KEC Conference 2026

ISSN: 3021-940X (Print)

DOI: 10.3126/injet.v3i2.95535

Predicting Fertilizer Quantity Using Geo-Spatial and Physicochemical Properties of Soil

Kriti Nyoupane[1], Santosh Giri[2*], Jebish Purbey[3]

[1]*Department of Electronics and Computer Engineering, Pulchowk Campus, IOE, TU, Lalitpur, Nepal,* kritineupane544@gmail.com
[2]*Department of Electronics and Computer Engineering, Pulchowk Campus, IOE, TU, Lalitpur, Nepal,* santoshgiri@pcampus.edu.np
[3]*Department of Electronics and Computer Engineering, Pulchowk Campus, IOE, TU, Lalitpur, Nepal,*74bex412.jebish@pcampus.edu.np

Abstract

Nepal's economy is highly dependent on agriculture, yet smallholder farmers benefit far less than they could because they rely on traditional practices or one-size-fits-all recommendations that ignore local soil and topographic variation. To address this, precision farming must be adopted nationwide, and the precise use of fertilizers for the right crop type is a critical part of this transition. This study presents a novel approach to recommend fertilizer quantities for the three major cereal crops, rice, maize, and wheat, including their OPV and hybrid subtypes, by modeling the problem as a multi-output regression task. The model uses the soil's physicochemical properties (pH, organic matter, total nitrogen, available phosphorus, available potassium, sand, silt, parent soil, clay, zinc, and boron) and geospatial properties (latitude, longitude, and elevation) as inputs. Around 386,000 soil data points from all provinces of Nepal were collected from the National Soil Science Research Center of Nepal Agricultural Research Council through a custom data extraction and transformation pipeline. A comparative analysis of Random Forest, XGBoost, and LightGBM was conducted against Linear and Ridge Regression baselines. Tree-based ensemble models consistently outperformed linear approaches, with LightGBM, XGBoost, and Random Forest achieving R^2 values of 0.976, 0.975, and 0.972, respectively, while both linear models only achieved 0.755. 10-fold cross-validation on a small subset suggests that performance across individual folds is consistent, rather than being overfit to the data. The proposed framework offers a scalable, data-driven path toward site-specific fertilizer recommendation across Nepal's diverse topography.

Keywords: Precision agriculture, fertilizer recommendation, multi-output regression, ensemble learning, soil physicochemical properties, geospatial analysis

1. Introduction

Agriculture is the backbone of Nepal's economy, employing more than 65% of the people and contributing approximately 38.81% to the national GDP [1]. Despite this dependence, smallholder farmers continue to operate under traditional practices or follow generic, one-size-fits-all fertilizer recommendations issued at the national level. These blanket recommendations, originally formulated decades ago, fail to reflect Nepal's highly diverse topography and the substantial variation in soil physicochemical properties across short distances. The result is inefficient fertilizer use, nutrient imbalance, soil degradation, and yield gaps that persist even where inputs are available [2].

Soil fertility in Nepal is known to vary strongly with elevation gradient and depth [3], yet most existing fertilizer prediction systems don't account for elevation as a parameter affecting soil fertility. This low spatial resolution misses key field-scale variability that drives nutrient dynamics. Recent advances in machine learning, particularly tree-based ensemble methods, offer a promising alternative by capturing the non-linear interactions between soil chemistry, micronutrients, soil texture, and topographic position. By learning directly from large geospatial soil datasets, such models can produce site-specific fertilizer recommendations at a resolution that blanket guidelines cannot match.

This study formulates fertilizer quantity recommendation as a multi-output regression problem, predicting Urea (split doses), DAP, MoP, organic matter, zinc, and boron simultaneously for rice, maize, and wheat (OPV and hybrid varieties) from soil physicochemical and geospatial inputs. The contributions of this work

* *Corresponding author*

are threefold. First, it introduces a topography-aware modeling framework that treats elevation as a continuous variable rather than a categorical zone. Second, it describes a reproducible pipeline that overcomes the single-query limitation of the NSSRC Soil Fertility API to assemble a high-quality dataset of approximately 386,000 samples covering all provinces of Nepal. Third, it formulates fertilizer prediction as a joint multi-output task, enabling balanced macronutrient and micronutrient recommendations within a single model.

2. Related Work

The application of machine learning to precision agriculture has grown rapidly in recent years, driven by the ability of these methods to model the non-linear relationships between soil characteristics, nutrient availability, and environmental conditions [4]. Polwaththa [5] further highlight that AI and ML techniques, when combined with IoT and remote-sensing data, can substantially improve crop yield prediction, irrigation planning, and fertilizer decision-making, while also identifying high implementation cost, limited farmer awareness, and rural infrastructure gaps as persistent barriers in developing countries. Ennaji [6] showed that ML methods can substantially improve nutrient use efficiency when trained on large, diverse datasets, while Awais [7] reported that Random Forest, Support Vector Machines, and neural networks outperform traditional statistical approaches for soil fertility prediction. However, both studies focus primarily on large commercial farms, leaving smallholder systems, which dominate countries like Nepal, underexplored.

Several recent platforms have attempted to deliver site-specific fertilizer advice. GeaGrow [8] uses neural networks with sensor inputs to estimate soil pH and macronutrients, while Musanase [9] developed a neural network for crop and fertilizer recommendation in Rwanda. Latha and Kumaresan [10] similarly demonstrated that integrating soil, weather, and pest data into deep-learning–based recommendation pipelines can achieve 90–93% accuracy for soil nutrient prediction and crop suggestion. Despite achieving high reported accuracy, these systems depend on dense sensor deployments or simulated training data, and they typically omit micronutrients, soil texture, and elevation variables that are critical in Himalayan soils.

Topographic heterogeneity remains substantially underexamined in the precision-agriculture literature. Standard ML models developed for industrial farms in flat regions often disregard elevation entirely [11], yet research in the Indian Himalayan region indicates that topography can explain up to 60% of the variation in soil property distribution, particularly soil depth and organic carbon [12]. Rengma [13] further demonstrate that altitude strongly influences soil organic matter and pH in mid-Himalayan terrain, reinforcing the need for elevation-aware models in Nepal.

Multi-output regression has been applied successfully to related soil-prediction problems. Kucuk [14] evaluated nine algorithms for soil moisture prediction across multiple soil depths, with Extra Tree Regression achieving the best R^2 of 0.81. Building on this body of work, the present study extends multi-output regression to fertilizer quantity prediction, integrates micronutrients and elevation explicitly, and operates at a national scale rarely addressed in prior literature.

3. Methodology

The proposed approach follows the pipeline shown in Figure 1, comprising data acquisition, preprocessing, exploratory analysis, multi-output model training, and evaluation.

3.1 Dataset Collection and Preprocessing

The dataset was obtained from the National Soil Science Research Center (NSSRC) Soil Fertility API. Because the API supports only single-location queries, a custom geospatial extraction pipeline was developed: GeoTIFF raster layers covering Nepal were sampled to generate approximately one million latitude–longitude coordinates, which were then queried sequentially to retrieve soil parameters and crop-with each soil and fertilizer attribute represented as a separate column. Measurement units (kg/ha, ppm, %) were stripped from cell values and appended to column headers to ensure that all numerical fields weretreated as continuous variables. Administrative identifiers (province, district, palika) were dropped

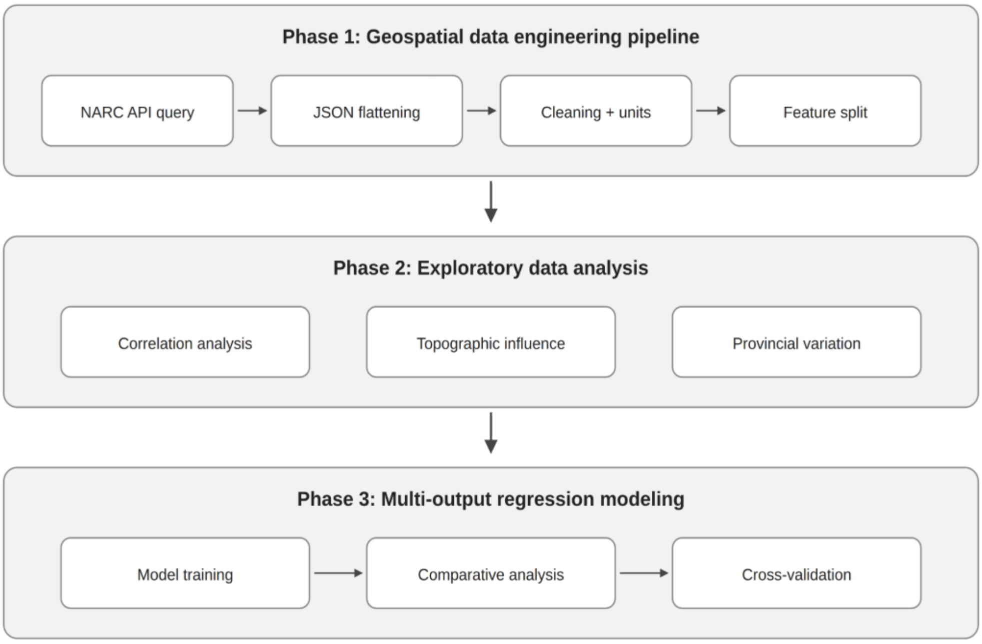

Figure 1. Methodology Pipeline

retained only for visualization. Records with missing values in essential soil or fertilizer fields were removed, reducing the dataset from 386,000 to 384,077 complete observations.

- **Input variables (X):** Longitude, latitude, elevation, pH, organic matter, total nitrogen, available potassium, available phosphorus, boron, zinc, sand, silt, and clay.
- **Target variables (Y):** Urea (split doses 1/2/3), DAP, MoP, organic fertiliser, zinc, and boron across maize (OPV and hybrid), rice (OPV and hybrid), and wheat (OPV).

Table 1. Summary of the Soil and Fertilizer Recommendation Dataset

Category	**Description**
Data Source	Nepal Agriculture Research Council (NARC) Soil API
Total Samples	386,000 raw → 384,077 after cleaning
Spatial Coverage	All 7 Provinces, 77 Districts, 742 Palikas
Geospatial Attributes	Latitude, Longitude, Elevation
Soil Physical Parameters	Sand, Silt and Clay (%), 9 Parent Soil Types (Colluvial Calcareous, Fluvial Non Calcareous, Fluvial Calcareous, Gneiss Migmatite, Lacustrine Non Calcareous, Quartzite, RK, Sandstone Greywacke Arkose, Slate Phyllite)
Soil Chemical Parameters	pH, Organic Matter (%), Nitrogen (%), Phosphorus (kg/ha), Potassium (kg/ha), Boron (ppm), Zinc (ppm)
Target Variables (Y)	40 variables: 5 crop-type combinations × 8 fertilizers (e.g., Maize OPV MOP, Rice Hybrid DAP, Wheat OPV UREA1)
Input Features (X)	14 variables: 3 geospatial + 7 chemical + 3 physical + 1 geological
Crop Types	Rice (OPV, Hybrid), Maize (OPV, Hybrid), Wheat (OPV)

Data Format	Nested JSON, flattened into tabular structure

3.2 Exploratory Data Analysis (EDA)

3.2.1Correlation Structure of Soil Properties

The correlation matrix in Figure 2 confirmed that there exists a strong relationship between soil parameters.

- The strongest correlation was observed between sand and silt content (r = 0.79), indicating an inverse relationship, in the study area, implying that an increase in sand content is accompanied by a sharp decrease in silt content.
- Organic matter and total nitrogen showed a strong positive correlation (r = 0.71), which aligns with established soil processes, as organic matter is the main source of mineralizable nitrogen.
- Elevation showed a very strong positive relationship with organic matter (r = 0.74).
- Elevation shows a strong positive relationship with potassium (r = 0.62) and total nitrogen (r = 0.57), suggesting that higher regions may be naturally more nutrient-dense or better at retaining these specific elements.

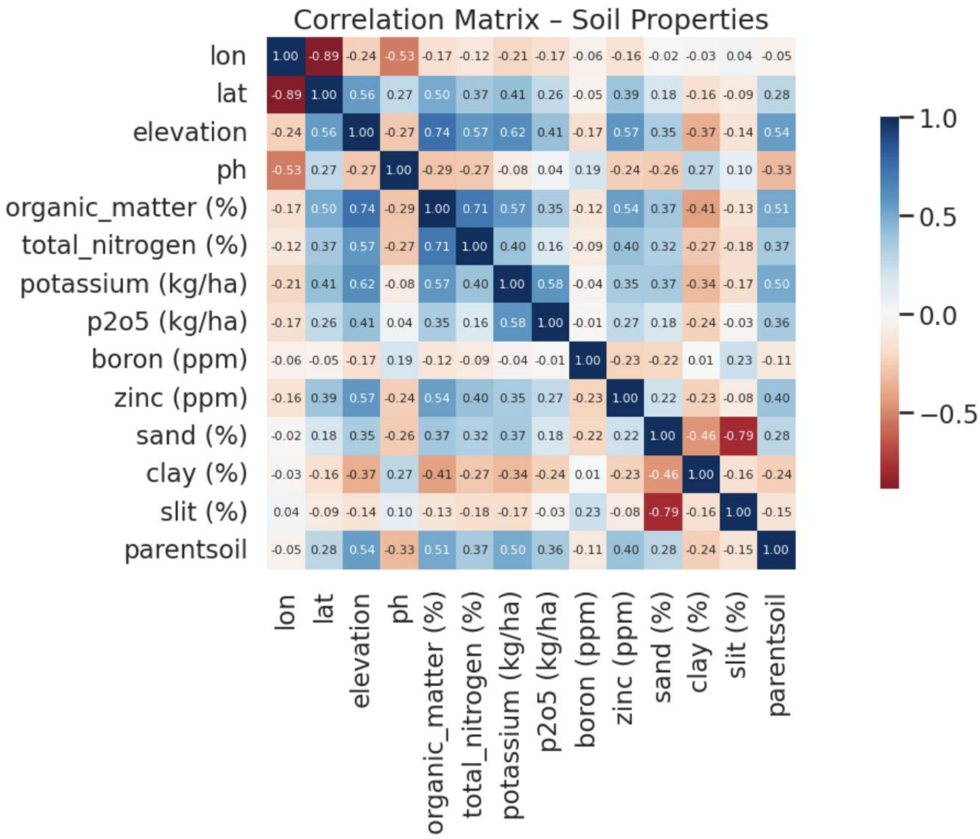

Figure 2. Correlation matrix of soil properties from around 384,000 preprocessed records.

3.2.2 Topographic Influence Visualizations

Spatial distribution maps were generated using QGIS to visualize the relationship between elevation and three key soil parameters (Figure 3, Figure 4, and Figure 5):

- Elevation vs. pH: A slight but consistent decline in pH was observed at higher altitudes, particularly in Eastern Nepal, consistent with increased leaching in wetter montane conditions.
- Elevation vs. organic matter: The sharpest elevation gradient was observed for organic matter, which increased substantially at higher altitudes due to slower decomposition in colder climates.

- Elevation vs. nitrogen: Nitrogen increased marginally with elevation, mostly in standard amounts among hill areas, most likely due to enhanced organic matter retention at higher altitudes.

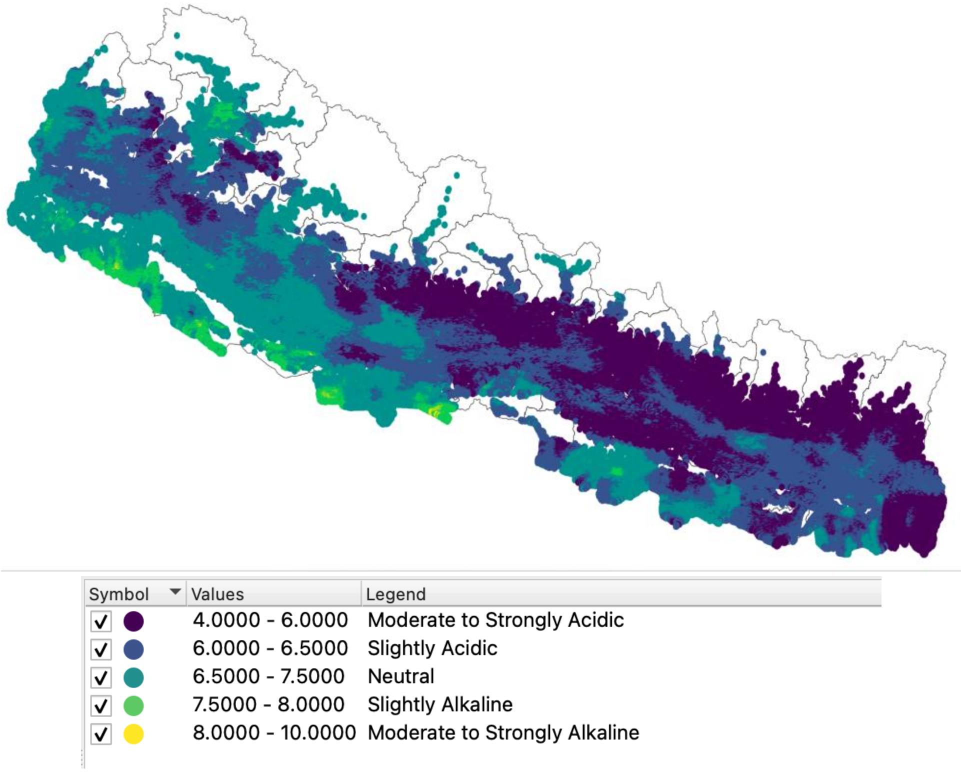

Figure 3. Spatial Distribution Map of pH in Nepal

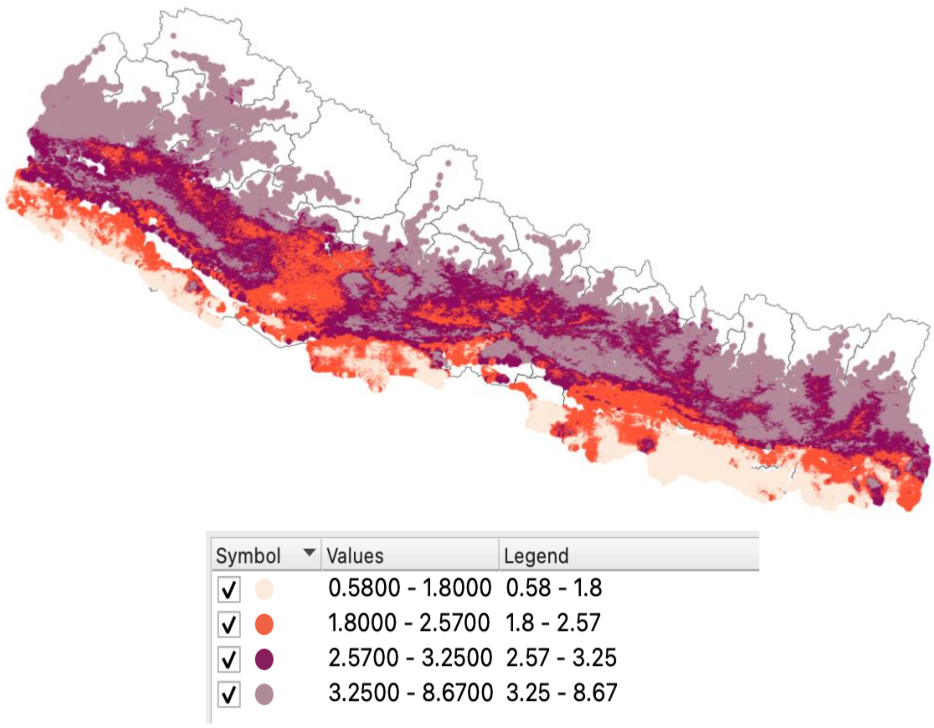

Figure 4. Spatial Distribution Map of organic matter in Nepal

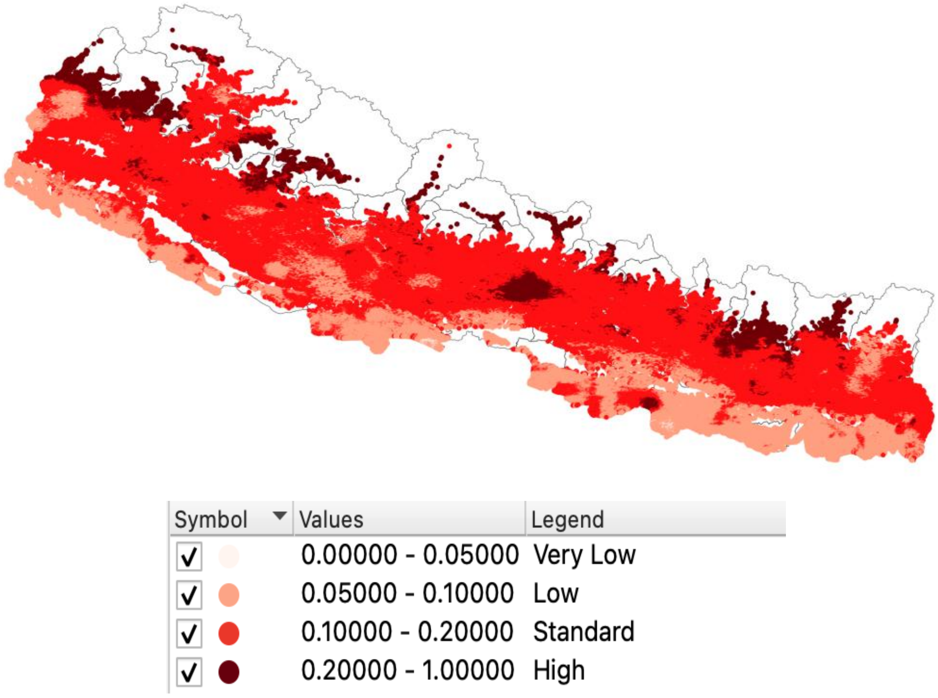

Figure 5. Spatial Distribution Map of nitrogen in Nepal

3.2.3 Provincial-Level Soil Variation

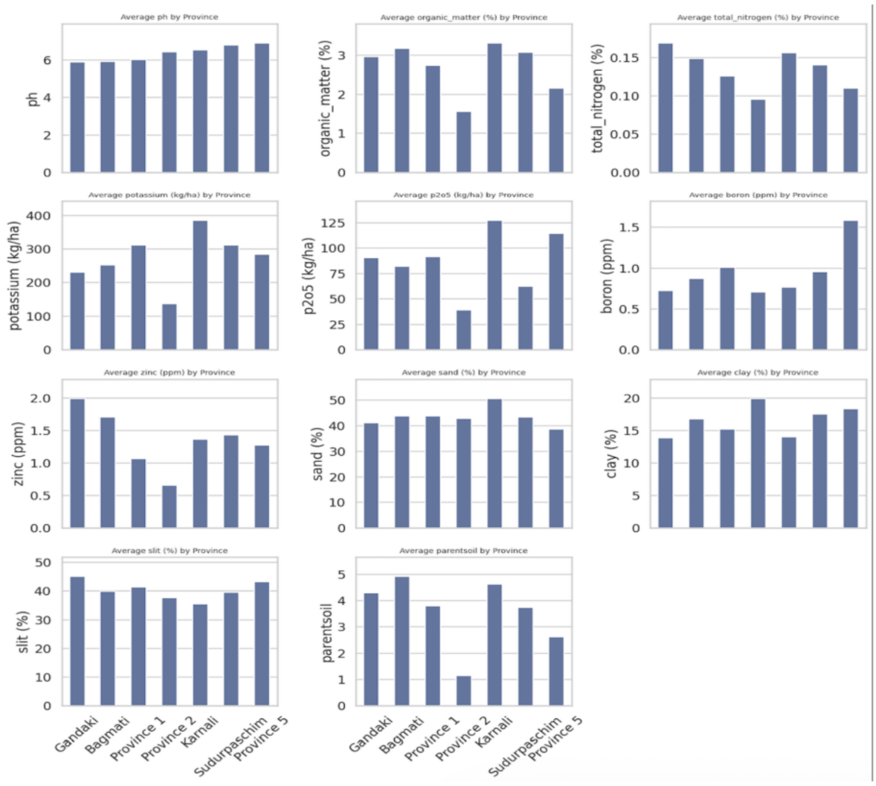

Figure 6. Relationship between 2 provinces and soil properties across the Nepalese dataset.

Province-wise summaries, as shown in Figure 6, showed clear spatial patterns:

- **Gandaki**: Bears the greatest overall nitrogen concentration (0.17%) and the best overall zinc concentration (2.0/ppm). Nevertheless, it is located on the lower soil acidity grounds, with a pH of about 5.9 and less clay content in the soil (14%)
- **Bagmati**: Characterized by a high organic matter (3.25 - 4.5%) and overall nitrogen (0.15 -0.2 percent) content, maintains a comparable acid level in the soil as Gandaki.
- **Province 1**: Has the largest amounts of nutrient deficiencies in the data set; the lowest average organic matter (1.6%), potassium (140 kg/ha), and zinc (0.6 ppm).
- **Province 2:** Has the lowest average phosphorus (P2O5) levels (40 kg/ha) and lower levels of boron (0.7 ppm), although it has a moderately alkaline pH of about 6.4.
- **Karnali**: Is also high in its macronutrient content, with the highest average potassium (means = 380 kg/ha) and a strong growth in phosphorus (means = 125 kg/ha), which is likely due to its high-altitude, silt-based profile.
- **Sudurpaschim**: Has the greatest concentration of boron (1.6 ppm) and alkalinity (pH corroboration of soil is about 7.0), as well as per cent sand (44) and clay (11.5) content.
- **Province 5:** Is largely defined by a high content of boron (in the form of an average of 1.6 ppm of boron) and a rather neutral pH of soil (average 6.9), which places it at the top end of the spectrum of chemically balanced provinces in regard to micronutrients.

3.3 Train-Test Split and Cross-Validation

The dataset was randomly split into 80% training and 20% test partitions using a fixed random seed for reproducibility. To mitigate the risk of spatial autocorrelation where geographically adjacent samples may leak information between train and test sets, 10-fold cross-validation was additionally performed on a held-out subset for stability assessment. All preprocessing steps (missing-value handling, unit standardization) were fitted on the training fold only and applied to the validation fold to prevent target leakage.

3.4 Multi-Output Regression Modeling

Five models were trained: Linear Regression, Ridge Regression, Random Forest, XGBoost, and LightGBM. The first two serve as linear baselines; the latter three are tree-based ensembles capable of capturing non-linear soil–nutrient interactions. Each model was wrapped in a MultiOutputRegressor to produce simultaneous predictions for all fertilizer targets. Key hyperparameters are listed in Table 2.

Table 2. Summary of the Hyperparameters of each model trained

Model	**Hyperparameters**
Ridge Regression	α = 1.0
Random Forest	n_estimators = 20, max_depth = 10, random_state = 42, n_jobs = -1 ,
XGBoost	n_estimators = 100, learning_rate = 0.1, random_state = 42, n_jobs = -1
LightGBM	n_estimators = 300, learning_rate = 0.05, num_leaves = 31, random_state = 42, n_jobs = -1

3.5 Evaluation Metrics

Three complementary regression metrics were used to evaluate model performance: the coefficient of determination (R^2), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE). Each metric captures a different aspect of predictive quality, and reporting all three together provides a more complete picture than any single measure could offer. R^2 explains how much variance in the target relative to a simple baseline (mean prediction), so mostly R^2 is used for crop-wise, fertilizer-specific performance evaluation of the model.

4. Result Analysis and Discussion

4.1 Overall Model Performance

Figure 7 compares global model performance across all targets. Tree-based ensemble methods substantially outperformed linear baselines: LightGBM achieved $R^2 = 0.976$ (RMSE = 0.010), XGBoost $R^2 = 0.975$ (RMSE = 0.010), and Random Forest $R^2 = 0.972$ (RMSE = 0.011), while Linear and Ridge Regression plateaued at $R^2 = 0.755$ (RMSE = 0.121). The large gap (≈0.22 R^2) confirms that the relationships between soil properties and fertilizer requirements are strongly non-linear and cannot be captured by linear models alone.

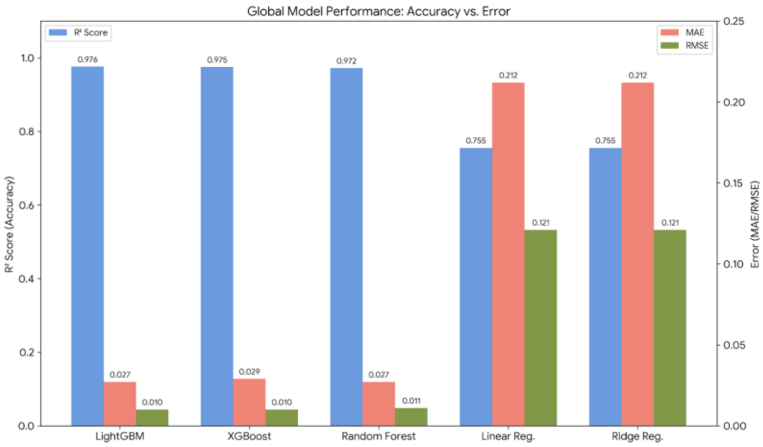

Figure 7. Global Model Performance of about 384,000 datapoints

4.2 Crop-Specific and Fertilizer-Specific Accuracy

Crop-wise performance, as shown in Figure 8, was consistently high for ensemble methods, with R^2 values of 0.983 for maize (Random Forest), 0.969 for rice, and 0.980 for wheat (LightGBM). Fertilizer-wise results in Figure 9 show that organic fertilizer and boron were predicted perfectly ($R^2 = 1.0$) across all ensemble models, indicating clear-cut threshold logic in the underlying recommendation rules. Macronutrient predictions were also strong: DAP reached $R^2 = 0.992$ and MoP $R^2 = 0.983$. Urea (split doses) showed slightly lower but still robust accuracy ($R^2 = 0.936$–0.958), consistent with the higher variability of nitrogen recommendations across ecological zones - a regime where linear baselines collapsed to $R^2 \approx 0.58$.

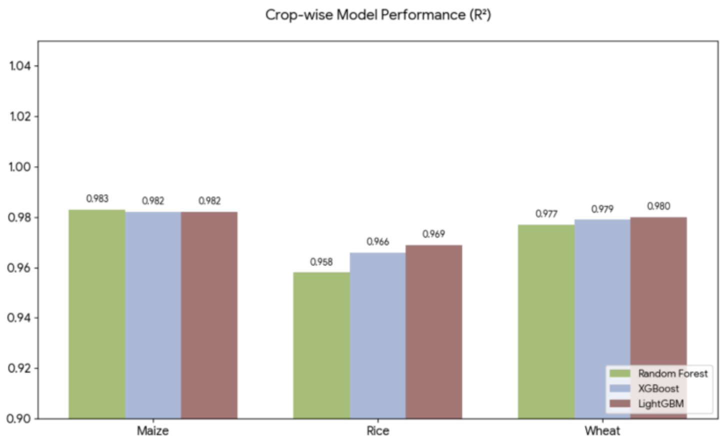

Figure 8. Crop Wise Model's Performance for about 384,000 datapoints

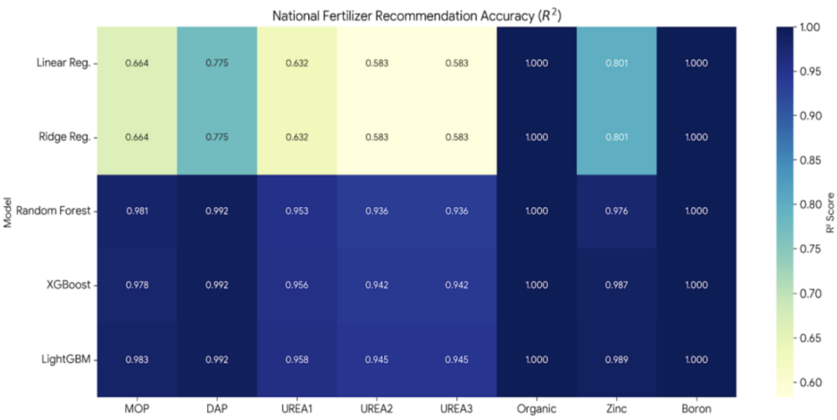

Figure 9. Fertilizer Wise Model's Accuracy for about 384,000 datapoints

4.3 Cross-Validation Stability

A 10-fold cross-validation experiment on an earlier two-province subset of approximately 100,000 samples Figure 10 confirmed model stability. Random Forest achieved the highest mean R^2 (0.988), with XGBoost and LightGBM closely following at 0.987 each. Linear and Ridge Regression averaged $R^2 = 0.894$. The narrow spread across folds and the close agreement between training and validation R^2 indicate that the ensemble models generalize well rather than overfit.

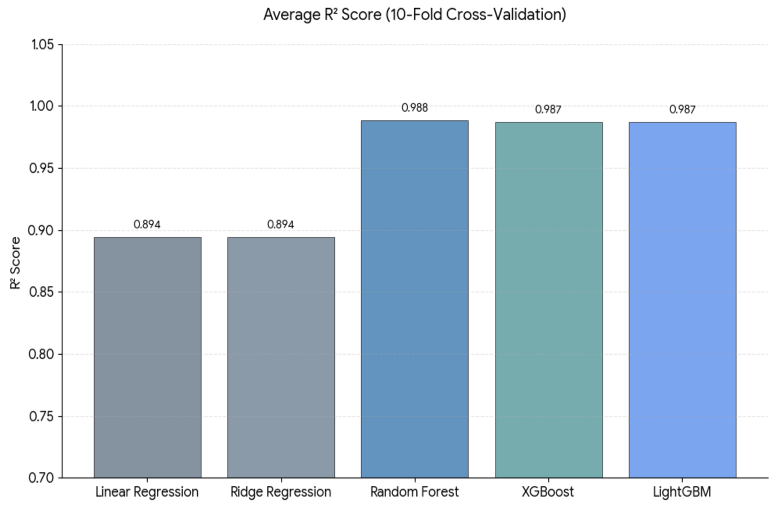

Figure 10. 10-Fold Cross Validation Averages R^2 for about 100,000 datapoints of only two provinces

4.4 Implications for Precision Agriculture

The findings have direct implications for improving fertilizer management in Nepal:

- Current blanket recommendations are not aligned with field realities. The dataset shows large nutrient variability across short distances, meaning fixed national doses are unsuitable for most farmers.
- Geospatial modelling is essential. Longitude, latitude, and elevation provide a more meaningful representation of nutrient needs than administrative boundaries.
- Farm-level decision-support becomes feasible. With the proposed ML framework, farmers could receive fertilizer recommendations tailored to their specific soil properties and location.
- Environmental and economic benefits: More precise application reduces nutrient wastage, lowers input cost, and improves soil sustainability.

These implications support transitioning from conventional fertilizer advisory systems toward data-driven, location-specific nutrient management.

4.5 Comparison with Existing Systems

The performance reported here exceeds the R^2 of 0.81 obtained by Kucuk [14] for multi-output soil moisture prediction and the 90 - 93% accuracy range cited by Latha and Kumaresan [11] for soil nutrient prediction. Unlike GeaGrow [7], which depends on physical sensor deployments, our model operates purely on geospatial coordinates and soil chemistry already available through national soil databases, making it deployable at scale without additional hardware. The explicit inclusion of elevation as a continuous predictor, absent from most global precision-agriculture models [12], aligns the framework with the topographic reality of Nepal's terrain.

4.6 Practical Deployment

The trained model can be deployed as a lightweight web or mobile service in which a farmer or agricultural extension officer enters GPS coordinates (or selects a location on a map) and receives crop-specific fertilizer quantities. NARC and the Department of Agriculture could integrate this service into existing advisory channels, including SMS-based delivery for areas with limited internet access. Because the model relies only on soil and geospatial inputs already collected by NSSRC, no additional field instrumentation is required, lowering the barrier to nationwide rollout.

5. Limitations

First, the dataset was retrieved from the NARC Soil Fertility API in 2022, and soil properties are not static; they vary with seasonal weather, irrigation, cropping cycles, organic input, and long-term land use. As established in the agronomy literature, fresh soil testing remains essential before any fertilizer application. The proposed model, therefore, learns a mapping $f(X) \rightarrow Y$ from soil and geospatial inputs to recommended fertilizer quantities, but it does not eliminate the need for current soil analysis to obtain an accurate input vector X reflecting present conditions. The model is best understood as a decision-support layer that operates on top of up-to-date soil measurements, not a replacement for them.

Second, hyperparameters across all five models were not exhaustively tuned due to computational constraints associated with the size of the dataset (≈384,000 records) and the multi-output nature of the prediction task. The Random Forest model in particular was configured with only 20 estimators and a maximum depth of 10 to keep training time tractable, which is considerably smaller than typical configurations reported in the literature. Most prior studies on soil-property prediction report Random Forest as the strongest performer; the slightly lower R^2 of Random Forest in this work (0.972) compared to LightGBM (0.976) and XGBoost (0.975) is therefore likely attributable to under-parameterization rather than a fundamental limitation of the algorithm. With more compute, formal tuning (e.g., grid search or Bayesian optimization) would likely improve all three ensemble models further, and Random Forest specifically may surpass the boosting methods.

Third, the targets predicted by the model are recommended fertilizer quantities returned by the NSSRC API rather than field-measured optimal doses validated against actual yields. The API itself derives its recommendations from raster-based soil maps with finite spatial resolution, which may introduce sampling bias, for example, over-representation of regions with denser raster coverage and under-representation of high-altitude or remote zones where soil samples are sparse. Consequently, the model partly reflects the recommendation logic encoded in the NSSRC system rather than ground-truth agronomic optima.

Fourth, a full 10-fold cross-validation was performed on a two-province subset of approximately 100,000 samples rather than the full nationwide dataset, again due to computational limitations. Extending cross-validation across all seven provinces remains a meaningful exercise for future stability assessment.

6. Conclusion and Future Work

This study presented a multi-output machine learning framework for site-specific fertilizer recommendation in Nepal, integrating soil physicochemical properties with geospatial features at a national scale. Using a custom data extraction pipeline, a dataset of 384,077 complete soil records covering all seven provinces was assembled from the NSSRC Soil Fertility API. Comparative evaluation across five models established that tree-based ensembles, LightGBM (R^2 = 0.976), XGBoost (R^2 = 0.975), and Random Forest (R^2 = 0.972), substantially outperform linear baselines (R^2 = 0.755), with 10-fold cross-validation confirming strong generalization. By treating elevation as a continuous predictor and predicting macronutrient and micronutrient quantities jointly, the framework moves beyond the static, region-based fertilizer guidelines currently in use.

Future work will proceed along three directions. First, verification of API-derived data against fresh soil samples collected in 2026 from the same coordinates will quantify the temporal drift between the 2022 source data and present-day soil conditions, and will indicate how frequently the model needs retraining. Second, an ablation study isolating the contribution of geospatial features (latitude, longitude, elevation) will be conducted by training and evaluating two parallel pipelines, one including geospatial inputs and one excluding them to empirically validate the central claim that topography improves fertilizer prediction. Third, systematic hyperparameter optimization of all ensemble models, particularly Random Forest with a larger number of estimators and unconstrained depth, will be carried out once additional computing is available, along with full nationwide cross-validation.

Acknowledgements

The authors gratefully acknowledge the support of the Department of Electronics and Computer Engineering, Pulchowk Campus, in conducting this research. Also, gratitude is extended towards the Nepal Agricultural Research Council (NARC), whose insights and datasets informed this proposal.

References

[1] DP Sherchan and KB Karki, Plant nutrient management for improving crop productivity in Nepal.

[2] Nishant K Sinha, Immanuel Chongboi Haokip, Jitendra Kumar, RH Wanjari, Shilpi Verma, Monoranjan Mohanty, Somasundaram Jayaraman, R Elanchezhian, and Rahul Mishra. Dhiraj Kumar, "Impact of fertilizer consumption on soil health and environmental quality in India," Indian Journal of Fertilizers, vol. 18, no. 10, pp. 992-1005, 2022.

[3] Vishwas Regmi, Babita Dhungana, Suraj Gosai, Swostika Dhungana, Aakash Adhikari, Shalin Subedi Sanket Kattel, "Analytical study on soil fertility status along with elevation gradient and depth of Chitwan–Mustang track, Nepal," Plant Physiology and Soil, vol. 2, no. 2, 2022.

[4] Seri Mastura Mustaza, Nasharuddin Zainal, Mohd Hairi Mohd Zaman, Asraf Mohamed Moubark Nurul Ayni Mat Pauzi, "NurulArtificial intelligence in precision agriculture: A review. Jurnal Kejuruteraan, 37(2):1025–1047, 2025," Journal Kejuruteraan, vol. 37, no. 2, pp. 1025-1047, 2025.

[5] STC Amarasinghe, AAYD Amarasinghe, AAY Amarasinghe KPGDM Polwaththa, "Exploring artificial intelligence and machine learning in precision agriculture: A pathway to improve efficiency and economic outcomes in crop production.," American Journal of Agricultural Science, Engineering and Technology, vol. 8, no. 3, pp. 50-59, 2024.

[6] Leonardus Vergütz, Achraf El Allali Oumnia Ennaji, "Machine learning in nutrient management: A review. Artificial Intelligence in Agriculture, 9:1–11, 2023.," Artificial Intelligence in Agriculture, vol. 9, pp. 1-11, 2023.

[7] Syed Muhammad Zaigham Abbas Naqvi, Hao Zhang, Linze Li, Wei Zhang, Fuad A Awwad, Emad AA Ismail, M Ijaz Khan, Vijaya Raghavan, Jiandong Hu Muhammad Awais, "Ai and machine learning for soil analysis: an assessment of sustainable agricultural practices," Bioresources and Bioprocessing, vol. 10, no. 1, p. 90, 2023.

[8] Oluwafolake Ojo, Mutiu Busari, Muftau Adebayo, Joshua Adejumobi, Daniel Folorunso, Femi Ayo,

Orobosade Alabi, Olusola Olabanjo Olusegun Folorunso, "Geagrow: a mobile tool for soil nutrient prediction and fertilizer optimization using artificial neural networks," Frontiers in Sustinable Food Systems, vol. 9, 2025.

[9] Anthony Vodacek, Damien Hanyurwimfura, Alfred Uwitonze, Innocent Kabandana Christine Musanase, "Data-driven analysis and machine learning-based crop and fertilizer recommendation system for revolutionizing farming practices," Agriculture, vol. 13, no. 11, 2023.

[10] P Latha and P Kumaresan, "Deep learning for soil nutrient prediction and strategic crop recommendations: An analytic perspective," Nature Environment and Pollution Technology, vol. 24, no. 1, pp. 1-12, 2025.

[11] Lucas K Johnson, Julia Silge, Hannah Frick, Max Kuhn, Colin M Beier Michael J Mahoney, "Assessing the performance of spatial cross-validation approaches for models of spatially structured data," arXiv preprint, 2023.

[12] Suresh Kumar, Bappa Das, Trisha Roy Justin George Kalambukattu, "Soil depth mapping in the hilly and mountainous landscape of the north-west himalayan region using machine learning techniques," Earth Science Informatics, vol. 18, no. 3, p. 492, 2023.

[13] Manohar Yadav, Justin George Kalambukattu, Suresh Kumar Nyenshu Seb Rengma, "Machine learning based digital mapping of soil properties in mid-himalayan terrain," 2023.

[14] Derya Birant, Pelin Yildirim Taser Cansel Kucuk, "Cansel Kucuk, Derya Birant, and Pelin Yildirim Taser. An intelligent multi-output regression model for soil moisture prediction: In the International Conference on Intelligent and Fuzzy Systems, ," Springer, pp. 474-481, 2021.