Yield Estimation of Potato Crop Using UAV Imagery and Machine Learning Algorithms

Asmita Banepali, Aman Manandhar, Ishwor Dhital, Sudipta Poudel and Pratima Thapa (Nepal)

Key words: Photogrammetry; Remote sensing; Yield Estimation; Multispectral UAV; Vegetation Indices (VI); Maching Learning

SUMMARY

Agricultural yield estimation plays a pivotal role in enhancing food security optimizing resource allocation, and supporting decision making in precision agriculture. In Nepal, where smallholder farming dominates and traditional yield estimation methods are often manual and time consuming, there is a pressing need for scalable, accurate, and efficient solutions. This study presents a data driven framework for estimating potato yield by integrating Unmanned Aerial Vehicle (UAV) remote sensing with machine learning techniques.

The research was conducted on a 3.7 hectare potato field in Dhulikhel Municipality, Kavrepalanchok District, Nepal. High-resolution imagery was acquired across five key phenological stages using RGB and multispectral sensors mounted on UAVs. To ensure spatial precision, nine Ground Control Points (GCPs) were established using Differential GPS, and 80 georeferenced sample plots were used for ground truthing. From the UAV imagery, 20 vegetation indices (VIs), including NDVI, EVI, CIrededge, and VARI, were computed and statistically analyzed for their correlation with actual yield data.

Both simple and multiple linear regression models were applied, alongside seven machine learning algorithms: Support Vector Machine (SVM), Random Forest, XGBoost, AdaBoost, Decision Tree, Gradient Boosting, and K-Nearest Neighbors (KNN). The multi-temporal multiple linear regression model performed best, achieving an R^2 of 82.61% and predicted R^2 of 77.54%. Among machine learning methods, Gradient Boosting delivered the highest accuracy (R^2 = 76.62%, RMSE = 0.1526). Multispectral based indices, especially CIrededge, showed stronger correlation with yield than RGB derived indices such as VARI, highlighting the advantages of multispectral sensors in agricultural

monitoring.

The study also revealed that multispectral-based vegetation indices had greater predictive power than those derived from RGB imagery, underlining the advantage of using multispectral UAV sensors in precision agriculture. For example, the Chlorophyll Index Red Edge (CI1) showed significantly higher correlation with yield than the RGB based VARI index. These results confirm the relevance of spectral data in capturing the physiological characteristics of crops that are closely related to productivity. Throughout the research, a comprehensive set of data was generated, including high-resolution orthomosaics for each growth stage, digital surface models (DSMs), digital terrain models (DTMs), vegetation index maps, plant height models, and yield sample polygons linked with ground-truth yield measurements. These datasets enabled rigorous statistical analysis, model training, and yield mapping, forming a valuable resource for future research and decision-making in precision agriculture

Overall, the integration of UAV-acquired spectral data with advanced analytics demonstrates a reliable, non-destructive, and spatially explicit method for yield estimation. This approach offers significant potential to modernize traditional agricultural practices in Nepal, providing smallholder farmers with timely, data-driven insights to support decision-making and improve productivity.