**Weather-based logistic regression models for predicting the risk of wheat blast epidemics**

Monalisa C. De Cól1, Mauricio A. O. Coelho2, Emerson M. Del Ponte1

1 Departamento de Fitopatologia, Universidade Federal de Viçosa, Viçosa MG Brasil.   
2 Empresa de Pesquisa Agropecuária de Minas Gerais (EPAMIG), Patos de Minas MG Brasil.

E-mail: [delponte@ufv.br](mailto:delponte@ufv.br)

The Brazilian Cerrado biome is currently deemed the agricultural frontier for the expansion of wheat cultivation. However, the endemic presence of a fungal disease, wheat blast, caused by the fungus *Pyricularia oryzae* *Triticum* lineage, constrains the expansion of wheat production in this region. This study was designed to develop empirical models for predicting wheat blast epidemics. Disease data were obtained mainly from EPAMIG field experiments in Patos de Minas, MG, Brazil, conducted from 2013 to 2019, featuring sequential plantings within a year. We also utilized an additional publicly accessible dataset from 42 cooperative fungicide trials (6 locations and 9 years). Each independent experiment (totaling 143 experiments) was categorized as either epidemic or non-epidemic based on the 20% incidence (percentage of diseased heads in the sample) as threshold in the untreated control plots. Daily meteorological variables including minimum (TMin), maximum (Tmax), and average temperature (Tmean), relative humidity (RH), and precipitation (P) were gathered from the NASA Power website for each trial and summarized. The wheat heading date (WHD) was used to delineate four window panes, each consisting of two seven-day intervals, one before and one after the WHD. These were combined with the weather-based variables, resulting in 56 predictors (14 weather variables × 4 windows). Logistic regression models were fit to the binary data, with variable selection accomplished using LASSO. The *bestglm* algorithm was employed to choose the optimal combination of variables based on the Bayesian information criterion. LASSO regression selected nine significant predictors, six of which were utilized and combined into four models. The final models comprised six variables: mean RH, the number of days with temperatures below 22ºC, and the interaction of T mean and mean RH, in the second week before WHD, the interaction of T mean and mean RH 1 week before WHD, and mean RH and total rainfall in the first week after WHD. These models included 1 to 4 predictors, with accuracies ranging between 0.80 and 0.83. Sensitivities varied from 0.74 to 0.91, specificities from 0.72 to 0.86, Kappa coefficients from 0.61 to 0.66, and AUC values from 0.84 to 0.90. The models, after further validation by specialists, may be used to assess the risk of wheat blast, thereby informing management decisions in the regions where the data were observed.

Financial support: FAPEMIG (Project # 123) and CAPES (Project #123)