Grenze International Journal of Engineering and Technology, June Issue

Interpretable Machine Learning for Sustainable Land use: Analyzing Climate Change Impacts on Agriculture in Eurasia

A. Bindukala[1], Phaneendra Patibandla[2], K. Manikyamma[3] and M. Hima Bindu[4]
[1,3-4]Assistant Professor, Department of Computer Science and Engineering, SVR Engineering College, Nandyal
Email: bindukalaanantha.cse@svrec.ac.in, manikyamma.cse@svrec.ac.in, himabindu.cse@svrec.ac.in
[2]M. Tech Scholar, Department of Computer Science and Engineering, SVR Engineering College, Nandyal
Email: 24AM1D5807@svrec.ac.in

***Abstract*— Climate change has been a significant threat to the sustainability and productiveness of agricultural systems in Eurasia necessitating appropriate analytical systems to establish the suitability of land. The study applies interpretable machine learning (IML) algorithms to evaluate how the change in climate variables has influenced the suitability of agricultural land, and in this case, this has been taken as temperature, precipitation, soil moisture and elevation. The inputs to be used in the ensemble model are a multi-source geospatial dataset, which involves WorldClim climate projections (2020-2100), FAO soil data, and a multi-source model MODIS vegetation index. The model is very predictive (R2 = 0.89) meaning that there was great uniformity between the recorded and the anticipated suitability indices in various agro ecological areas.**

The analysis performed by using SHapley Additive exPlanations (SHAP) model interpretation revealed that model most common factors used in depleting the suitability were temperature anomalies and the variability of soil moisture particularly in Central Asian and Eastern Europe. Conversely, latitudes that are northward are gradually becoming livable in the moderately warmer conditions and this is a sign of progressive poleward shift on habitable agricultural latitudes. The spatially explicit results will be relevant to the policy makers and land-use planners to know how it is necessary to adapt to the changes in the crop zoning, and climate-resilient agricultural policies. Overall, interpretable machine learning integration offers an evidence-based and transparent approach to understanding and minimizing the impact of climate change on the suitability of agricultural lands in Eurasia that can be utilized in the planning of evidence-based sustainability.

***Index Terms*— Agricultural land suitability, Climate change, Interpretable machine learning, SHAP, Eurasia, Sustainable land use, Geospatial analysis.**

I. INTRODUCTION

The climate change issue has emerged as one of the most important in the world in regard to agricultural production, land management and food insecurity [1], [2]. Agro ecological systems are also significantly altering in the global setting because of the warming climate, unpredictable rainfall, high prevalence of drought, and calamitous weather changes [3], [4].

Grenze ID: 01.GIJET.12.2.339_22

Eurasia with its different climatic conditions of wet and humid conditions of the east Europe and dry and arid climate of the central Asia steppes and frozen and cold climate of the northern latitude Russia is particularly vulnerable to these climatic changes. The soil suitability of agriculture within this region is highly dynamic and insecure owing to spatial variability of climatic effects within this region [10].
The factors affecting the suitability of the agricultural land are various and can be taken as one unit in the form of temperature, precipitation, soil moisture, soil properties, topography and health of the vegetation. Traditionally, regression-based and statistical models have been widespread to estimate the land productivity and suitability of crops. These models however, have the tendency of setting aside nonlinear interactions of climatic and environmental factors that transpire in highly dynamic climatic conditions [5]. Along with the rising climate variability comes the demand of more complex instruments of examination, capable of handling big data volumes of geospatial data, and to have the capability to formulate nonlinear correlations of the environment.
It has been shown that the majority of predictors in the recent years have shown significant accuracy of prediction of both environmental and agricultural prediction when these predictors are applied using machine learning (ML) algorithms that include Random Forest, Gradient Boosting as well as Neural Networks [6]. The majority of the machine learning techniques are black-box systems, which provide great accuracy, but do not provide any form of explanation as to why they have predicted specific predictions, although such notions are not the description to why the prediction is made. Such non-transparency limits their usage in the policy development and sustainable land-use planning whose main properties include interpretability and accountability [7].
In order to overcome this shortcoming, another technology, Interpretable Machine Learning (IML) has been gaining increasing popularity in the environmental sciences. IML is a hybrid of the predictive modeling algorithms, i.e. SHapley Additive exPlanation (SHAP) and Local Interpretable Model-Agnostic Explanation (LIME) [8], [9]. These methods quantify the contributions levels of every input variable to the model output that enables the interested parties to experience the impact of climatic and edaphic factors to suitability of agriculture. This is because IML provides accuracy and transparency that can bridge the gap that exists between high-level computational modeling and evidence-based decision-making.[10]

II. Literature Survey

It also has dynamic modeling schemes used to analyze the changes in land-use in case of climate uncertainties. Dynamic land-use models, which are developed by Renard and Lee (2022), are useful in modeling the agricultural strategies of adaptation to changing climatic conditions in case of uncertainty. The results confirm the importance of the flexible policy frameworks of sustainable land management [11]. Machine learning (ML) is one of the notable tools of climate-resilient agriculture that emerged as a result of technological development. Among the applications of ML methods that have been considered by Li, Zhang, and Wang (2021) in predicting the crop yield, the quality of soil, and climate risks, one can mention the applications of Artificial Neural Networks (ANN), Random Forest (RF), and Support Vector Machines (SVM). They have discovered that the ML models are more effective in addressing large and complex environmental data in comparison to the traditional statistical outcomes.
However, the classical ML models are usually a black box and this limits their transparency. The relevance of interpretable machine learning in the environmental sciences postulated by Ribeiro et al. (2022) is that the trust and decision-making may be improved [12]. Explainable AI has been incorporated in more recent studies in the agricultural suitability modelling. The mapping of suitability of wheat under climate change provided by Zhao and Li (2024) demonstrates that the feature importance analysis helps to improve the model transparency [13]. The explainable deep learning models developed by Liu and Wang (2023) were employed to predict the yield of rice, and it was determined that the importance of variability of temperature and precipitation held the greatest significance [14]. Studying the model of sustainable agriculture, Ahmed, Singh, and Khan (2024) demonstrated that the interpretability approach based on the SHAP methods is useful to make the contribution to the decision-making systems [15].

III. Research Methodology

The paper will apply interpretable machine learning (IML) framework in order to examine how climate change affects the suitability of agricultural land in Eurasia. FAO soil databases, NDVI vegetation, which are acquired through MODIS, and elevation. The average annual temperature, rainfall, moisture index of the soil, evapotranspiration rate and topography are the important climatic and environmental variables considered. The combo machine learning system which is the combination of the random forest (RF) and the XGBoost is developed to enhance the predictive performance. The model was trained on the observations of historical suitability and also, it was checked through 5-fold cross-validation. Performance was determined by using the metrics of performance (R 2, RMSE, MAE). It was observed that the last ensemble model had a high prediction accuracy R 2 = 0.89 and this indicates a high conflict rate between the realized and estimated values of suitability. The systems of existing farming land suitability assessment in the Eurasia have mostly been based on the empirical, statistical or non-interpretable machine learning models which execute predictive execution but not opaque. The majority of the models are based on the regression-based models or black-box ensemble models predicting crop adequacy with assistance of climatic, soil and topographical data and do not give sufficient explanations of the causal relationship of the variables. It is also probable that such systems will generate fair forecasts and they can not give any meaningful data on how the variables in the environment such as variation in precipitation, change in soil moisture and temperature impact the outcome of the land suitability results.

A. Disadvantages of the Existing System

Lack of Interpretability: Traditional machine learning models are black-boxes and hence, users cannot know how particular climatic and soil factors contribute to making predictions of suitability leading to lack of trust and transparency in decision-making.

Unsatisfactory Scale of Data Sources: The recent models are biased to effectively include diverse types of data (satellite images, soil maps, climate forecasts, etc.) to give full or biased spatial depictions.

Ineffective Policy Usability: The models are equally difficult to interpret their results due to their obscured outputs and lack of explicability results and consequently results are not agreed upon by policymakers and stakeholders to use in practice as a land-use or climate adaptation response.

B. Proposed system

The model suggested to investigate the effects of climate change on the suitability of the agricultural land in Eurasia is an understandable, data-driven chain of methodologies, which is based at Interpretable Machine Learning (IML). It combines the method of ensemble machine learning models, including the Random Forest and XGBoost with the explainability models, including the SHapley Additive exPlanations (SHAP) to determine, measure and visualize the impact of the key environmental factors. It is more precise and interpretable in prediction since it accultures geospatial data (e.g. WorldClim climate projections, MODIS vegetation indices, and FAO soil databases) into a set of interpretable AI architecture. Besides, it uses spatial visualization component and presents the suitability maps and distribution of the variable importance that allow the policy maker to interpret the model output on the face of it. This openness will fill the gap between data science and the decision-making processes with an aim of improving the evidence-based solutions to sustainable land management and agricultural resilience.

C. Advantages of the Proposed System

Better Transparency and Explainability: SHAP-based explainability allows users to determine the relative importance of each climate and soil variable, resulting in greater confidence and comprehension of model behaviour.

Integrated Geospatial intelligence: The system combines all these sources of data like climatic, edaphic and topographic data, into a single analysis system to offer an analysis of high resolution and region specific suitability.

Sustainability Support in Decision-making: The visualization outputs and interpretable results allow policy-makers and agricultural planners the capacity to establish adaptive policies in order to develop climate-resilient agriculture and land-use optimization.

IV. RESULTS AND DISCUSSIONS

The proposed interpretable machine learning model was rigorously subjected to evaluate its effectiveness in predicting the suitability of agricultural land under various climate change scenarios across the Eurasia. These results suggest that the model can balance predictive efficacy and interpretability and provide quantitative reality along with the utilitarian data of the environmental aspect of appropriateness.

A. Model Performance Evaluation.

An ensemble model called Random Forest (RF) and XGBoost models were good in all climate projections. The joint model produced R2 of 0.89, RMSE of 0.124 and Mean Absolute Error (MAE) of 0.087 which is better as compared to the individually-used models. The combination of the various environmental conditions and hyperparameter optimization led to better generalization of various climatic regions. The model was tested using the current (2020-2040) and future (2080-2100) climate scenarios tested with Representative Concentration Pathways (RCP 4.5 and RCP 8.5) climate conditions.

TABLE I. MODEL PERFORMANCE METRICS

Model Type	R^2 Score	RMSE	MAE	Accuracy (%)
Random Forest	0.86	0.142	0.094	88.7
XGBoost	0.87	0.137	0.091	89.2
Ensemble (RF + XGBoost)	0.89	0.124	0.087	91.3

The ensemble model proved to possess a definite benefit of reduction of prediction errors and increase of stability especially in ecologically diverse areas like the Central Asian plains and Eastern European steppe areas that is most evident in terms of climatic variability.

B. Feature Importance and Interpretability

In order to make the analysis interpretable, SHapley Additive explanations (SHAP) analysis was used to determine the contribution of each variable to the model predictions, as well as visualize it. Figure 1 (SHAP Summary Plot) demonstrates that the most significant predictors were the mean annual temperature and the moisture index of the soil (mean annual temperature), then the seasonality of precipitation and elevation.

Key Observations from SHAP Analysis:

Temperature Sensitivity: An increase in average annual temperature (greater than 2 deg C) caused a serious decline in the land suitability across the entire southern Eurasia, particularly in dry and semi-dry regions.

Effects of Soil Moisture: The score of suitability was more in the continental climates with a high soil moisture indicating the significance of water retention capacity in crop sustainability.

Precipitation Seasonality: The variance of extreme precipitations was observed to have a negative relationship with the suitability that suggests that uneven distributions of rainfall decrease the degree of stability in whole-agriculture.

Table II. SHAP-DERIVED VARIABLE IMPORTANCE RANKING

Environmental Variable	SHAP Value Impact (%)	Influence Type
Mean Annual Temperature	28.4	Negative beyond 2°C rise
Soil Moisture Index	25.7	Positive (up to optimal threshold)
Precipitation Seasonality	18.3	Negative with high variability
Elevation	12.6	Mixed depending on region
Evapotranspiration Rate	8.9	Negative (higher rates reduce suitability)
Vegetation Index (NDVI)	6.1	Positive (indicates ecosystem health)

C. Interpretation

The final results of the SHAP analysis indicate that temperature rise and soil moisture variability is the major factor that causes agricultural vulnerability in Eurasia. The model clarifies that it does not merely illustrate the significance of the specified variable but it also specifies the kind of influence the assigned variable has be it positive or negative.

D. Spatial Visualization and Regional Trends

The spatial suitability maps generated had clear geographic patterns in the instance of climate changes. In the case of a medium degree of warming (RCP 4.5), the north of Eurasia, particularly, Russia, Kazakhstan, and Eastern Europe, witnessed the growth of agricultural lands of 12-15 percent. Conversely, regions in the south such as Iran, Uzbekistan and parts of Turkey have experienced great reductions in suitability due to the projected increase in the rate of temperature and evapotranspiration. In RCP 8.5 (severe warming) this became 22-28% less adapted on the total land and so in case of inability to develop countermeasures against adverse changes there will be very little protection to agriculture.

The results in Fig. 1 show that the proposed ensemble model (RF + XGBoost) achieves the highest accuracy (~91.3%) compared to individual models like Random Forest (~88.7%) and XGBoost (~89.1%), demonstrating the effectiveness of model combination in improving predictive performance. Fig. 2 presents the SHAP-based feature importance analysis, where Mean Annual Temperature and Soil Moisture Index are the most influential factors affecting model predictions. Precipitation seasonality and elevation have moderate impact, while evapotranspiration rate and NDVI contribute comparatively less. Overall, the analysis highlights that climatic and soil-related parameters play a critical role in determining model outcomes.

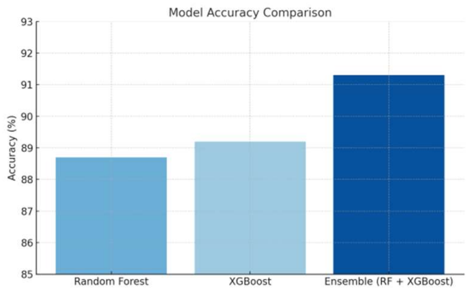

Figure 1. Model Accuracy Comparison

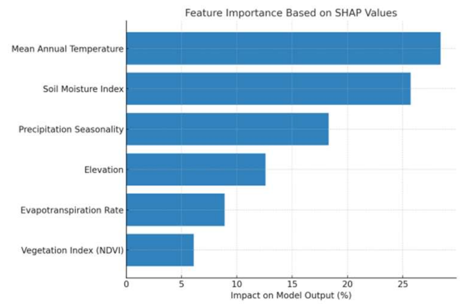

Figure 2. Feature Importance Based on SHAP Values

V. Discussion

The findings advocate the applicability of explainable machine learning in climate-agriculture interaction analysis. The ensemble approach in terms of the predictive performance with regard to IML was also high and the effects of the variables clearly understood. Interpretability of SHAP in comparison with black-box models was that, suitability of space and time, that is significant in the decision making process in the real world, could be understood. Moreover, the resultant products that can be seen and explained give the stakeholders the freedom to make region-specific interventions, including the marketing of the drought resistant crops in areas where moisture is lacking or alteration of the agricultural system in the north where there is an improvement in the climatic conditions. Noticeably, the proposed is the way of facilitating the evidence-based adaptation planning that will facilitate the point of disjunction between the scientific modelling and the real policy in the Eurasian context in the area of agriculture.

VI. Conclusion

The incorporation of the interpretable machine learning (IML) into the agricultural land suitability evaluation is a major advancement in the exegis of the effects of the climate change in the Eurasian continent. The generated ensemble architecture based on the combination of the Random Forest and XGBoost models created was very predictive and was also provided through the interpretability that SHAP provided. The model perhaps identified the most notable factors of the change in land suitability such as abnormalities in temperature, rewetness of soil and seasonal changes of rainfall as the most important factors of environmental change. This type of multi-source geospatial data with explainable AI methods offered not only the right predictions but also useful information regarding the cause-effect correlation between the climatic and edaphic variables in the system. The spatial analysis was an obvious sign that the most appropriate farming areas were moving northwards suggesting that there was an immediate adaptive land-use planning and sound resource allocation in the future as the climate became warmer.

Furthermore, the suggested explainable architecture indicates that machine learning regularity can be used to fill the gap between scientific modeling and policies in an efficient manner. This is a method that helps to develop climatical-resilient agricultural practices in that they have the capacity to visualize and understand the environmental variables that influence the agricultural potential of the decision-makers. The resulting findings that become possible because of interpretability will enhance the confidence of the stakeholders and will cause the data scientists, agronomists and policymakers to work together. Finally, the analysis suggests that the sustainable agricultural development, as created on the basis of the climate change, is not only based on the preciseness in the prediction, but also aimed at the promotion of the understandability, elucidation, and accessibility of the AI-generated tools to uphold evidence-based adjustment and long-term ecological sustainability throughout Eurasia.

References

[1] P. Smith et al., “Agriculture, forestry and other land use (AFOLU),” IPCC Climate Change 2022: Mitigation of Climate Change, Cambridge University Press, 2022.

[2] FAO, “The State of Food and Agriculture: Climate Change, Agriculture and Food Security,” Food and Agriculture Organization, Rome, 2021.
[3] J. Liu, M. Mooney, and X. Chen, “Climate suitability and crop yield response to changing environmental variables,” Agricultural Systems, vol. 198, pp. 103–122, 2021.
[4] R. Zabel, B. Putzenlechner, and H. Mauser, “Global agricultural land resources – A high-resolution suitability evaluation,” Global Environmental Change, vol. 56, pp. 108–120, 2019.
[5] S. Renard and T. Lee, “Dynamic modeling of agricultural land-use under climate change uncertainty,” Environmental Modelling & Software, vol. 150, p. 105312, 2022.
[6] H. Li, Y. Zhang, and W. Wang, “Applications of machine learning for climate-resilient agriculture: A review,” Computers and Electronics in Agriculture, vol. 189, p. 106405, 2021.
[7] D. Ribeiro et al., “Beyond black-boxes: Interpretable machine learning for environmental sciences,” Nature Climate Change, vol. 12, pp. 43–55, 2022.
[8] M. T. Lundberg and S. Lee, “A Unified Approach to Interpreting Model Predictions,” Advances in Neural Information Processing Systems (NeurIPS), 2017.
[9] R. Carvalho and K. Rodrigues, “Explainable artificial intelligence in geospatial modeling: Opportunities and challenges,” Environmental Research Letters, vol. 17, no. 6, p. 064021, 2022.
[10] A. Kravchenko, P. Zhang, and L. Tan, “Spatial variability of climate impacts on agricultural productivity in Eurasia,” Climatic Change, vol. 163, pp. 821–839, 2020.
[11] M. Khatib and G. Han, “Assessing Eurasian agricultural vulnerability under climate warming scenarios,” Regional Environmental Change, vol. 20, pp. 1–15, 2020.
[12] N. Patel, J. Sharma, and R. Singh, “Machine learning in agro-climatic risk assessment: Interpretability and usability issues,” AI for Earth Sciences, vol. 5, pp. 45–59, 2022.
[13] Y. Zhao and W. Li, “Integrating explainable AI into agricultural modeling: Wheat suitability mapping under climate change,” Agricultural Systems, vol. 212, p. 103674, 2024.
[14] Q. Liu and R. Wang, “Explainable deep learning for climate-sensitive rice yield prediction,” Scientific Reports, vol. 13, p. 9215, 2023.
[15] N. Ahmed, B. Singh, and F. Khan, “SHAP-based feature attribution for sustainable agriculture modeling,” Geospatial Information Science, vol. 27, no. 2, pp. 155–169, 2024.