Development of an Open-Source Web-Based Tool for Alfalfa Yield and Quality Prediction

Zhou Zhang, Matthew Digman, University of Wisconsin-Madison

RATIONALE & OBJECTIVES

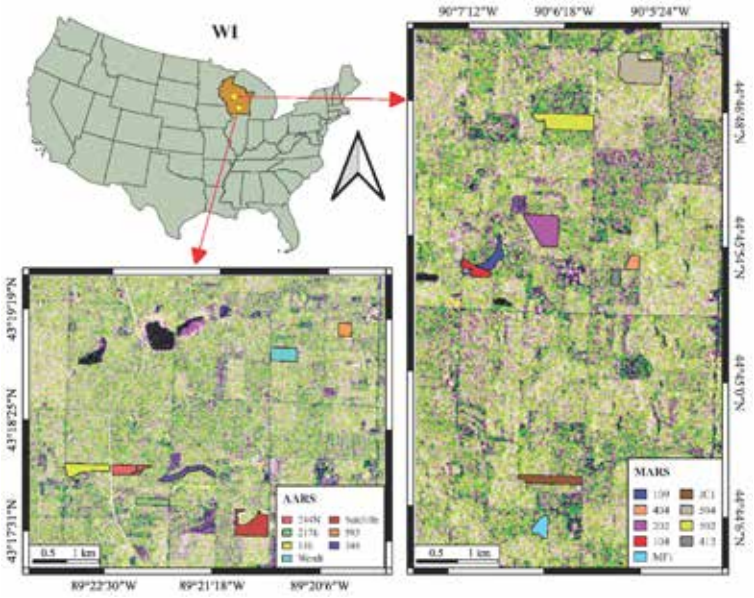
- Develop machine learning models to estimate in-season alfalfa yield and quality using multi-source data and validate the results on Wisconsin Alfalfa Yield and Persistence database.
- Create an open-source web-based tool to provide end-users with in-season alfalfa yield and quality maps to assess their spatial variability for precision management.

STUDY DESCRIPTION

Locations:

Arlington Agricultural Research Station (7 sites), Marshfield Agricultural Research Station (9 sites).

Figure 1. The distribution of sampling sites.



Data:

The features extracted from Sentinel-1 Ground Range Detected (GRD) products, environmental factors from Daymet V4 product and cutting order.

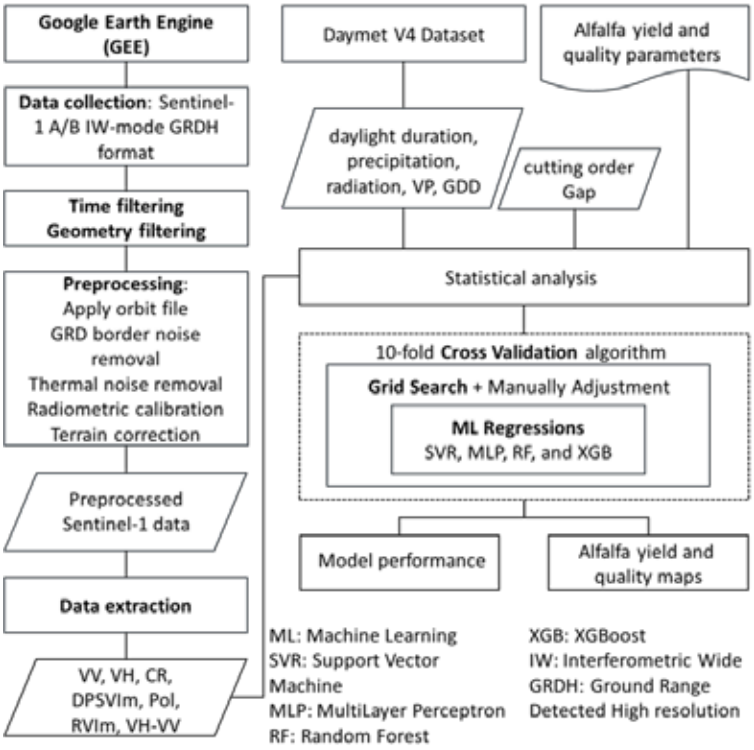
Machine Learning Models:

Support Vector Regression (SVR), Multilayer Perceptron (MLP), Random Forest (RF), and Extreme Gradient Boosting (XGB).

Yield and Quality:

Dry matter yield (DMY), crude protein (CP), neutral detergent fiber (NDF), NDF digestibility (NDFD), and acid detergent fiber (ADF).

Figure 2. Workflow of Objectives 1.



Development environment:

HTML+ JavaScript+ Python (Flask)+ PythonAnywhere service.

RESULTS

Objectives 1:

- Estimated and measured alfalfa yield and quality traits (Figure 3).
- The XGB model consistently performed the best for all alfalfa yield and traits.
- The DMY estimates highly agreed with the measured values with an average R^2 of 0.64 and RMSE of 0.71 tons/ha. The best result for estimating CP was with an average R^2 of 0.72. The performance in estimating alfalfa fiber indicators was achieved with the highest average R^2 of 0.54, 0.64, and 0.60, respectively.

Figure 3. Scatterplots of the estimated and measured alfalfa yield and quality traits, where a - e are for DMY, CP, ADF, NDF, and NDFD, respectively. Estimated values in the plots for individual samples and evaluation metrics were averaged over the 20 times cross validation process.

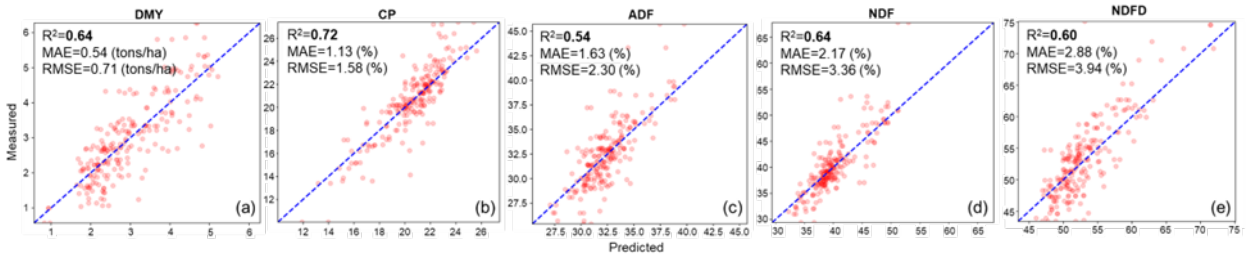
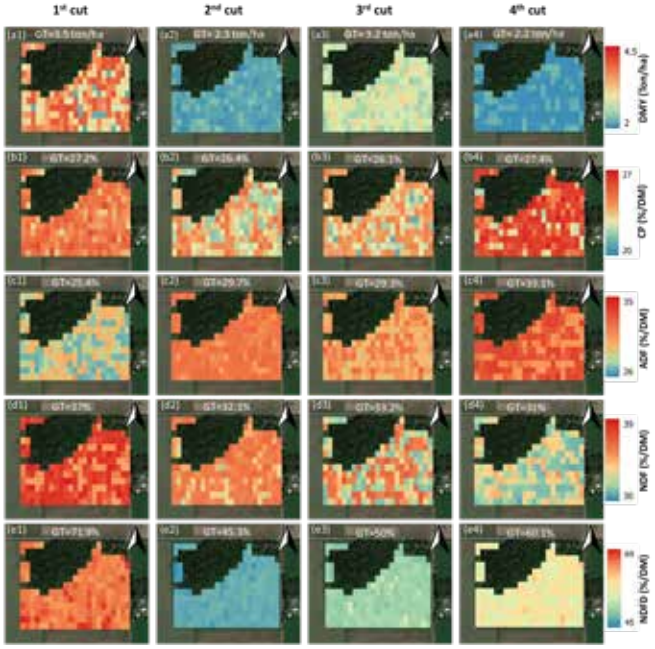


Table 1. R^2 of the four machine learning models in estimating alfalfa yield and quality traits.

Input	Traits	SVR	MLP	RF	XGB
With Environmental factors	DMY	0.53	0.53	0.59	**0.64**
	CP	0.23	0.24	0.59	**0.72**
	ADF	0.27	0.26	0.35	**0.54**
	NDF	0.23	0.24	0.49	**0.64**
	NDFD	0.23	0.27	0.40	**0.60**
Without Environmental factors	DMY	0.40	**0.47**	**0.47**	0.45
	CP	**0.19**	0.18	0.13	0.11
	ADF	**0.15**	**0.15**	0.07	0.07
	NDF	0.01	**0.04**	-0.00	-0.10
	NDFD	0.12	0.12	**0.18**	0.14

- Performance for the fiber indicators and CP was not ideal only with the SAR features with the best R^2 all <0.2.
- Performance was significantly higher when environmental factors were added, especially for the quality traits.

Figure 4. Visualization of estimated alfalfa DMY and quality traits of Sutcliffe in 2020.



- The average values of each map corresponded well with their field-level ground data.
- The estimated maps also demonstrated the spatial variations of each field.

Objectives 2:

- The webtool: https://alfalfabse.pythonanywhere.com.
- The webtool can present the saved yield maps of the two study locations in the database.

Figure 5. Some saved yield map of fields in location A.

- The users can also define an area and the date that they wish to show the yield for.

Figure 6. Yield map of a newly defined area.

CONCLUSIONS/SUGGESTIONS

- The potential of using Sentinel-1 radar features and environmental factors in estimating alfalfa yield and quality was proved in this study.
- Field maps that were generated for alfalfa DMY and quality using the model estimates depicted in-field variations, providing a valuable reference for decision-making of field management practices and alfalfa harvest time.

Alfalfa
U.S. Alfalfa Farmer Research Initiative

Funding for this project was provided by the U.S. Alfalfa Farmer Research Initiative of the National Alfalfa & Forage Alliance.