INTERNATIONAL JOURNAL FOR INNOVATIVE RESEARCH IN MULTIDISCIPLINARY FIELD
ISSN(O): 2455-0620 [Impact Factor: 9.47]
Monthly, Peer-Reviewed, Refereed, Indexed Journal with IC Value : 86.87
Volume - 12, Issue - 11, November - 2026

DOIs:10.2015/IJIRMF/202611002 --:-- Research Paper / Article / Review

AI-Based Land Suitability Analysis and Recommendation System Using Machine Learning and GIS

R.M.VAMSIKRISHNA[1], A.SUSHWANTH[1], A.GIRIVARDHAN[1], G.UDAY[1], MS.RICHATIWARI[2]
Students[1], Assistant Professor[2]. Department of (Data Science) Computer Science and Engineering
Hyderabad Institute of Technology & Management,
Gandimaisamma medchal checkpost road, Basuragadi Village, Hyderabad, Telangana 501401
Email:- ramovakr4@gmail.com

Abstract - *Sustainable land-use planning relies on land suitability assessment, because individual parcels differ in how well they serve agricultural, residential, commercial, industrial, or mixed-use purposes. Conventional assessment methods rest largely on manual analysis and expert judgement, which makes them slow and hard to extend to large spatial datasets. This paper describes an AI-Based Land Suitability Analysis and Recommendation System that couples machine learning with Geographic Information System (GIS) techniques to assess land for several purposes. The system examines spatial, environmental, terrain, soil, water, infrastructure, and risk-related factors and produces suitability scores for five land-use categories: agriculture, residential, commercial, industrial, and mixed-use. A machine-learning pipeline predicts these scores and returns a recommended land-use category together with a confidence indicator. A nearby-land analysis adds spatial context for the chosen parcel, and SHAP (SHapley Additive exPlanations) highlights the features that raise or lower each prediction. Through an interactive web interface, users can analyze a parcel by its Land ID and inspect it on a GIS map. Several machine-learning models are compared experimentally using MAE, RMSE, R^2, accuracy, precision, recall, and F1-score. By uniting predictive modeling, spatial analysis, recommendation, and explainability on one platform, the framework offers an integrated decision-support approach to land assessment.*

Keywords - *Land Suitability Analysis, Machine Learning, Geographic Information System (GIS), Spatial Analysis, Land-Use Recommendation, Explainable AI, SHAP, Spatial Prediction, Decision Support System)*

I. INTRODUCTION

Land is a basic natural resource on which farming, housing, commerce, industry, and mixed-use development all depend. How suitable a given parcel is for any of these uses is shaped by soil characteristics, water availability, climate, terrain, accessibility, infrastructure, and environmental risk. Reliable suitability assessment is therefore a prerequisite for sound land-use planning and sustainable development.

Conventionally, suitability is judged by manually weighing many geographical and environmental factors. This is slow, demands substantial expert knowledge, and becomes impractical once a large number of parcels must be examined. Growing access to geospatial datasets and machine learning techniques makes it possible to automate this process and to base land-use decisions on data.

A number of recent studies have used machine learning and geospatial methods for land suitability, mostly in agricultural or crop-specific settings. Many of these approaches, however, address only one land-use category, are tied to particular crops, or need extra infrastructure such as IoT sensors and a continuous stream of environmental data. What is missing is an integrated system that can assess suitability for several development purposes and, at the same time, supply spatial context and interpretable predictions.

This paper introduces an AI-Based Land Suitability Analysis and Recommendation System that brings together machine learning and Geographic Information System (GIS) techniques. The system assesses five land-use categories

(Agriculture, Residential, Commercial, Industrial, and Mixed-Use). It draws on a range of land, environmental, infrastructure, terrain, and spatial factors to compute a suitability score for every category, and it recommends the category whose predicted score is highest.

Besides predicting suitability, the system reports a confidence indicator and compares a parcel with geographically close ones through a nearby-land analysis. Explainable Artificial Intelligence (XAI) based on SHAP is used to show which features push an individual prediction up or down. Results are delivered through a web platform with interactive GIS visualization, so that users can look up parcels, review their suitability results, see what drove a prediction, and browse the surrounding land records.

The aim of the framework is to offer a single decision-support platform that unites machine learning-based prediction, multi-purpose land assessment, spatial analysis, recommendation, and explainability. Such an approach can make land-use assessment more systematic and data-driven, and it lays the groundwork for later incorporation of real-time geospatial and environmental information.

2. LITERATURE SURVEY

Land suitability has traditionally been evaluated from soil, terrain, climate, and other environmental factors, using conventional land evaluation and multi-criteria decision-making methods. Mustafa et al. built a framework for analyzing land suitability for different crops with remote sensing and GIS, following the FAO land evaluation approach. By drawing on soil and land characteristics to rate crop suitability and derive cropping patterns, the study showed how useful geospatial information is for systematic land evaluation.

Prakash T.N [1] proposed a fuzzy multi-criteria decision-making method for agricultural land suitability analysis, using geoinformation and fuzzy techniques to deal with the uncertainty inherent in land evaluation. The study shows that suitability assessment gains from combining several criteria instead of relying on one land characteristic.

Sinchana et al.[2] surveyed land suitability mapping based on geospatial technologies, covering how remote sensing, GIS, and machine learning are applied in such studies. Their review points to a shift away from conventional suitability mapping toward data-driven methods that can handle many spatial and environmental variables at once.

El Behairy et al. [3] put forward a GIS approach integrated with artificial intelligence to evaluate wheat suitability in arid regions. The framework paired GIS with a Mamdani fuzzy inference system fed by chemical, physical, and soil-fertility indices. It partitioned the study area into suitability classes and showed that combining AI with GIS is useful for crop suitability assessment. The application, however, was confined to wheat cultivation in an arid region.

Vincent et al.[4] developed a sensor-driven, AI-based agricultural recommendation model for assessing land suitability. IoT sensors, a Raspberry Pi, cloud infrastructure, and neural-network modeling were used to collect environmental readings and produce farming recommendations. The work illustrates what real-time sensing combined with AI can offer, but it also makes the approach dependent on sensing hardware, network connectivity, and cloud resources.

Bendahmane et al. [5] designed an approach for predicting land suitability on smart farms using IoT and deep learning. Their system joins IoT devices, MQTT, cloud computing, and deep-learning methods to make predictions from environmental data for smart-farming use. It shows the value of temporal, real-time data in land assessment, yet its architecture presupposes sensor infrastructure and uninterrupted data collection.

Pratama and Hikmawati [6] studied machine learning for predicting agricultural land suitability, concentrating on preprocessing, feature selection, and data balancing. They tested Random Forest, Decision Tree, and K-Nearest Neighbors in an effort to raise predictive performance. Their findings underline how much careful data preparation and feature selection matter when machine learning is applied to land suitability datasets.

Ganati and Sitote [7] examined machine-learning techniques for predicting land suitability for wheat and barley. They compared several machine-learning methods for crop-specific suitability prediction and showed that data-driven models

suit agricultural land assessment. The study is nevertheless limited to wheat and barley, whereas land-use planning may call for assessment across several development purposes.

Mugiyo et al. [8] carried out a scoping review of land suitability methods, covering both conventional and modern approaches, with particular attention to neglected and underutilised crops. The review brings out how varied the methodologies are and how important it is to choose criteria and evaluation methods that fit the intended land use.

Abdullah et al. [9] reported a systematic literature review on biophysical, socio-economic, and local-knowledge factors in agricultural land suitability assessment. It covers GIS-based land evaluation, multi-criteria decision-making, geospatial analysis, and hybrid methods, and stresses that a comprehensive assessment needs several sources of information to be combined.

Taken together, the reviewed work shows considerable progress in GIS-based land evaluation, machine learning, fuzzy inference, IoT-based assessment, and AI-driven agricultural recommendation. Yet most of it still concentrates on agriculture or single crops, and some methods depend on specialized sensor and cloud infrastructure. The literature also reveals a need for systems that pair multi-purpose land-use assessment with spatial context and interpretable AI. Prompted by this, the present work builds a unified framework that evaluates Agriculture, Residential, Commercial, Industrial, and Mixed-Use suitability through machine learning, GIS-based spatial analysis, nearby-land information, and SHAP-based explainability.

3. RESEARCH GAP

The related studies reviewed above point to the following research gaps:

- Limited multi-purpose land suitability analysis:-
  Most existing studies concentrate on agriculture or particular crops such as wheat and barley, and few address several land-use categories within a single system.
- Lack of integrated land-use recommendations:-
  Existing approaches usually predict agricultural suitability only, instead of giving one unified recommendation across Agriculture, Residential, Commercial, Industrial, and Mixed-Use development.
- Dependence on IoT and sensor infrastructure:-
  Some current systems need IoT sensors, Raspberry Pi devices, cloud platforms, and continuous environmental data collection, which raises infrastructure and deployment demands.
- Limited explainability of predictions:-
  Many machine-learning-based systems emphasize predictive accuracy but say little about which land characteristics drove a given prediction.
- Insufficient integration of GIS and machine learning:-
  Geospatial data feature in several studies, but there is room for closer coupling of ML-based suitability prediction with interactive GIS visualization.
- Limited spatial context from surrounding lands:-
  Existing approaches typically assess a single parcel or crop in isolation, without explicitly using neighboring land records and their recommendations as context.
- Lack of an end-to-end decision-support platform:-
  Previous studies tend to focus on the prediction model itself, rather than bringing data preprocessing, ML prediction, explainable AI, recommendation generation, nearby-land analysis, and interactive GIS visualization together on one platform.

4. PROBLEM STATEMENT

Assessing land suitability means analyzing many factors, including soil properties, climate, water availability, terrain, accessibility, infrastructure, and environmental risk. Existing methods are frequently limited to agricultural uses or specific crops, and others rely on manual evaluation, multi-criteria analysis, or additional sensing and cloud infrastructure. As a result, it is hard to obtain a scalable, unified assessment of land for different development purposes.

What is needed, therefore, is an integrated, data-driven system that can assess individual land parcels for multiple land-use categories and give an understandable recommendation grounded in several spatial and environmental factors. It should also take the geographical context of neighboring parcels into account, report confidence information, and explain what influenced each prediction.

This work sets out to build an AI-based land suitability analysis and recommendation system able to process multi-dimensional land and spatial data and predict suitability scores for Agriculture, Residential, Commercial, Industrial, and Mixed-Use categories. The system must recommend the most suitable category, analyze nearby parcels by geographic proximity, and make its predictions interpretable through SHAP-based Explainable AI. Results are to be delivered through an interactive, GIS-enabled web interface that supports systematic, data-driven land-use assessment.

5. PROPOSED METHODOLOGY

The proposed system is an integrated machine-learning and GIS framework for multi-purpose land suitability assessment. Its methodology brings together land attributes, environmental factors, infrastructure information, spatial coordinates, machine-learning prediction, recommendation generation, nearby-land analysis, and Explainable AI. The workflow proceeds through the stages described below.

A. Data Collection

A structured dataset of land attributes is used, covering geographical location, soil properties, climate conditions, terrain characteristics, water availability, environmental factors, and infrastructure accessibility. Together, these attributes describe each land record and feed the suitability prediction process.

B. Data Preprocessing

The collected data are preprocessed to raise their quality and make them ready for machine learning. This step covers treating missing values, detecting duplicate records, fixing inconsistent entries, and preparing the relevant features for training. The processed dataset is then split into training and testing sets so that model performance can be assessed.

C. Machine Learning Model Development

Random Forest, Extra Trees, XGBoost, and LightGBM are the machine learning algorithms used to learn the relationship between land characteristics and suitability outcomes. Model quality is measured with Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and the coefficient of determination (R^2), and, where applicable, accuracy, precision, recall, and F1-score are also used to judge the land-use recommendations. The model chosen for deployment depends on the evaluation results and on what the application requires.

D. Land Suitability Prediction

Given the land attributes as input, the trained model outputs suitability scores for the agriculture, residential, commercial, industrial, and mixed-use categories. The system compares these scores and recommends the category with the top predicted score. The output contains the score for each category, the recommended land use, and supporting information where it is available.

E. GIS-Based Spatial Analysis

GIS technology is built in to show suitability predictions on a map. Latitude and longitude coordinates place each land record on an interactive map. The system offers location-based prediction and nearby-land analysis, letting users browse neighboring records, compare their suitability scores, and review their recommendations. This helps users grasp how land suitability is distributed across space.

F. Explainable AI Using SHAP

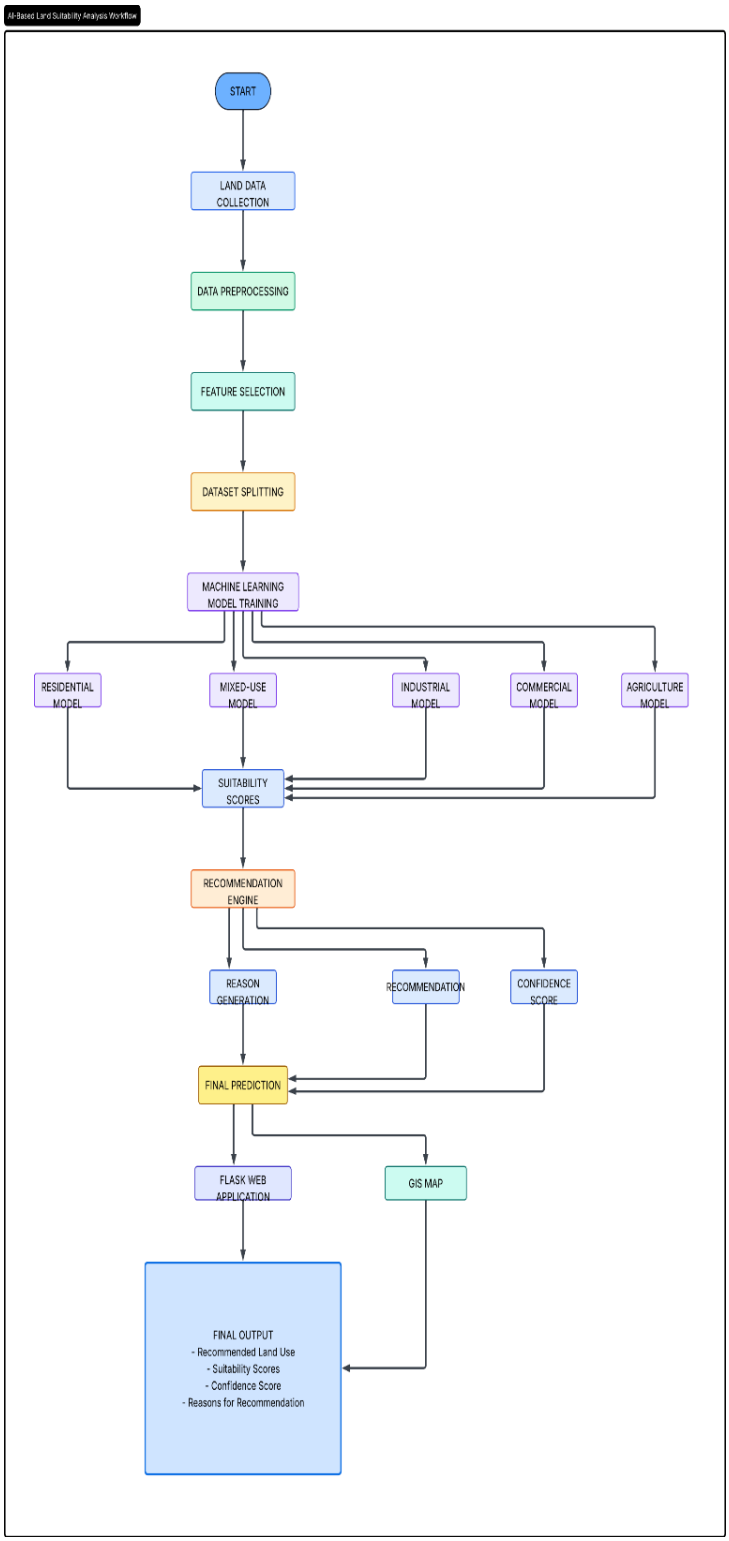
To interpret the machine learning predictions, the system uses SHapley Additive exPlanations (SHAP). SHAP quantifies how much each input feature contributes to a model's output, so users can see which land characteristics shaped a particular prediction. This makes the recommendation process easier to interpret and gives decision makers further information to work with.

G. Web-Based System Integration

The system is built with Python and Flask for backend processing and with HTML, CSS, and JavaScript for the frontend. Flask API endpoints link the user interface to the machine learning models and the land records, while Leaflet and configured map layers supply interactive geographic visualization. The interface shows suitability scores, recommendations, spatial information, and any available explanations in an accessible form.

WROK FLOW:-

Fig 1: Ai based land suitability analysis workflow



6. IMPLEMENTATION

The proposed AI-based land suitability system was implemented as a web-based GIS application that combines spatial data processing, machine learning prediction, recommendation, and interactive visualization.

6.1. Dataset Preparation

- A spatial land suitability dataset was assembled, with Land ID, latitude, longitude, soil, climate, terrain, water availability, infrastructure, accessibility, and environmental factors for each record.
- The data were cleaned and preprocessed prior to model training.
- Missing values, categorical attributes, and numerical features were each handled with suitable treatment.

6.2. Data Preprocessing

- Numerical features were made ready for model training.
- Categorical variables were encoded.
- Features relevant to suitability prediction were selected.
- Spatial coordinates were kept so that GIS-based visualization and location-based analysis remain possible.

6.3. Machine Learning Models

- A separate suitability prediction model was built for each of the following:
  - Agriculture
  - Residential
  - Commercial
  - Industrial
  - Mixed-Use
- Ensemble machine-learning algorithms (Extra Trees, XGBoost, LightGBM, and Random Forest) were evaluated.
- Together, the models produce suitability scores for the five land-use categories.

6.4. Recommendation Engine

- The predicted suitability scores are fed into the recommendation engine.
- The engine compares the scores of all five categories.
- The category with the highest predicted suitability becomes the recommended land use.
- Each recommendation is accompanied by a confidence score and supporting reasons.

6.5. Flask Web Application

- A Flask-based backend exposing REST APIs was developed.
- The backend is responsible for:
  - Land ID prediction
  - Latitude/longitude-based prediction
  - Land dataset retrieval
  - SHAP explainability
  - GIS map integration
- The frontend talks to the Flask backend by means of API requests.

6.6. GIS Visualization

- An interactive GIS map was built with Leaflet and Mapbox/OpenStreetMap-based map layers.
- Land locations appear on the map according to their geographic coordinates.
- The map can display the different suitability categories.
- A user can pick a location or enter a land ID to get its suitability analysis.

6.7. Explainable AI

- SHAP-based explainability was added to reveal which features influence each prediction.
- For any individual land prediction, the system can supply feature importance and explanations.

6.8. Final Output

For every land parcel analyzed, the system returns:

- Land ID
- Location
- Agriculture suitability score
- Residential suitability score
- Commercial suitability score
- Industrial suitability score
- Mixed-Use suitability score
- Recommended land use
- Confidence score
- Reasons for recommendation
- GIS location and visualization

7.RESULT AND ANALYSIS

The system was evaluated on the four machine-learning models using regression and recommendation performance metrics. The results indicate that ensemble-based models can predict land-use suitability scores effectively and produce sound land-use recommendations

Table I. Performance Evaluation of Machine Learning Models for Land Suitability Prediction

	Model	MAE	RMSE	R^2	Recommendation Accuracy	Precision	Recall	F1-Score
1	Extra Trees	1.1452	1.7824	0.9693	91.25%	91.33%	91.25%	91.25%
2	XG Boost	1.4033	1.8753	0.9670	88.50%	88.59%	88.50%	88.50%
3	Light GBM	1.4420	1.8878	0.9665	88.20%	88.30%	88.20%	88.22%
4	Random Forest	2.3686	3.2521	0.9036	86.40%	87.06%	86.40%	86.21%

Analysis:

- Extra Trees reached the highest recommendation accuracy of the evaluated models, at 91.25%.
- It also recorded the highest R^2 value (0.9693), which indicates a strong fit for the suitability-score prediction task.
- Extra Trees likewise had the lowest MAE (1.1452) and RMSE (1.7824).
- XGBoost and LightGBM also performed well, with recommendation accuracies of 88.50% and 88.20%, respectively.
- Random Forest performed comparatively worse, with an R^2 of 0.9036 and a recommendation accuracy of 86.40%.
- The GIS interface incorporates the prediction results, presenting land-use suitability scores, recommendations, confidence values, and spatial visualization.
- Five land-use categories are supported: Agriculture, Residential, Commercial, Industrial, and Mixed-Use.

Second Evaluation:

A second evaluation gave the results below:

Table II. Performance Comparison of Machine Learning Models in the Second Evaluation

	Model	MAE	RMSE	R^2	Recommendation Accuracy
1	**LightGBM**	1.7405	2.2126	0.9544	87.17%
2	**XGBoost**	1.9546	2.4887	0.9432	86.33%
3	**Extra Trees**	3.0650	3.8812	0.8614	77.50%
4	**Random Forest**	3.4484	4.3093	0.8330	76.67%

In this second evaluation, LightGBM delivered the best results of the four models, with a recommendation accuracy of 87.17% and an R^2 of 0.9544

Advantages:

1. Improved Land Assessment: Assesses land suitability for several purposes at once.
2. Accurate Predictions: Relies on machine learning models to estimate suitability.
3. Multiple Recommendations: Covers agriculture, residential, commercial, industrial, and mixed-use planning.
4. GIS Visualization: Shows land locations and suitability results on an interactive map.
5. Explainable Results: Applies SHAP to pinpoint the factors behind each prediction.
6. Time Saving: Cuts the manual effort needed for land assessment.
7. Better Decision-Making: Lets users compare suitability scores before choosing a land-use category.
8. User-Friendly Interface: Delivers predictions and map-based results through a web application.

Limitations:

1. Data Dependency: Prediction accuracy is tied to the quality and availability of the dataset.
2. Limited Data Coverage: Missing or out-of-date land information can distort predictions.
3. Model Limitations: Machine learning models cannot be expected to predict every land condition correctly.
4. Geographical Accuracy: Wrong latitude or longitude values can lead to errors in map visualization.
5. Environmental Changes: Shifts in climate, soil, water availability, and land use can alter suitability.
6. Limited Real-Time Data: Without real-time data, the system may not match current conditions on the ground.
7. Prediction Interpretability: SHAP explains how features contribute but does not guarantee that a recommendation is right.
8. Decision Support Only: The system offers recommendations; final land-use decisions still need field verification and expert assessment.

8. CONCLUSION AND FUTURE SCOPE

The AI-Based Land Suitability Analysis and Recommendation System combines machine learning with Geographic Information System (GIS) technologies to assess land for agricultural, residential, commercial, industrial, and mixed-use purposes. It produces suitability scores, recommends the most appropriate land-use category, and presents the results on an interactive map. SHAP-based explanations let users see which factors shaped each prediction. The system enables quicker land assessment and better-informed decisions; its results, however, hinge on dataset quality, geographical accuracy, and rigorous model evaluation.

The system could be improved in the following ways:

I. **Real-Time Data Integration:** Bring in up-to-date soil, weather, and environmental data.
II. **Satellite Image Analysis:** Apply remote sensing and satellite imagery to strengthen land assessment.
III. **Advanced Machine Learning:** Investigate further algorithms and optimization techniques to boost prediction performance.

IV. **Improved GIS Mapping:** Raise geographical accuracy and offer finer, land-level visualization.
V. **Mobile Application:** Build a mobile application so land suitability results are easy to access.
VI. **Additional Risk Analysis:** Add flood risk, soil erosion, drought, and other environmental hazards to the analysis.
VII. **Field Validation:** Check system predictions against real-world land conditions and expert assessments.
VIII. **Expanded Geographic Coverage:** Evaluate the system on datasets from other regions to widen its applicability

REFERENCES

1. T. N. Prakash, "Land Suitability Analysis for Agricultural Crops: A fuzzy multicriteria decision making approach," 2003.
2. G. S. Sinchana, A. L. Choodarathnakara, and G. A. Arpitha, "A review on land suitability mapping using geospatial technology," *The Mysore Journal of Agricultural Sciences*, vol. 57, no. 3, pp. 55–79, 2023.
3. T. Bendahmane, K. Rezeg, A. Merizig, O. Kazar, and S. Harous, "IoT and Deep learning-based approach for an efficient land suitability prediction in smart farms," *International Journal of Computing and Digital Systems*, vol. 17, no. 1, pp. 1–14, Jan. 2025.
4. A. A. Mustafa *et al.*, "Land suitability analysis for different crops: A Multi Criteria Decision Making Approach using Remote Sensing and GIS," *Researcher*, vol. 3, no. 12, pp. 61–63, 2011.
5. R. a. E. Behairy *et al.*, "Artificial Intelligence Integrated GIS for land suitability assessment of wheat crop growth in arid zones to sustain food security," *Agronomy*, vol. 13, no. 5, p. 1281, Apr. 2023.
6. D. R. Vincent, N. Deepa, D. Elavarasan, K. Srinivasan, S. H. Chauhdary, and C. Iwendi, "Sensors driven AI-Based Agriculture Recommendation model for assessing land suitability," *Sensors*, vol. 19, no. 17, p. 3667, Aug. 2019.
7. N. Kumar and D. K. Gowda, "Sensors driven AI-Based Agriculture Recommendation model for assessing land suitability," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 12, no. 5, pp. 399–400, May 2023.
8. H. Mugiyo *et al.*, "Evaluation of Land Suitability Methods with Reference to Neglected and Underutilised Crop Species: A Scoping Review," *Land*, vol. 10, no. 2, p. 125, Jan. 2021.
9. A. S, M. S, and S. R, "AI-based Smart Crop Recommendation System for Sustainable Agricultural Production: a data-driven approach to minimize resource use and maximize yield," *Madras Agricultural Journal*, vol. 112, no. 4–6, pp. 135–139, Jan. 2025.
10. N. Abdullah, A. Zolkafli, and M. Omar, "ADVANCEMENTS AND INTEGRATION OF BIOPHYSICAL, SOCIO-ECONOMIC, AND LOCAL KNOWLEDGE IN AGRICULTURAL LAND SUITABILITY ASSESSMENTS: A SYSTEMATIC LITERATURE REVIEW," *Mağallaẗ Zirāʿaẗ Al-rāfidayn*, vol. 53, no. 1, pp. 168–192.
11. B. A. Ganati and T. M. Sitote, "Predicting land suitability for wheat and barley crops using machine learning techniques," *Scientific Reports*, vol. 15, no. 1, p. 15879, May 2025.
12. D. R. Vincent, N. Deepa, D. Elavarasan, K. Srinivasan, S. H. Chauhdary, and C. Iwendi, "Sensors driven AI-Based Agriculture Recommendation model for assessing land suitability," *Sensors*, vol. 19, no. 17, p. 3667.