Forage Biomass Prediction using UAV-Driven Multi-Spectrum and Machine Learning

K.J. Han[1], M.W. Alison[2], X. Meng[3], B. Awasthi[3], K.C. Manisha[3], and V. Sehgal[1]
[1]School of Plant, Environmental, and Soil Sciences, LSU AgCenter, Baton Rouge, LA
[2]Macon Ridge Research Station, LSU AgCenter, Winnsboro, LA
[3]Department of Geography and Anthropology, LSU, Baton Rouge, LA

Introduction

Estimating pasture yield is crucial for improving management efficiency. Traditionally, this involves labor-intensive and time-consuming methods like multiple spot clippings, drying biomass samples, and weighing. Although simplified methods such as rising plate meters or grazing sticks are available, they often have significant error potential due to their limited representativeness of the entire pasture. A viable alternative could be using a prediction model developed with multispectral data from a UAV-mounted camera. The Random Forest (RF) algorithm, one of the most popular machine learning methods for regression tasks, is known for its robustness, ability to handle large datasets, and high accuracy.

Materials and Methods

Annual Ryegrass and Oats Yield Data

Yield data collected at the LSU AgCenter Southeast and Macon Ridge Research Stations includes 595 data points for annual ryegrass (Lolium multiflorum) and 739 composite data points for ryegrass and oats, spanning from January to May 2003.

Independent Variables: Red, Red Edge, NDVI (Normal Difference Vegetation Index), NDRE (Normalized Difference Red Edge), CIRE (Chlorophyll Index Red Edge), SR (Surface Reflectance)

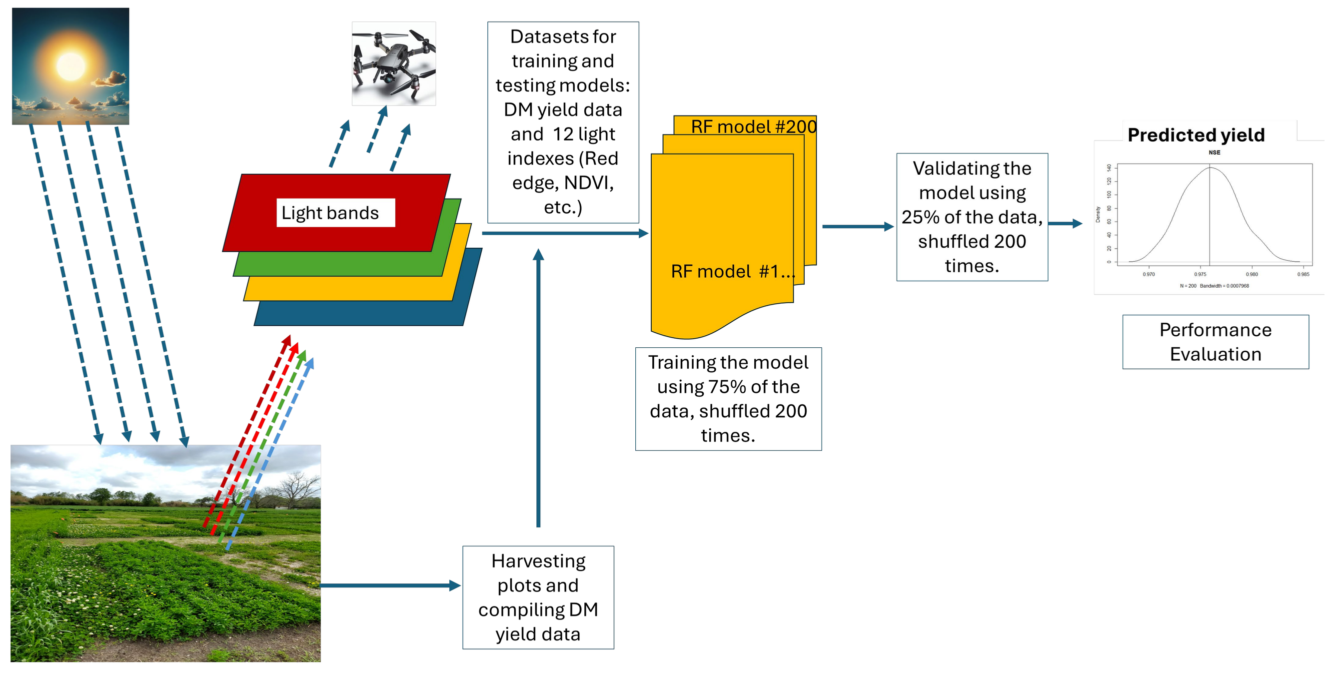

Fig. 1. Process for model development and validation

Model Development: Data Split for Training and Validation: 3:1 ratio and model development (Fig. 1)

Model Evaluation: Cross-validation with RMSE (Root Mean Square Error), MAE (Mean Absolute Error), & Correlation between prediction vs observation.

Results & Discussions

Model Fitting:

The observed yield data showed a moderate skew towards the higher yield side across three harvests. As shown in Fig. 2, among the light indexes, Red Edges and NDVI (Normalized Difference Vegetation Index) were the most significant predictors for yield estimation.

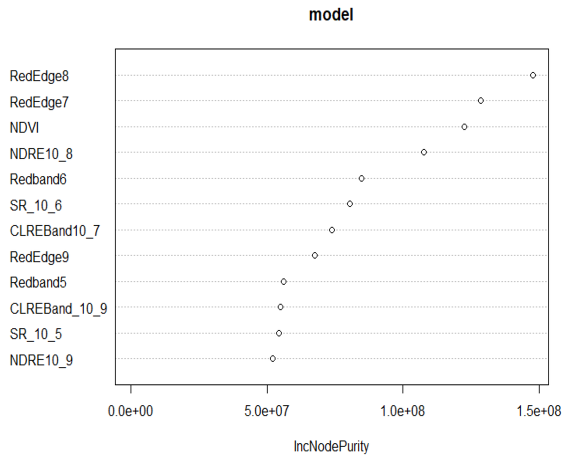

Fig. 2. Importance of light indexes in order.

Assessment of model performance

The scatter plots of DM yields, comparing predicted yields to observations, showed linear alignments in ryegrass-only (Fig. 3a) and composite modeling attempts (Fig. 3b).

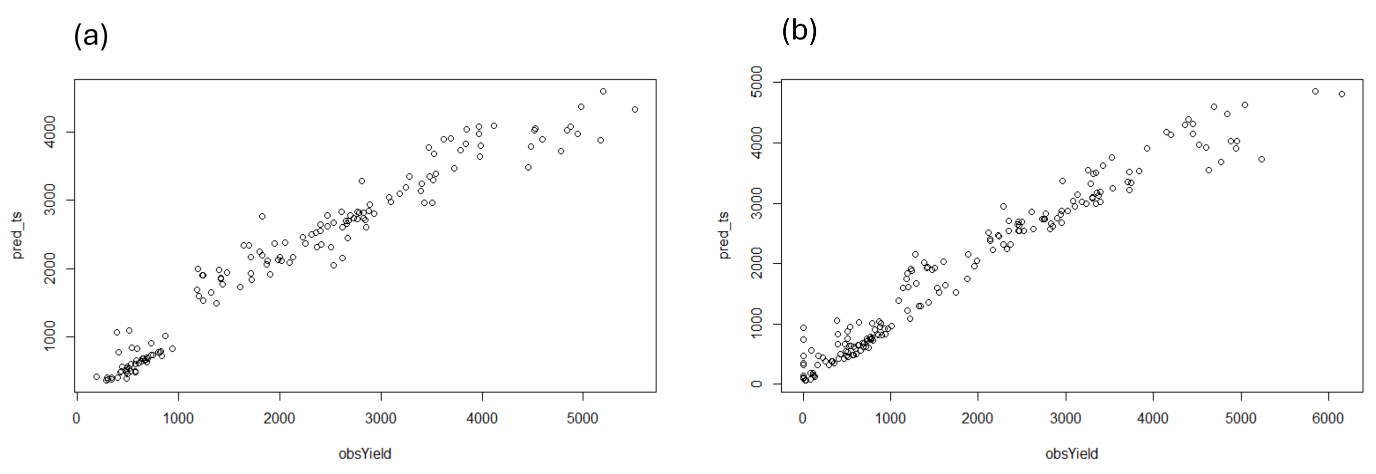

Fig. 3. Scatter plots of predicted vs observed DM yields of ryegrass only (a) and composite (b).

Figure 4 illustrates the mean predicted yield of the test data after 200 ensemble iterations (red dots) for the ryegrass-only model (a) and the composite model (b).

In summary, the average RMSE was 363 lbs per acre for the ryegrass-only model and 411 lbs per acre for the composite model (Table 1). Similarly, the composite model's MAE was higher than that of the ryegrass-only model. The average correlations between prediction and observation were similar for both models.

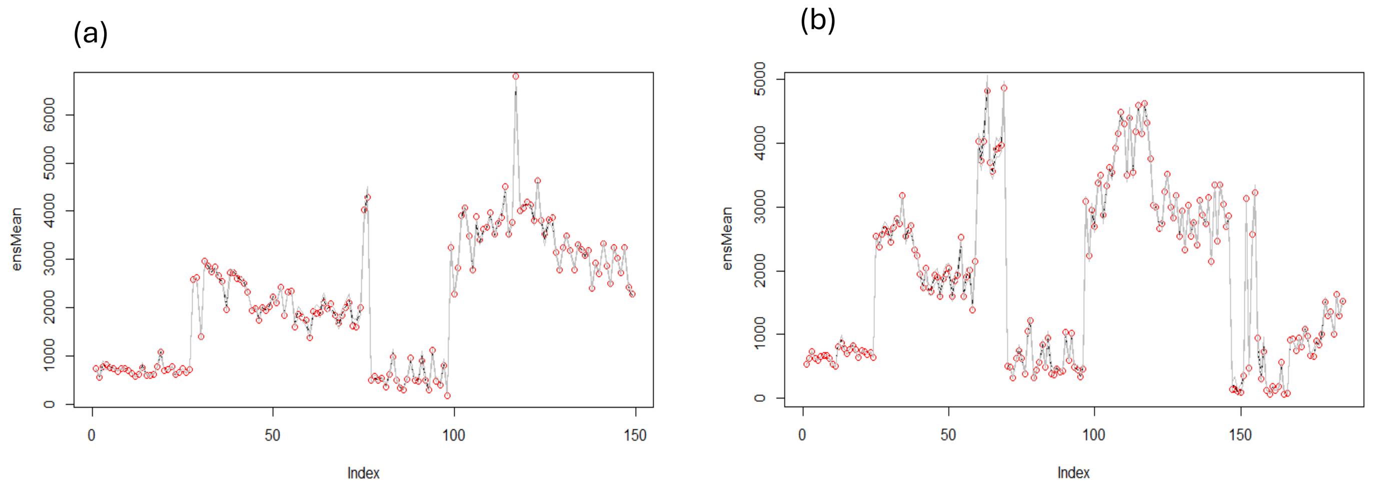

Fig. 4. Yield predictions of 200 ensemble of the ryegrass only (a) and composite model (b).

Table 1. Summary of performance metrics of ryegrass only and composite models

Performance metrics	Ryegrass only	Composite
RMSE, lbs/acre	363	411
MAE, lbs/acre	241	262
Correlation	0.975	0.970

Implications

These yield prediction models seem suitable for estimating forage biomass production of overseeded annual ryegrass and potentially small-grain forage. By incorporating additional datasets, such as multispectral and yield data from warm-season grass pastures, the models can be expanded to cover a wider range of forage types. This approach is expected to enhance accuracy and significantly reduce the labor and time required compared to traditional methods, making it a promising tool for agricultural stakeholders.