A Precision Soil Health Assessment Framework Using LIBS and Machine Learning to Support Sustainable Nutrient Management and Crop Protection

Yingchao Huang
University of Alberta, Canada

Shubho Mohajan
University of Alberta, Canada

Amina E. Hussein
University of Alberta, Canada

Miles Dyck
University of Alberta, Canada

Abstract

Accurate and timely assessment of soil nitrogen and carbon content is fundamental to soil health evaluation, sustainable nutrient management, and integrated pest management (IPM)–based crop protection strategies. Reliable quantification of these key soil properties enables optimized fertilizer application, enhances crop resilience, and reduces excessive chemical inputs that can increase pest pressure and environmental impact. This study presents a machine-learning framework for precise estimation of soil nitrogen and carbon content using Laser-Induced Breakdown Spectroscopy (LIBS), designed to support sustainable agricultural decision-making under realistic field conditions and limited sample availability.

A robust feature selection strategy was developed to identify informative elemental emission lines based on Partial Least Squares (PLS) coefficients, improving the quality and relevance of spectral features. PLS was subsequently applied to extract key components from the selected emission lines, which served as inputs for model training, validation, and testing. The proposed framework integrates two complementary learning pathways: (1) a Multi-Layer Perceptron (MLP) model for direct prediction and (2) a similarity-based learning mechanism that employs cosine similarity to weight query spectra relative to training spectra. Final estimates are obtained through a weighted fusion of outputs from both pathways, enhancing prediction robustness and accuracy.

The framework was evaluated using 43 soil samples collected from agricultural fields in Breton, Alberta, and benchmarked against established methods, including PLS Regression (PLSR), LASSO, and Gaussian Process Regression (GPR). The proposed approach achieved relative errors of 13.04% for nitrogen and 13.45% for carbon, with root mean square errors of 0.072 w/w% and 0.33 w/w%, respectively, outperforming the literature methods. Ablation and sensitivity analyses further demonstrated the importance of emission line selection, feature extraction, model integration, and key hyperparameters.

These results highlight the potential of the proposed LIBS-based machine-learning framework as a practical and reliable tool for soil health monitoring, precision nutrient management, and sustainable crop protection, contributing to data-driven IPM strategies and environmentally responsible agricultural practices.