

An Experimental Investigation on Accessible Machine Learning Models for Examining the Possibility of Environmental Changes On Land for Farming Relevance

Mr.¹Vinay Patel G L, ²Sahana S S

¹Assistant Professor, Department of MCA, BIET, Davanagere

²Student, 4th Semester MCA, Department of MCA, BIET, Davanagere

ABSTRACT

This research explores the pressing global challenge of food security in the context of climate change, with a focus on its impact on agricultural land suitability. The primary goal is to forecast potential risks linked to declining land suitability and shifting irrigation trends, which have direct implications for global food availability. Aligned with the United Nations' Sustainable Development Goals to combat hunger and malnutrition, this study concentrates on Central Eurasia—a region marked by distinct socio-economic vulnerabilities. Using interpretable machine learning models, we examine how various carbon emission scenarios influence land suitability for agriculture. Our predictive model achieves an accuracy of 86% and a mean average precision of 72% in a multi-class classification of land suitability. By highlighting the most at-risk zones in Eastern Europe and Northern Asia, the findings offer valuable guidance for policymakers. These insights support data-driven decisions related to the efficient distribution of essential resources such as water and fertilizers, helping to avert future food crises. Overall, the study confirms the potential of machine learning in anticipating and mitigating the adverse effects of climate change on food production systems.

Keywords: *Climate Change, Agricultural Land Suitability, Machine Learning, Environmental Impact, Cropland Prediction, Sustainable Farming.*

I. INTRODUCTION

Agricultural sustainability is increasingly being threatened by the accelerating pace of environmental and climatic changes. The implications of climate change on the usability and productivity of agricultural land have emerged as a critical area of concern, especially in regions heavily dependent on farming for economic stability and food security. Traditional methods of assessing land suitability often fall short in capturing the complex and dynamic interactions between climatic variables and land characteristics. In this context, the application of accessible and interpretable machine learning (ML) models offers a promising solution for examining environmental influences on land relevance for

farming. This study is a forward looking investigation into the utility of machine learning approaches to predict changes in agricultural land suitability under varying environmental scenarios. By integrating climate data, soil characteristics, and historical land use patterns, the research explores how ML models can be utilized not only to assess current suitability but

also to forecast future changes that could impact farming relevance. Our objective is to enable early detection of risk-prone regions and facilitate informed decision-making for agricultural planning and resource allocation. Focusing on Central Eurasia as a case study—an area vulnerable to environmental stress and undergoing rapid climatic shifts—we employ a combination of data-driven models and interpretable ML techniques to identify

key factors influencing land degradation and transformation. The significance of this research lies in its potential to support global food security goals, reduce the risk of land mismanagement, and contribute to the development of adaptive strategies for sustainable agriculture. The proposed model is not only capable of making accurate predictions but also emphasizes explainability, thereby ensuring its applicability for policymakers, environmentalists, and agronomists alike.

II. RELATED WORK

Understanding the impact of environmental and climatic changes on agricultural land suitability has gained increasing attention in recent years. Researchers have explored a variety of modeling approaches, particularly through climate simulation and data-driven techniques, to assess potential land degradation and shifts in farming potential.[1]

The **Climate Model Intercomparison Project Phase 6 (CMIP6)** provides foundational projections for assessing future climate behavior. These projections simulate long-term climate trends, incorporating scenarios such as RCP 4.5, RCP 6.0, and RCP 8.5. Shoaib et al. used CMIP data to analyze crop yield changes in China using linear regression models based on rainfall and temperature trends. Similarly, Müller et al. examined multiple CMIP5 and CMIP6 datasets to evaluate the effect of pressure and temperature variations on crop productivity.[2]

In terms of datasets, resources like the **Global Food Security Support Analysis Data (GFSAD1km)** offer high-resolution insights into cropland distribution and irrigation patterns. The **ERA5** dataset and **Land Use MIP (LUMIP)** projections also provide valuable indicators such as soil types and land cover fractions, essential for modeling agricultural suitability.

Machine learning models have been applied extensively to this domain. Studies by **Dikshit et al.** assessed several ML algorithms—including Artificial Neural Networks (ANN), Support Vector

Machines (SVM), Random Forests (RF), and Adaptive Neuro-Fuzzy Inference Systems (ANFIS)—for drought prediction across various regions. Deep learning models, as explored by Dharani et al., have also proven effective for crop yield classification and regression tasks.

In the area of land cover prediction, **Diaconu et al.** implemented ConvLSTM networks to predict vegetation indices from satellite imagery. Additionally, **Yadav et al.** utilized SVM and RF to assess soil fertility, while **Hounkpatin et al.** employed classical ML techniques to investigate soil chemical properties in West Africa.

Despite the progress, current models face notable challenges:

Complexity of environmental data makes model training and interpretation difficult.

Limited data availability inconsistencies in labeling can reduce prediction accuracy.

Scalability and explainability of advanced ML models are often insufficient for direct policy applications.

This research addresses these limitations by developing an interpretable and accessible ML framework to examine climate-driven land suitability changes, focusing on ease of use, scalability, and actionable insightstechniques improved within-project performance but yielded mixed results in cross-project scenarios due to data heterogeneity across software systems.[3]

To improve generalization across projects, researchers like Zhang et al. and Kaliraj et al. explored the use of ensemble learning and optimized feature selection, respectively. Ensemble methods combined the strengths of multiple classifiers to enhance predictive robustness, while optimized feature selection aimed to identify universal software metrics that are effective across diverse projects[4]. Further

innovations included Bayesian Regularization techniques by Mahajan et al.

III. EXISTING SYSTEM

The Climate Model Intercomparison Project Phase 6 (CMIP6), launched in 2013, serves as a comprehensive platform for simulating future climate scenarios through numerical modeling. These simulations provide insight into long-term trends in Earth's atmospheric, terrestrial, and oceanic systems from 2015 to 2100, influenced by projected human activities. Shoaib et al. utilized CMIP data to assess how crop yields in China might vary under different climate conditions—specifically RCP 4.5, RCP 6.0, and RCP 8.5—by applying linear regression on temperature and precipitation data derived from the World Bank.

Similarly, Müller et al. conducted an extensive evaluation of 79 climate projections from CMIP5 and CMIP6, analyzing pressure and temperature trends and their consequences on agricultural productivity. Complementing this, the Intergovernmental Panel on Climate Change (IPCC) provided a seminal report offering high-level guidance on managing rising global temperatures and preserving cropland suitability. A common methodology in climate research involves combining outputs from multiple CMIP models to form ensemble predictions. These ensembles estimate the average and variability of climate variables across models, producing more robust and comprehensive forecasts.

In the domain of open-access data, several high-resolution datasets provide valuable information on land suitability. For instance, the Global Food Security Support Analysis Data (GFSAD1km) offers 1-kilometer resolution land classification based on the 2010 landscape, distinguishing between irrigated and non-irrigated croplands. The ERA5 dataset contributes additional context by including variables like soil types, useful for

identifying regions with organic-rich soils. Furthermore, the Land Use Model Intercomparison Project (LUMIP)—a CMIP extension—offers annual projections of land cover types, such as cropland, pasture, and urban areas, through the fracLut variable.

Geospatial data analysis has become a vital tool in this field, supported by advancements in machine learning and deep learning techniques. Recent studies have applied models like Long Short-Term Memory (LSTM) and Multi-Layer Perceptron (MLP) for climate prediction tasks, such as temperature forecasting. Dikshit et al. evaluated several machine learning models—including ANN, SVM, ANFIS, Extreme Learning Machines (ELM), decision trees, and random forests—for their effectiveness in drought prediction across different continents.

Additionally, Dharani et al. investigated deep learning techniques, emphasizing their potential in crop yield prediction through classification and regression frameworks. Diaconu et al. employed ConvLSTM networks to forecast vegetation indices (e.g., NDVI) and RGB imagery for land cover analysis. In the soil domain, Yadav et al. demonstrated how SVM and RF models could be used to evaluate soil fertility. Hounkpatin et al. further validated the application of traditional ML techniques to analyze the chemical composition of soil in Benin.

Limitations of Existing Systems

Despite significant progress, existing systems face several challenges:

Data Complexity: Many machine learning models struggle with interpreting large-scale, multidimensional climate and geospatial datasets.

Data Scarcity: The predictive accuracy of ML models heavily depends on the volume and quality of available data, which may be insufficient or inconsistent in certain regions.

Labeling Errors: The reliability of predictions hinges on properly labeled training data. Mislabeling or noise in the dataset can significantly degrade model performance.

IV. PROPOSED SYSTEM

This study introduces a meta-classification framework aimed at identifying the key determinants that influence agricultural land suitability, including changes in irrigation practices and land use configurations. The proposed approach provides critical insights into how land suitability is expected to evolve over the coming decades by isolating the primary environmental and climatic factors driving these changes.

Specifically, the research: Analyzes the relationship between climatic indicators and the associated risks to productive land use; Investigates the trends in agricultural land development, focusing on emerging irrigation methods and potential degradation in land viability; Applies both machine learning and deep learning techniques to uncover the variables most impactful in determining farming potential.

The findings serve as a strategic resource for policymakers and land management professionals, offering evidence-based guidance to create adaptive plans that respond to the intertwined effects of climate change and agricultural productivity.

Advantages

The proposed system follows a structured three-phase methodology:

Data Collection and Preprocessing – Involves gathering and cleaning historical and projected climate and agricultural data.

Model Training – Employs accessible and interpretable ML algorithms to learn from the processed data.

Result Evaluation and Prediction – Uses the trained models to simulate future cropland distribution under various Shared Socioeconomic Pathways (SSP) and climate scenarios.

This framework delivers accurate, data-driven forecasts, leverages existing datasets effectively, and facilitates reliable projections of agricultural land suitability for future planning.

System Architecture

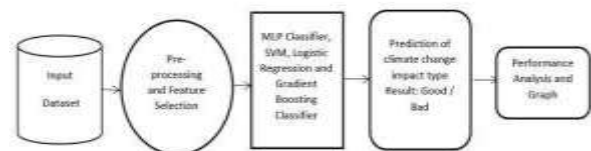


Fig 1. System Architecture

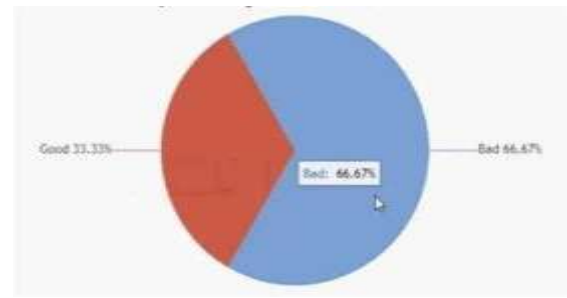
V. MODULE DESCRIPTION MODULES DESCRIPTION:

The system is structured into two main modules: the **Remote User Module** and the **Service Provider Module**. The Remote User Module is designed for individuals such as farmers, researchers, and agricultural analysts who seek to assess the impact of climate change on agricultural land suitability. Users can register, manage their profiles, and input data related to their geographical region, soil characteristics, and climate variables. Based on this input, the system utilizes interpretable machine learning models to predict the suitability of the land under future climate conditions. The prediction result is categorized as either "Good" or "Bad," indicating whether the land is favorable for agriculture. In addition, the system provides interpretable insights, highlighting the key environmental features that influenced the prediction, thereby helping users understand the reasoning behind the outcome.

The Service Provider Module, functioning as the administrator panel, is intended for backend users such as system administrators and research analysts. This module offers full access to user management, where all registered users and their activities can be monitored. It also maintains a repository of all prediction records, including the input parameters and the generated outcomes. A key feature of this module is the visualization dashboard, which presents data trends and prediction results through interactive graphs, charts, and geospatial maps. These visualizations enable a broader understanding of land suitability patterns across various regions and scenarios. Moreover, the service provider has access to the complete prediction datasets and can download them for further analysis or model refinement. This modular structure ensures both usability for public users and control and transparency for administrators

VI. RESULT

The proposed system is structured into two primary modules: the Remote User module and the Service Provider module, each serving distinct roles within the application. The Remote User module is intended for farmers, agricultural researchers, or environmental analysts who need to assess the impact of climate change on land suitability. Within this module, users can manage their user profile, which includes registration, login, and personal detail management. The core functionality is the Prediction Page, where users input relevant environmental data such as temperature, rainfall, soil type, and vegetation index. The system processes this data through a trained machine learning model and outputs a prediction result, classified as either “Good” (indicating suitable conditions for farming) or “Bad” (indicating potential land degradation or unsuitability for agriculture).



The Service Provider module is designed for administrators, policymakers, or domain experts who oversee system usage and data trends. This module provides access to all user profiles and their associated prediction activities. It includes a comprehensive dashboard displaying all prediction results submitted by users. The module also features graphical representations, enabling detailed analysis of prediction patterns, regional trends, and climate impact statistics. Furthermore, the service provider can download datasets containing historical prediction data in structured formats for external analysis, research, or model retraining. Together, these modules create an integrated system that promotes climate-aware decision-making in agriculture through accessible machine learning technologies.

VII. CONCLUSION

This project presents an effective and intelligent solution for enhancing crop quality and managing weed infestation using advanced computational techniques. By integrating Machine Learning and Deep Learning models, the system achieves high accuracy in classifying crop health and identifying invasive weeds. The use of feature extraction methods, pre-trained deep neural networks, and dataset balancing techniques like SMOTE significantly improved model performance. The system not only provides reliable prediction results but also supports decision-making in modern agriculture through a user-friendly platform for both farmers and service providers. With strong results on benchmark datasets, the solution proves to be practical, scalable, and environmentally sustainable. It holds great potential to advance precision

farming, reduce chemical usage, and improve overall agricultural productivity.

REFERENCES

- J. A. Foley, R. DeFries, G. P. Asner, C. Barford, G. Bonan, S. R. Carpenter, F. S. Chapin, M. T. Coe, G. C. Daily, H. K. Gibbs, J. H. Helkowski, T. Holloway, E. A. Howard, C. J. Kucharik, C. Monfreda, J. A. Patz, I. C. Prentice, N. Ramankutty, and P. K. Snyder, "Global consequences of land use," *Science*, vol. 309, no. 5734, pp. 570–574, 2005.[1]
- H. C. J. Godfray, J. R. Beddington, I. R. Crute, L. Haddad, D. Lawrence, J. F. Muir, J. Pretty, S. Robinson, S. M. Thomas, and C. Toulmin, "Food security: The challenge of feeding 9 billion people," *Science*, vol. 327, no. 5967, pp. 812–818, 2010. 4 [2]
- D. Tilman, C. Balzer, J. Hill, and B. L. Befort, "Global food demand and the sustainable intensification of agriculture," *Proc. Nat. Acad. Sci. USA*, vol. 108, no. 50, pp. 20260–20264, Dec. 2011. [3]
- V. Gitz, A. Meybeck, L. Lipper, C. D. Young, and S. Braatz, "Climate change and food security: Risks and responses," *Food Agricult. Org. United Nations (FAO)*, Rome, Italy, Tech. Rep. 110, 2016, pp.2–4. [4]
- L. S. Huning and A. AghaKouchak, "Global snow drought hot spots and characteristics," *Proc. Nat. Acad. Sci. USA*, vol. 117, no. 33, pp. 19753–19759, Aug. 2020. [5]
- J. R. Porter, L. Xie, A. J. Challinor, K. Cochrane, S. M. Howden, M. M. Iqbal, D. B. Lobell, and M. I. Travasso, *Food Security and Food Production*. Cambridge, U.K.: Cambridge Univ. Press, Jan. 2014, pp. 485–533. [6]
- K. Riahi et al., "The shared socioeconomic pathways and their energy, land use, and greenhouse gas emissions implications: An overview," *Global Environ. change*, vol. 42, pp. 153–168, Jan. 2017. [7]
- D. Peano, T. Lovato, and S. Materia, "CMCC-ESM2 model output prepared for CMIP6 LS3MIP," *Earth Syst. Grid Fed.*, Rome, Italy, Tech. Rep., 2020. [Online]. Available: <https://www.ipcc.ch/srccl/cite-report/> [8]
- A. Voldoire et al., "Evaluation of CMIP6 deck experiments with CNRMCM6-1," *J. Adv. Model. Earth Syst.*, vol. 11, no. 7, pp. 2177–2213, 2019. [9]
- S. Yukimoto, H. Kawai, T. Koshiro, N. Oshima, K. Yoshida, S. Urakawa, H. Tsujino, M. Deushi, T. Tanaka, M. Hosaka, S. Yabu, H. Yoshimura, E. Shindo, R. Mizuta, A. Obata, Y. Adachi, and M. Ishii, "The meteorological research institute Earth system model version 2.0, MRIESM2.0: Description and basic evaluation of the physical component," *J. Meteorol. Soc. Japan. Ser. II*, vol. 97, no. 5, pp. 931–965, 2019.[10]
