119-AGR

Machine Learning-Driven Estimation of Winter Cereal Rye Biomass across the Eastern Continental US

Abstract

Cereal rye is a widely adopted winter cover crop in the US due to its winter hardiness and its ability to enhance the sustainability of our cropping systems by reducing nitrate leaching, conserving soil and water, and suppressing weeds. Cereal rye biomass dictates the degree to which these ecosystem services are provisioned in agroecosystems. Therefore, accurate cereal rye biomass predictions is crucial for effective agricultural planning and sustainability efforts, particularly in response to variable climatic conditions. This study addresses a significant gap in the regional-scale prediction of cereal rye biomass by incorporating daily meteorological data and evaluating multiple predictive modeling techniques. The primary objective was to develop a robust predictive model for cereal rye biomass using an enhanced dataset that combines cereal rye biomass observations with high-resolution climatic variables derived from the GridMET dataset. Our methodology involved integrating cereal rye biomass data with daily weather metrics, employing several statistical and machine learning models to assess which best predicts biomass outcomes. We compared the performance of multi-linear regression, decision trees, Random Forest, and Feedforward Neural Networks (FNN). Comparatively, the Random Forest model exhibited superior accuracy, achieving an R^2 value of 0.81 and a Normalized Root Mean Square Error (NRMSE) of 0.08%, which highlighted its effectiveness in capturing the complex relationships between biomass and climatic factors. Further refinements, through Recursive Feature Elimination (RFE) and parameter tuning, significantly enhanced the model's predictive capability. To interpret the influence of individual predictors, Partial Dependence Plots (PDPs) were generated, offering insights into the dependency of biomass on specific climatic variables. Utilizing the refined Random Forest model, we conducted spatial predictions of rye biomass across the Eastern Continental United States (East-CONUS) for the climatological period of 1990-2020. The resulting county-level biomass prediction maps were iterated for four scenarios: early and late planting and termination dates. These maps provide actionable insights that can aid farmers and agricultural planners in improving cover crop biomass predictions under varying climatic conditions hence, optimizing planting/termination schedules. As next steps, the study recommends exploring the implications of the model, such as, evaluating the role of winter cereal rye to optimize soil health metrics, weed control and soil moisture conservation. This research contributes to the precision agriculture tools necessary to adapt farming practices to climate variability, thereby enhancing sustainability and resource management.

Keywords: Winter Rye Biomass, Machine Learning, Random Forest, Feature Elimination, Biomass Prediction maps, Soil/Water conservation.