A MULTIDIMENSIONAL FEATURE SELECTION AND MODEL OPTIMIZATION FRAMEWORK FOR ACCURATE ESTIMATION OF FRACTIONAL VEGETATION COVER FROM FULLY POLARIMETRIC SAR DATA

Author's: K. Manikandan[1], N. Vanjulavalli[2], K. Sujith[3]

Abstract

Fractional Vegetation Cover (FVC) is a critical biophysical parameter widely used in precision agriculture, environmental monitoring, hydrology, ecosystem assessment, and climate change studies. Accurate estimation of FVC enables improved crop yield prediction, soil erosion modeling, drought assessment, and carbon cycle analysis. While optical remote sensing techniques have traditionally been employed for vegetation monitoring, they suffer from cloud cover limitations, atmospheric disturbances, and restricted temporal availability. Fully Polarimetric Synthetic Aperture Radar (PolSAR) systems provide an all-weather, day-and-night imaging capability, making them highly suitable for continuous vegetation monitoring.This proposes a comprehensive framework for estimating Fractional Vegetation Cover from fully polarimetric SAR data using multidimensional feature extraction, feature selection, and model optimization. The methodology integrates polarimetric features such as intensity measures (HH, HV, VV), polarimetric ratios, coherence parameters, entropy, anisotropy, alpha angle proxies, and texture-based descriptors. These multidimensional features are derived from simulated PolSAR covariance representations to emulate realistic scattering behavior of vegetation and soil.To improve prediction accuracy and reduce redundancy, multiple feature selection strategies are applied, including filter-based (Mutual Information), embedded methods (L1-regularized regression), and wrapper-based approaches (Recursive Feature Elimination). A consensus-based feature selection strategy ensures robust dimensionality reduction while preserving physically meaningful polarimetric descriptors.The selected feature subsets are then used for regression-based FVC estimation. Various machine learning models such as Random Forest Regression, Gradient Boosting Regression, Support Vector Regression (SVR), and ElasticNet are optimized using cross-validation and hyperparameter tuning techniques. Model performance is evaluated using statistical metrics including Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and coefficient of determination (R^2).