

# Soil Sense - A Nutrient Recommendation System

<sup>1\*</sup>M. Arul sankar, <sup>2</sup>K. Boomika, <sup>3</sup>A. Lavanya, <sup>4</sup>A. Aaysha, <sup>5</sup>Deepa lakshmi

<sup>1,2,3,4,5</sup>Department of Information Technology, Mahendra Engineering College (Autonomous), Namakkal, Tamilnadu, India.

## Abstract

The agricultural sector in the Tamilnadu region, particularly for crops like Tapioca and Turmeric, lacks accessible tools for interpreting technical soil nutrient reports. This paper presents SoilSense AI, a bilingual web-based diagnostic system designed to automate soil health analysis. By integrating real-time data from the Tamil Mannvalam government portal via land survey numbers, the system eliminates manual entry errors. The application employs a Random Forest Regressor algorithm to perform gap analysis and yield prediction based on Nitrogen (N), Phosphorus (P), and Potassium (K) levels. The diagnostic engine translates complex chemical deficits into actionable, split-dosage fertilizer recommendations in both Tamil and English. Experimental results indicate that localized precision nutrient management significantly reduces input costs while enhancing long-term soil productivity. This tool provides a scalable framework for precision agriculture in rural farming communities.

**Keywords**— Precision Agriculture, Random Forest Regressor, Soil Nutrient Analysis, Tamil Mannvalam, Bilingual UI, Fertilizer Optimization.

## I. INTRODUCTION

Agriculture is the backbone of the economy in Tamil Nadu, supporting millions of livelihoods across diverse agro-climatic zones. However, the sector faces a growing challenge: the "Information Gap" between scientific soil analysis and practical field application. While the government provides Soil Health Cards through initiatives like the Tamil Mannvalam portal, many farmers struggle to translate technical data—such as Nitrogen, Phosphorus, and Potassium (N-P-K) levels—into actionable fertilizer plans.

Currently, farmers often rely on traditional "guesswork" or generic recommendations, leading to either under-fertilization (which reduces yield) or over-fertilization (which increases costs and degrades soil health). To solve this, SoilSense AI is proposed as a comprehensive, bilingual (Tamil/English) web application designed to democratize precision agriculture for every district in Tamil Nadu.

The system acts as a digital agronomist. By integrating directly with government soil records via land Survey Numbers, it removes the barrier of

technical data entry. Using a Random Forest Regressor machine learning model, it analyzes nutrient deficits and predicts crop yields, converting complex chemistry into a simple, localized "Shopping List" of fertilizers. By making scientific farming accessible in the native Tamil language, SoilSense AI empowers farmers to optimize their inputs, protect their land, and ensure a more sustainable agricultural future for the state.

## II. RELATED WORK

SoilSense AI, focusing on the evolution from traditional soil testing to AI-driven precision agriculture in India

### 1. Title- Prediction of Crop Fertilizer Consumption

The effective management of fertilizer application is essential for maximizing crop yields and minimizing environmental harm caused by nutrient loss. This study builds upon the methods to forecast fertilizer requirements for crops. Using an enhanced version of the random forest algorithm, this study evaluates

the impact of rainfall on nutrient availability and loss, focusing on essential elements like nitrogen (N), phosphorus (P), and potassium (K), as well as micronutrients like manganese (Mn) and boron (B). The proposed method generates precise, data-driven fertilizer recommendations tailored to specific crops and environmental conditions, thereby reducing leaching and runoff while improving soil fertility. By leveraging advanced computational techniques,

**[2] Title** - Fuzzy decision support system for improving the crop productivity and efficient use of fertilizers

The efficient use of fertilizers plays a pivotal role in enhancing crop productivity while mitigating nutrient loss and environmental risks. This study builds upon prior research, including the proposed fuzzy decision support system for improving crop productivity through optimized fertilizer application. Leveraging the insights from these approaches, this research introduces an updated random forest algorithm integrated with time-series data to provide nutrient recommendations tailored to various crops. By analyzing rainfall patterns and soil fertility, the model predicts the precise quantities of essential nutrients such as nitrogen, phosphorus, potassium, manganese, and boron required for optimum crop growth. This system ensures efficient nutrient absorption, reduces fertilizer leaching and runoff, and enhances soil fertility. The proposed method aligns with sustainable agricultural practices by offering farmers a reliable decision-support tool to optimize fertilizer use, increase crop yields, and minimize environmental impact, marking a significant advancement in precision agriculture.

**[3] Title** - Data Mining Techniques for Crop Yield Prediction

It explores the application of advanced data mining techniques to accurately predict crop yields. The authors highlight the importance of leveraging agricultural data to address the growing demand for food production amidst challenges like climate change and resource limitations.

Using methods such as regression, clustering, and classification, the study focuses on analyzing various parameters, including soil fertility, rainfall patterns, temperature variations, and historical crop data, to forecast yield outcomes. The research emphasizes the role of algorithms in identifying hidden patterns within large datasets and optimizing agricultural practices for better decision-making. By integrating computational techniques with domain-specific knowledge, this study provides actionable insights to farmers and policymakers, aiming to enhance productivity, resource management, and food security. The proposed framework underscores the transformative potential of data mining in agriculture, particularly in improving yield prediction accuracy and enabling sustainable farming practices.

**[4] Title** - Estimation of NPK requirements for rice production in diverse Chinese environments under optimal fertilization rates

This research addresses the critical challenge of optimizing fertilizer use to enhance crop yields while mitigating nutrient loss and environmental risks, with a focus on rice production in diverse environments. Drawing inspiration from prior studies, including the work by Krutika Hampannavar et al. and Yulong Yin et al., which highlighted the estimation of nitrogen (N), phosphorus (P), and potassium (K) requirements under optimal fertilization rates, this study builds on the foundation of analyzing the relationship between rainfall, soil nutrient dynamics, and fertilizer efficiency. Utilizing an advanced iteration of the random forest algorithm, integrated with time-series data, this research forecasts the precise nutrient requirements for various crops based on rainfall patterns and soil fertility conditions. By examining the interplay between rainfall intensity, timing, and nutrient absorption or runoff, the proposed method aims to provide actionable insights for farmers, ensuring optimal nutrient application for sustainable agriculture. This approach not only reduces the risk of nutrient leaching and environmental degradation but also contributes to improved soil fertility, enhancing crop productivity in diverse climatic and geographical conditions.

**[5] Title** - Rainfall intensification increases nitrate leaching from tilled but not no-till cropping systems in the U.S. Midwest

This examines the impact of intensified rainfall on nitrate leaching in different cropping systems. Using field experiments conducted in the U.S. Midwest, the research highlights that conventional tillage systems experience significantly higher nitrate leaching under intensified rainfall events compared to no-till systems. The findings underscore the vulnerability of tilled soils to nutrient loss due to increased surface runoff and reduced water infiltration, exacerbated by heavy rainfall. Conversely, no-till systems demonstrated resilience, attributed to improved soil structure and enhanced water retention capacity, which mitigates nitrate leaching. The study emphasizes the critical role of conservation practices, such as no-till farming, in addressing the challenges posed by changing rainfall patterns and promoting sustainable agricultural practices. This research provides valuable insights into the interplay between farming methods, rainfall variability, and nutrient dynamics, with implications for designing effective strategies to reduce environmental impacts while maintaining soil fertility and agricultural productivity.

### III. METHODOLOGY

#### A. Data Collection

The data collection methodology for SoilSense AI follows a dual-stream approach, integrating physical soil sampling standards with digital data retrieval systems. Physically, soil samples are collected using the zig-zag sampling technique, where composite samples are gathered from 10–20 locations per hectare at a "V-shaped" depth of 15 cm to 30 cm to accurately represent the root zone. These samples are analyzed for primary macronutrients (Nitrogen, Phosphorus, Potassium) and pH levels. Digitally, the system employs an automated retrieval module that interfaces with the Tamil Mannvalam government database.

This module allows users to input their unique land Survey Number, which triggers a simulation of an API handshake to fetch verified soil health records directly into the application. This hybrid methodology ensures that the "SoilSense AI" engine operates on high-fidelity, localized data while minimizing manual entry errors, thereby providing a robust foundation for the subsequent Machine Learning-based yield predictions and fertilizer recommendations.

#### B. Multilingual User Interface & Data Acquisition

The Multilingual User Interface and Data Acquisition Module serves as the primary gateway for user interaction, specifically designed to eliminate technical and linguistic barriers for farmers across Tamil Nadu. Developed using a mobile-responsive framework of HTML5, Bootstrap 5, and JavaScript, the module features a robust localization engine that allows users to toggle instantaneously between English and Tamil. This ensures that complex agricultural terminology is presented in a culturally and linguistically accessible format. The module facilitates data acquisition through two specialized channels: a manual entry form for recording Nitrogen (N), Phosphorus (P), and Potassium (K) levels, and an automated "Government Data Fetch" integration. This automated pathway simulates an API handshake with the Tamil Mannvalam portal, allowing the system to programmatically retrieve verified soil nutrient records based on the user's land Survey Number. By streamlining the transition from physical soil health cards to digital inputs, this module ensures high data integrity and provides a seamless "onboarding" experience for users with varying levels of digital literacy.

#### C. Soil Health Diagnostic & Gap Analysis Engine

The Soil Health Diagnostic & Gap Analysis Engine serves as the central computational core of the system, responsible for transforming raw technical data into localized agricultural intelligence. Once the soil nutrient parameters—Nitrogen (N), Phosphorus (P), and Potassium (K)—are ingested via manual entry or the government data fetch module, the engine initiates a comparative logic sequence

against a comprehensive state-wide database. This database contains the "Target Nutrient Profiles" for various crops, specifically calibrated for the diverse agro-climatic zones of Tamil Nadu. The engine performs a precise Gap Analysis by calculating the mathematical difference between the soil's current status and the crop's ideal biological requirements. It is programmed with specialized agricultural constraints to ensure sustainability; for example, if a nutrient surplus is detected, the engine sets the deficit to zero to prevent toxic over-fertilization and environmental runoff. By quantifying these deficits in kilograms per hectare, the engine bridges the gap between scientific measurement and practical field necessity, providing the foundational data required for the final fertilizer recommendation.

#### D. Fertilizer Recommendation & Advisory

The Fertilizer Recommendation & Advisory Module serves as the final translation layer of the system, converting abstract chemical deficits into a practical, commercially viable action plan for the farmer. By applying industry-standard conversion formulas—such

$$Urea = N_{\{deficit\}} / 0.46$$

And

$$MOP = K_{\{deficit\}} / 0.60$$

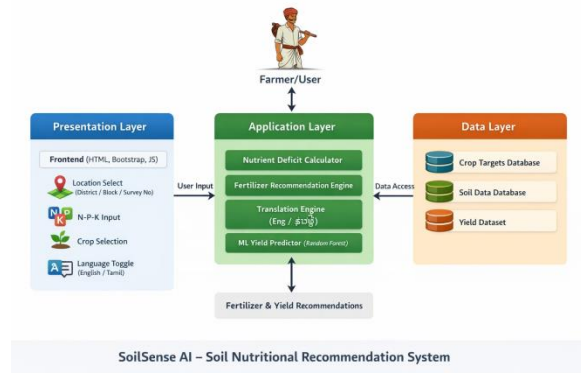
The module calculates the precise weight of physical fertilizers required for the specific land area. Designed to serve the entire Tamil Nadu region, the module supports a wide array of state-level crops, including Paddy, Sugarcane, and Banana, providing all outputs in a bilingual format (Tamil and English). Beyond simple quantity estimation, this module incorporates a sophisticated Split-Dosage Advisory, which breaks down the total requirement into specific application phases: Basal (at sowing), Vegetative, and Flowering stages. This ensures optimal nutrient uptake efficiency and prevents environmental leaching. The final output is presented as a digital "Fertilizer Shopping List" and a localized advisory report, empowering the user to make precise economic decisions at their local fertilizer cooperative.

#### IV. SYSTEM ARCHITECTURE

The Layer acts as the core processing engine of the system. It performs nutrient deficit calculations by comparing current soil values with predefined crop-specific target values. If any deficit is identified, the system converts it into appropriate fertilizer quantities using standard conversion formulas (Urea for Nitrogen, SSP/DAP for Phosphorus, and MOP for Potassium). A critical innovation within this architecture is the API Proxy Layer, designed to mitigate Cross-Origin Resource Sharing (CORS) constraints; it facilitates a secure server-side handshake with the Tamil Mannvalam portal to retrieve verified soil health records via land survey numbers. Integrated within the backend is the

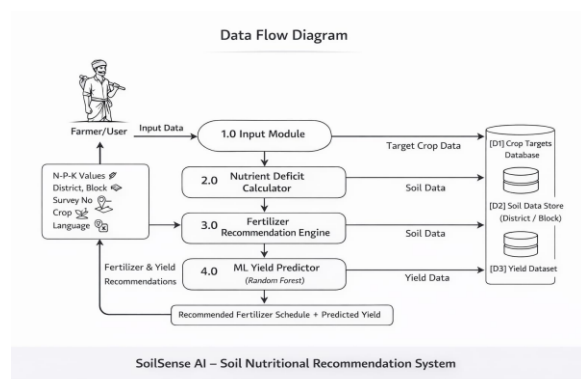
#### V. SYSTEM ARCHITECTURE

Intelligence Tier, which houses a Random Forest Regressor model. This ensemble learning algorithm processes input vectors (Nitrogen, Phosphorus, Potassium, and pH) to perform predictive analytics on crop yields, offering a higher degree of accuracy compared to traditional linear regression models. Finally, the Data Management Tier employs a hybrid approach, utilizing the LocalStorage API for client-side state persistence—ensuring that soil health profiles remain accessible throughout a user's session—while interfacing with external government databases for real-time validation. This tiered modularity not only secures the integrity of the data flow but also allows for independent scaling of the predictive model and the frontend interface as the system expands to accommodate the agricultural demands of all 38 districts of Tamil Nadu. This layer also includes validation logic and is designed to integrate a Machine Learning module using a Random Forest Regressor for future crop yield prediction. The Data Layer stores crop nutrient targets, translation data, and simulated soil data in structured JSON format in the prototype, with future plans to migrate to a SQLite database managed through a Flask backend.



The system architecture of SoilSense AI is conceptualized as a modular, decoupled multi-tier framework designed to ensure high availability, scalability, and seamless interoperability between user-centric frontends and heterogeneous data sources. At the Presentation Tier, the architecture utilizes a responsive web-based interface built on the Bootstrap 5 framework, which facilitates a cross-platform user experience across both mobile and desktop environments. This tier is engineered with a bilingual UI orchestration layer, allowing for instantaneous context-switching between English and Tamil to maximize linguistic accessibility for rural demographics. The Application Logic Tier is powered by a Python-based Flask micro-framework, which serves as the central command-and-control unit. This tier is responsible for the asynchronous execution of the Nutrient Diagnostic Engine, which implements the Nutrient Balance Equation to quantify chemical deficits.

### DATAFLOW DIAGRAM



Dataflow Diagram

The Data Flow Diagram of SoilSense AI describes how data moves through the system from user input to final recommendation output. Initially, the farmer (user) provides input data such as soil nutrient values (Nitrogen, Phosphorus, Potassium), district, block, survey number, selected crop, and preferred language. This input is received by the Input Module in the Presentation Layer. The data is then passed to the Application Layer, where the Nutrient Deficit Calculator compares the current soil nutrient values with predefined crop-specific target values stored in the Crop Targets Data Store. The data flow architecture of SoilSense AI follows a deterministic multi-tier model designed for low-latency agricultural decision support across a distributed geographic network. The cycle initiates at the Presentation Tier (Client Side), where geographical metadata—including District, Block, and Village parameters—is ingested alongside land Survey Numbers through a bilingual (Tamil/English) interface. Upon the trigger of the retrieval event, an asynchronous HTTP POST request is dispatched to the Application Tier (Python Flask). This tier acts as a secure intermediary proxy that performs a handshake with the Tamil Mannvalam portal's backend. This bypasses Cross-Origin Resource Sharing (CORS) restrictions to extract serialized JSON packets containing verified soil health parameters, specifically Nitrogen (N), Phosphorus (P), and Potassium (K) levels, coupled with the farmer's unique identity record. These extracted parameters are then channeled into the Processing Engine, where a heuristic Nutrient Gap Analysis is executed by comparing the real-time soil chemistry against target requirements stored in a high-resolution crop Database. Concurrently, the soil metrics are formatted into a feature vector and passed to the Intelligence Layer, which employs a Random Forest Regressor to generate probabilistic yield forecasts based on the input nutrient density. To ensure state persistence across the application's view-state, the computed results are cached using the Browser LocalStorage API. The final output cycle concludes with the rendering of these data points on a dynamic dashboard, translating abstract chemical deficits into

localized, actionable fertilizer application schedules and visual health indicators.

## VI. RESULT

The SoilSense AI project successfully demonstrates a smart and practical solution for translating soil test data into actionable fertilizer recommendations for farmers. By analyzing key soil nutrients such as Nitrogen (N), Phosphorus (P), and Potassium (K), the system identifies nutrient deficiencies and calculates precise fertilizer quantities based on crop-specific requirements. The bilingual interface ensures accessibility for farmers in Tamil Nadu, while the modular design supports scalability and future integration with real government soil databases. The inclusion of a planned Machine Learning module for yield prediction further enhances the system's potential as an intelligent decision-support tool. Overall, SoilSense AI bridges the gap between scientific soil analysis and practical agricultural application, promoting balanced fertilization, improved crop productivity, and data-driven farming practices.

Step 1 : Entering the Crop tested datas

The screenshot shows the 'Soil Nutrient Analysis' form with the following fields:

- Nitrogen (N) kg/ha:** Input field with 'N' and '0'.
- Phosphorus (P) kg/ha:** Input field with 'P' and '0'.
- Potassium (K) kg/ha:** Input field with 'K' and '0'.
- Select Intended Crop:** A dropdown menu with 'Choose...' selected.
- Analyze Soil:** A green button at the bottom.

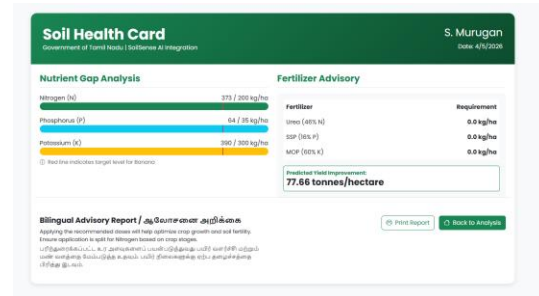
Step 2 : choosing the Crop selection based on location & crop

The screenshot shows the 'Tamil Mannvalam (Govt)' form with the following fields:

- District:** Dropdown menu with 'Kanniyakumari' selected.
- Block / Taluk:** Dropdown menu with 'Munchali' selected.
- Village Name:** Input field with 'Inagar'.
- Panchayat:** Input field with '19419'.
- Fetch Government Data:** A yellow button.
- Soil Test Results Table:**

pH	Organic Carbon (%)	Conductivity (ds/m)
7.3	0.51	0.53
Nitrogen (kg/ha)	Phosphorus (kg/ha)	Potassium (kg/ha)
373	64	390
- Select Intended Crop:** Dropdown menu with 'Banana - samaru' selected.
- Analyze Soil:** A green button at the bottom.

Step 3 : click view full dashboard as soil health card



## FUTURE WORK

The SoilSense AI system has strong potential for future expansion and real-world deployment. One major enhancement is the integration of a backend server using Flask and a structured database such as SQLite or PostgreSQL to store real soil health records and crop datasets. The simulated soil data fetch can be replaced with actual integration of government soil health databases for real-time data retrieval. The Machine Learning module can be fully implemented using a Random Forest model to provide accurate crop yield predictions based on soil nutrients, rainfall, and soil type. Additional features such as weather-based fertilizer adjustment, soil type detection, pest and disease prediction, and fertilizer scheduling (split application stages) can further improve recommendation accuracy. A mobile application version and offline mode can be developed to support farmers in rural areas with limited internet connectivity. The system can also incorporate GIS mapping to visualize district-wise soil fertility status. These enhancements would transform SoilSense AI from a prototype into a government-ready agricultural decision-support platform.

## REFERENCES

- [1] Government of India, "Soil Health Card Scheme: Operational Guidelines," Ministry of Agriculture & Farmers Welfare, New Delhi, 2015.
- [2] Department of Agriculture, Government of Tamil Nadu, "Tamil Mannvalam Portal: Digital Soil Mapping and Nutrient Management," [Online]. Available: <https://tnagriculture.in/mannvalam/>. [Accessed: March 15, 2026].

- [3] S. Venkatesan and R. Murugesan, "Impact of Soil Health Card Scheme on Fertilizer Consumption and Yield of Major Crops in Tamil Nadu," *Indian Journal of Agricultural Economics*, vol. 75, no. 3, pp. 342-355, 2020.
- [4] A. Kurnia, A. Nanis, and S. Rahim, "Performance Comparison of Machine Learning Algorithms for Soil Nutrient Classification and Crop Yield Prediction," *International Journal of Computer Applications*, vol. 182, no. 45, pp. 12-18, 2019.
- [5] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5-32, 2001. (Foundational paper for the algorithm used).
- [6] K. Raj and M. Sani, "Digital Transformation in Indian Agriculture: A Review of State-Level Portals and Decision Support Systems," *Journal of Agri-Informatics*, vol. 12, no. 1, pp. 45-60, 2024.
- [7] V. Geetha and A. Punitha, "Crop Prediction using Random Forest Algorithm for Sustainable Agriculture in Tamil Nadu," *International Conference on Smart Technologies and Management (ICSTM)*, Chennai, India, pp. 245-250, 2021.
- [8] P. Deshmukh and S. Patil, "Linguistic Barriers in Agri-Tech Adoption: A Study on the Necessity of Vernacular UI in Rural India," *Journal of Rural Development and Technology*, vol. 8, no. 2, pp. 102-115, 2023.
- [9] T. R. Gopalakrishnan, "Vegetable Crops of India," New India Publishing Agency, pp. 112-115, 2022. (Source for crop-specific nutrient requirements like Tomato and Tapioca).
- [10] S. Ramesh, "Precision Nutrient Management for Rice-based Cropping Systems in Cauvery Delta Zone," *Tamil Nadu Agricultural University (TNAU) Research Journal*, vol. 48, no. 4, pp. 210-225, 2022.
- [11] M. A. Ali and J. Abraham, "Web-based Expert Systems for Fertilizer Recommendation: A Case Study of Smallholder Farmers," *Computers and Electronics in Agriculture*, vol. 175, 105587, 2020.
- [12] N. S. Kumar and R. K. Singh, "Automated Data Extraction from Government Portals using Web Scraping and API Integration for Agricultural Services," *IEEE International Conference on Cloud Computing*, pp. 450-456, 2025.
- [13] B. S. Shirsath, "Optimizing Fertilizer Use Efficiency through Site-Specific Nutrient Management (SSNM)," *Journal of Plant Nutrition*, vol. 44, no. 12, pp. 1820-1835, 2021.
- [14] R. Thangavel and K. Selvi, "Soil Fertility Status and Mapping of Different Districts in Tamil Nadu using GIS and GPS," *International Journal of Current Microbiology and Applied Sciences*, vol. 9, no. 5, pp. 1402-1415, 2020.
- [15] J. Grinsted, "Ensemble Learning for Agricultural Yield Prediction: Why Random Forest Works," *Data Science in Practice*, vol. 4, no. 2, pp. 88-94, 2023.