Report for FY24-MG-002: Improving the stability of forecasting models for Fusarium head blight

Note: Report must cover the *entire* duration of the project period of performance, May 1, 2022-July 30, 2026.

1. **What are the major goals and objectives of the research project?**
   The overall goal of this project is to create better models for predicting Fusarium head blight (FHB). The objectives were to:
   1) Continue expanding the FHB data matrix with observations from the IM-CP (**Paul Lab**), but also estimate how many observations are needed for model stability (**DeWolf Lab**);
   2) Search for a stable FHB predictive model or ensemble of models that generalizes well to any environment (in the future or in other states) (**DeWolf Lab**);
   3) Formally assess the stability of variable selection in logistic regression (LR) and random forest (RF) models proposed for FHB prediction (**DeWolf Lab**);
   4) Assess the predictive stability of LR and RF models developed to date (**DeWolf Lab**).

2. **What was accomplished under these goals or objectives?** *(For each major goal/objective, address these three items below.)*

   What were the major activities?
   For **Obj 1**, over the last four years, several new observations were collected from trials conducted by PIs in several U.S. states. Integrated management trials were conducted as part of the IM_CP in AL, DE, IL, IN, KS, KY, LA, NC, MD, MI, MO, MN, ND, NE, OH, PA, SD, TN, VA, and WI. In each trial, at least two commercial wheat cultivars, classified as susceptible, moderately susceptible, or moderately resistant, were planted. FHB index, incidence, and DON data were collected from untreated, naturally infected, non-irrigated plots of each cultivar and edited for inclusion in the master data file for FHB risk model development and validation.

   For details on Obj 2, 3, and 4, see Dr. DeWolf's report.
   The modeling team met several times over the last four years to discuss modeling approaches and share preliminary results. Over the last year, we (Led by Dr. Shah in the DeWolf lab) used bootstrap simulations of different sample sizes to evaluate the impact of sample size on models developed using different approaches (see Dr. DeWolf's Performance Report for details).

   What were the significant results?
   Obj 1. Over the last four years, we collected 882 new data points representing unique combinations of cultivar x block x trial from the IM_CP and added them to the master dataset. Each new observation included the mean FHB index and corresponding DON data from a unique environment, including cultivar resistance class, trial location, wheat market class, cropping and tillage history, and local weather conditions. These data are being mined specifically from trial-resistance-class cases that were neither treated with fungicides nor inoculated with *Fusarium graminearum* spore suspensions.

   For details on Obj 2, 3, and 4, see Dr. DeWolf's report.

List key outcomes or other achievements.

Obj 1. We collected new IND data and associated weather data from a range of wheat-growing locations, likely representing unique FHB-weather patterns not yet in our master dataset.

For details on Obj 2, 3, and 4, see Dr. DeWolf's report.

Preliminary stability analyses highlighted the importance of increasing the sample size used for model development to improve model accuracy and stability.

3. **What opportunities for training and professional development has the project provided?**
   The project provided training to a postdoctoral researcher, partially funded by the project, in data management and data mining. The postdoc gained hands-on experience in extracting and curating FHB index data, as well as cleaning and organizing disease datasets.

4. **How have the results been disseminated to communities of interest?**
   A manuscript was published early in the funding cycle, and abstracts/talks/posters were presented at USWBSI Forum and a gathering of plant disease epidemiologists.

5. **What do you plan to do during the next reporting period to accomplish the goals and objectives?**
   Not applicable, this is the final report for the agreement.