the future with the optional integration of IoT-based sensors, which will allow making decisions at the field level more accurately. The marketplace, disease detection and knowledge hub modules add to the influence of S.A.M. beyond what is possible using a classic decision support: Farmers The direct sales turn to the advantage of farmers, who are not directed to intermediaries to increase the margin of profit.

- The early detection of the disease will minimize the loss of crops and enhance yield stability.
- Prediction of the yield also makes it possible to allocate resources and to plan ahead.
- Encourages more accessibility to farmers in areas with alternative languages due to a multilingual chatbot service.
- Offline-first design means it can be used in low connectivity areas

## 9. ISSUES AND CONSTRAINTS

- Information Quality and Depth- Agricultural data in certain areas can be rough, obsolete, or incomplete, which reduces the accuracy. This doubt is clearly conveyed to the user interface.
- Forecast Propagation Errors - There is a possibility of weather forecast errors getting into the system of making the recommendations. To counter this, instead of absolute predictions, the system will give conservative ranges.
- Limited Labeled Data - The shortage of labeled data is overcome by using appropriateness indices and rules that are operated by experts to augment machine learning models.
- Practice Heterogeneity - The agricultural practice in different regions does not allow its generalizability. Refinements of recommendations need to be made with local calibration and user feedback.
- Accessibility Constraints- There are connectivity gaps and language barrier which are overcome by offline and multilingual user interfaces.

## 10. FUTURE WORK

- Satellite Indices: Combine remote sensing data with recommendations to authenticate suggestions and early indications of crop stress.
- Active Learning: Use lightweight user feedback indicators to repeatedly improve models.
- Transfer Learning: Use domain adaptation methods to alleviate the cold-start problem in new areas.
- Market Integration: Include commodity price trends and layers of market access to make decisions economically viable.
- Accessibility Improvements: Increase voice interfaces and localization packages in the name of wider inclusivity.
- Satellite and Drone Imagery: Fine grained crop health monitoring integration.
- Blockchain-based Marketplace: Transparent transactions in the F2C platform.
- Financial Services Integration: Insurance recommendations and credit scoring on the yield predictions.

Community-based Feedback Loops: Pest outbreak, local practice and market price trends forewarnings in the form of crowdsourced data.

Human-in-the-loop High-tech NLP Models: To enable the chatbot further to become a voice-enabled conversational farm assistant.

## 11. CONCLUSIONS

The Smart Agri Manager (S.A.M.) illustrates that an entirely software-based solution is capable of providing services related to farm advisory as actionable and is accessible, scalable and cost-efficient. The system is safe and provides interpretable and context-aware advisories that are feasible to smallholder farmers and student-driven deployments by combining rule-aware logic, explainable models, and conservative defaults. The modular design enables upgrades to be made in a more incremental manner, including the addition of satellite imagery, sensor inputs or market data in the future without redesigning the architecture significantly. In addition to technical contributions, S.A.M. has an educational value, providing a reproducible framework that can be implemented in mini-projects, hackathons, and classroom demonstrations, therefore, helping fill the existing gap between theoretical and technical teaching and real-world agricultural usage. In the future, the application of user feedback loops, transfer learning, and satellite-based monitoring will help to adapt more closely to local conditions and become more accurate in predictions. Through such improvements, S.A.M. can become a scalable and community-based decision support system of sustainable farming. Moreover, the fact that S.A.M. is being expanded to other fields, including AI-driven disease detection, yield forecasting, knowledge hub, or a farmer-to-consumer marketplace also highlights its ability to serve as a full-scale digital ecosystem. This growth will make certain that besides receiving agronomic recommendations, farmers will be able to access both economic opportunities as well as risk management plans and ready-to-use knowledge resources within their respective languages. Finally, S.A.M. is not merely the decision support tool, but it is hoped that it will become a driver of the digital transformation of the agricultural sector. S.A.M. can help achieve higher levels of

resilience, better food security and sustainable rural development in the emerging economies by empowering the farmers with transparent low-cost, and data-driven solutions.

## REFERENCES

- [1]. FAO, “FAOSTAT—Food and Agriculture Data,” Food and Agriculture Organization of the United Nations.
- [2]. [World Bank, “Agriculture & Rural Development Data”.
- [3]. CGIAR, “Digital innovations for smallholder agriculture”.
- [4]. ICAR, “Agro-Climatic Zones and Cropping Systems”.
- [5]. European Commission JRC, “Agri datasets and services”.
- [6]. USDA NRCS, “Soil Survey Manual; Soil Taxonomy”.
- [7]. FAO, “Efficient plant nutrition and fertilizer use”.
- [8]. IEEE, “AI systems—data quality considerations”.
- [9]. ISRIC, “SoilGrids—Global gridded soil information”.
- [10].NOAA, “GHCN—Historical climate; forecast resources”.
- [11].Eurostat, “Agri-environmental indicators”.
- [12].IMD (India), “District Rainfall and Monsoon Reports”.
- [13].Government data protection guidelines—agri advisories.
- [14].FAO, “Crop Water Information (Kc)”.
- [15].ICAR, “Crop rotations and soil amendments”.
- [16].R. G. Allen et al., “FAO-56: Crop Evapotranspiration”.
- [17].ECMWF/Copernicus, “ERA5-Land reanalysis”.
- [18].USGS/NASA, “Landsat Missions”.
- [19].S. Lundberg and S. Lee, “A Unified Approach to Interpreting Model Predictions (SHAP)”.
- [20].Anbumani P, Dhanapal R. Enhanced Blockchain-based Key Generation using Butterfly Optimization Algorithm for Efficient Data Sharing in Cloud Computing. *Journal of Information Technology Management*. 2023 Mar 1;15(Special Issue: Digital Twin Enabled Neural Networks Architecture Management for Sustainable Computing):21-42.
- [21].MS. M. SANGEETHA DR. N. K. KARTHIKEYANDR. P. TAMIJESELVY “A Pioneering Approach for the Prognosis of Cancer from Micro Array Data Using Bi-BPCA and ID3 Algorithm” *International Journal of Research in Computer Applications & Information Technology* Volume 10, Number 20 41643- 41649 *International Journal of Applied Engineering Research*.