Epidemiology

Risk Assessment Models for Wheat Fusarium Head Blight Epidemics Based on Within-Season Weather Data

E. D. De Wolf, L. V. Madden, and P. E. Lipps

First author: Department of Plant Pathology, The Pennsylvania State University, Buckhout Laboratory, University Park 16802; and second and third authors: Department of Plant Pathology, The Ohio State University/OARDC, 1680 Madison Ave., Wooster 44691.

ABSTRACT

De Wolf, E. D., Madden, L. V., and Lipps, P. E. 2003. Risk assessment models for wheat Fusarium head blight epidemics based on within-season weather data. Phytopathology 93:428-435.

Logistic regression models for wheat Fusarium head blight were developed using information collected at 50 location-years, including four states, representing three different U.S. wheat-production regions. Non-parametric correlation analysis and stepwise logistic regression analysis identified combinations of temperature, relative humidity, and rainfall or durations of specified weather conditions, for 7 days prior to anthesis, and 10 days beginning at crop anthesis, as potential predictor variables. Prediction accuracy of developed logistic regression models ranged from 62 to 85%. Models suitable for application as a disease warning system were identified based on model prediction accuracy, sensitivity, specificity, and availability of weather variables at crop anthesis. Four of the identified models correctly classified 84% of the 50 location-years. A fifth model that used only pre-anthesis weather conditions correctly classified 70% of the location-years. The most useful predictor variables were the duration (h) of precipitation 7 days prior to anthesis, duration (h) that temperature was between 15 and 30°C 7 days prior to anthesis, and the duration (h) that temperature was between 15 and 30°C and relative humidity was greater than or equal to 90%. When model performance was evaluated with an independent validation set ($n = 9$), prediction accuracy was only 6% lower than the accuracy for the original data sets. These results indicate that narrow time periods around crop anthesis can be used to predict Fusarium head blight epidemics.

Additional keywords: disease forecasting, *Fusarium graminearum*, *Gibberella zeae*, head scab.

Epidemics of Fusarium head blight (FHB), caused by *Gibberella zeae* (Schwein.) Petch (anamorph: *Fusarium graminearum* Schwabe), have had a deleterious impact on wheat (*Triticum aestivum* L. em. Thell) production in many regions of North America. Accounts of severe epidemics of FHB, or scab, were recorded by pioneering plant pathologists, such as J. C. Arthur (4), A. D. Selby (26), and others (1,5,25) during the late 1800s and early 1900s. In recent years, epidemics of FHB have occurred in many wheat-producing states of the United States, including North Dakota, Minnesota, South Dakota, Ohio, Indiana, Michigan, Missouri, Kansas, and Arkansas (16). Epidemics of FHB result in severe losses through direct reduction in grain yield and increased grain cleaning costs (16). Fusarium-damaged grain is of particular concern because of mycotoxins (i.e., deoxynivalenol) produced by *G. zeae* in diseased kernels (16,23). Estimated losses to growers, grain handlers, and industries that utilize wheat-related products in North Dakota, Minnesota, and South Dakota during 1993 alone have exceeded $1 billion (16). Severe losses have continued to occur throughout many of the wheat-production regions of the United States (21).

FHB has proven to be a difficult disease to manage because of limitations in control options (7). Despite these difficulties, progress has been made by recent efforts to improve host resistance (3,22) and efficacy of chemical and biological controls (17). Regardless, FHB remains a major concern for wheat producers. A disease forecasting system providing wheat producers with reliable and timely management recommendations is desirable. A timely disease warning also would provide valuable time for grain handlers and food processors to deal with the prognosis of disease and the potential for mycotoxin contaminated grain by establishing the necessary infrastructure to appropriately test for, and manage, the damaged grain.

One approach to developing a disease forecasting system is through risk assessment and management. Generally, a disease model is developed to estimate the probability (i.e., risk) of a undesirable event occurring at a given location and time (31). Fusarium head blight is well suited for risk assessment modeling because of the severity of epidemics, compounded losses resulting from mycotoxin contamination, and relatively narrow time periods of pathogen sporulation, inoculum dispersal, and host infection (8,10).

Research regarding the response of *G. zeae* to different environmental conditions has been previously reported (2,29,32,33). Infection is favored by extended periods (48 to 72 h) at >90% relative humidity (RH) with temperatures between 15 and 30°C (2). When wetness, or high moisture events, are discontinuous, infection can still occur but infection efficiency is reduced (2,6). Similarly, *G. zeae* perithecia development and sexual reproduction is favored by temperatures between 15 and 29°C, but is limited at temperatures >30°C (33). Moisture level also may influence the perithecia development. Experiments on the effects of water potential on perithecia development indicate that in vitro perithecia production is the greatest at –1.5 MPa, is limited below –4.0 MPa, and is reduced as water potential approaches 0 MPa (29).

Previous efforts to predict FHB epidemics have occurred in Asia (15,19,35), South America (18), and North America (10,12). Based on information gathered between 1953 and 1963 in Japan, Nakagawa et al. (19) reported correlation coefficients between amounts of *Fusarium* spp.-damaged grain and monthly weather summaries. They then provide a linear regression model for predicting FHB-damaged grain based upon a monthly summary of temperature, rainfall, and sunshine prior to heading. Efforts to predict FHB have been extensive in China, including 19 documented

Corresponding author: E. D. De Wolf; E-mail address: edd10@psu.edu

forecast models (15). Most of these models use combinations of temperature and rainfall to predict disease levels. However, Zhao et al. (35) proposed a prediction system that provides long-range disease forecasts based on relationships between sea temperatures and FHB severity. Moshchini and Fortungo (18) used 22 observations of disease severity from the Pampeana region of Argentina to develop linear equations for FHB predictions. The authors reported how different combinations of temperature, RH, and precipitation could be used to predict FHB severity in that region based on linear equations. In the United States, Francl et al. (10) have utilized a cooperative effort among research institutions to identify factors critical to the development of FHB epidemics. These findings suggest that information regarding both inoculum level and environmental conduciveness are important for accurate epidemic prediction. A model developed in Ontario, Canada (12) utilizes weather data obtained pre-anthesis and weather forecasts postanthesis to predict deoxynivalenol levels in the harvested grain.

Our objectives were to (i) identify weather variables associated with epidemic development using existing records of FHB epidemics in the United States, (ii) identify periods of crop growth during which available weather variables were most critical, and ultimately (iii) develop risk-assessment models for wheat FHB that provide information important to disease management.

MATERIALS AND METHODS

Researchers from North Dakota, Ohio, Missouri, and Kansas contributed information about FHB severity, hourly weather, and crop growth from hard red spring wheat, hard red winter wheat, and soft red winter wheat production regions where FHB epidemics have occurred (Table 1). Differences in wheat cultivars as well as methods of assessing disease severity (or other measures of the impact of the disease) among the regions resulted in no consistent quantitative scale for the magnitude of FHB epidemics. However, within each region there was a separation between years with a major epidemic of FHB and years with little or no disease. Thus, each location-year was classified as to whether there was a major epidemic (binary code of "1") or not ("0"). This classification corresponded to cases with FHB field severities of ≥10% ("1") or <10% ("0"). In one location (Kansas), FHB was measured strictly in terms of yield loss. Only 2 years of data were used from this state, with yield losses of 0 and 24% (the former being classified as no epidemic and the latter being classified as a major epidemic). All cases used to develop the models came from replicated cultivar, fungicide, or other field evaluations. The total data set used to develop the risk-assessment models consisted of 50 cases from four states, 12 locations, and more than 18 years of data collections (Table 1). Epidemics were recorded in 18 of the 50 cases.

Hourly weather observations of temperature (degree Celsius), RH (percent), and precipitation (millimeters) were collected by automated weather-monitoring equipment near the replicated field plots. Proximity of weather-monitoring equipment to field experiments ranged from approximately 10 m to 2 km. Instrumentation varied among locations. All weather observations were evaluated graphically to identify potential errors or missing data. Location-years with errors in weather observations or time periods with missing data were eliminated from the data sets. Hourly weather observations for 30 days before and after crop anthesis were summarized into 24-h observations. To minimize the impact of splitting natural wetness or high-RH periods, the 24-h ("day") summaries began and ended at 1200 h (noon). Crop growth stage notes were specific to replicated plots from which disease data were collected.

Predictor variables. Information from previous research efforts was used to construct predictor ("independent") variables possibly useful during the modeling (2,9,10,18,19,24,29,32,33). These vari-

ables were obtained from (i) the calculation of daily minimums, maximums, averages, and sums; (ii) the duration (h) that pre-specified conditions were met; (iii) the functional response to temperature and moisture; and (iv) the calculation of interaction terms between predictor variables. A total of 49 variables were constructed from the original hourly observations of temperature, RH, and precipitation. Only those variables identified as most useful in later portions of the analysis will be discussed in detail. All variables were scaled between 0 and 1, by dividing by the observed maximum to ensure uniformity of variable scale, and to facilitate the calculation of interpretable interaction terms.

Nonparametric analysis was used to identify variables potentially associated with the binary representation of FHB epidemic status (27). Variables with a Kendall correlation (SAS Institute Inc., Cary, NC) coefficient of <0.23 (significance level of >0.05) were dropped from the modeling process unless they were deemed to contain potentially important biological information not already represented in the variables with greater correlation coefficients. This process reduced the number of potential variables to 25.

Logistic regression models. Logistic regression (SAS Institute Inc.) was used to model the relationship between FHB epidemics and predictor variables (13). As with the more familiar linear regression, the goal of logistic regression is to develop the best fitting, most parsimonious, yet biologically rational model of a given relationship. However, the dependent variable for logistic regression is binary, or dichotomous, with an assumed binomial distribution of errors. Predictions of the logistic model can be used as estimates of the probabilities of events given a set of independent variables. This predicted probability lends itself naturally to risk assessment analysis.

Variables identified in the correlation analysis were used as predictor variables to develop logistic regression models for classifying FHB epidemics. A stepwise logistic procedure initially was used only as a guide, and multiple logistic models with one, two, or more predictor variables were evaluated. The prediction accuracy (percentage of correctly classified cases) was determined for each evaluated model, and predictor variables that produced the highest accuracy were further evaluated by considering their interactions. For instance, if X_1, X_2, and X_3 represent three predictor variables, then two-way interactions (e.g., X_1X_2), and the three-way interaction ($X_1X_2X_3$) were calculated. Box plots of the interaction terms and predictor variables were used to assess potential separation of epidemic and nonepidemic location-years based on a given variable. The interaction terms then were used in a stepwise logistic regression analysis with the previously identified predictor variables to determine the models with the highest prediction accuracy.

TABLE 1. Information on the data sets used to develop logistic models of predicting epidemics of Fusarium head blight of wheat based on weather data

State	Location	Years[a]	Wheat class[b]
Ohio	Wooster	16	SRW
Ohio	Hoytville	2	SRW
Ohio	South Charleston	1	SRW
North Dakota	Fargo	7	HRS
North Dakota	Cando	4	HRS
North Dakota	Langdon	4	HRS
North Dakota	Carrington	2	HRS
North Dakota	Dazey	1	HRS
Missouri	Novlty	5	SRW
Missouri	Columbia	3	SRW
Missouri	Lamar	3	SRW
Kansas	Powhattan	2	HRW

[a] Number of years included in this analysis from location indicated in previous column.
[b] SRW = soft red winter wheat, HRS = hard red spring wheat, and HRW = hard red winter wheat.

Assuming that p is the probability of an epidemic, the logistic model with X_1 and X_1X_2 as predictors can be written as

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_1 X_2 \qquad (1)$$

in which β_0 to β_2 are parameters. In nominal use, a location-year is classified, for validation purposes, as an epidemic if the predicted p based in equation 1 is greater than 0.5. However, the highest accuracy may correspond to a different p. The critical p for classifying a location-year as an epidemic (p^*) is determined by calculating sensitivity and specificity for each predicted p, and selecting the p with the largest summation of sensitivity and specificity (28). Sensitivity is defined as the percentage of correctly classified epidemics and specificity is the percentage of correctly classified nonepidemics. Prediction errors (misclassified location-years) of the models with the highest prediction accuracies were analyzed to aid in the evaluation of model performance.

The accuracy of the identified models was assessed with a second validation data set. This data set consisted of nine location-years representing locations from Ohio, North Dakota, and Pennsylvania, and were obtained after models were developed. Weather and disease variables were recorded using methods described for data used in model development. Validation data remained independent of the data set used to develop the logistic regression models.

RESULTS

Variable associations with scab epidemics. Correlation analysis and stepwise logistic regression procedures identified 25 variables that were individually related to, or related in combination with, FHB epidemics across the 50 location-years (Table 2). Kendall correlations for these 25 variables ranged from 0.11 to 0.54. Many of these variables were intercorrelated and, thus, could not be used together in a model to classify location-years. For instance, T157 and T15307 values were very close for most years because there were few hours during this time of year (in the years studied) with temperatures greater than 30°C. In general, the magnitude of the correlations were higher for the anthesis variables than for the pre-anthesis variables. The variable most correlated with epidemics of FHB was TRH9010 (Table 2), the hours of RH ≥ 90% that were also between 15 and 30°C for the 10 days following anthesis. This variable also was the most significant single variable in a logistic regression model. Variables that summarized humidity at thresholds <90%, or weighted RH values approaching the 90% threshold, had reduced correlation coefficients than those that used the ≥90% threshold. Similarly, variables that summarized temperature values weighted above or below the reported optimums (2,33) also had reduced correlation coefficients.

Stepwise logistic regression identified three variables as jointly having a strong relationship with FHB epidemics and nonepidemics. The first was TRH9010 and the other two summarized weather over the 7 days prior to crop anthesis (T15307 and

TABLE 2. Weather variables potentially of value in predicting epidemics of Fusarium head blight of wheat were identified based on nonparametric correlation coefficient, F statistic for individual variable inclusion in a logistic model, or inclusion as part of group variables in a multiple-logistic model

Variable	Definition[a]	Kendall correlation[b]	F value[c]
Pre-anthesis weather variables[d]			
MT7	Average daily minimum T (°C)	0.28	6.07
T157	Duration (h) T ≥ 15°C	0.25	4.48
T15307	Duration (h) of 15 ≤ T ≤ 30°C	0.29	5.99
RH7	Average daily RH (%)	0.25	5.90
RHR7	Sum of RH response	0.20	2.80
TRHR7	Combination of T and RH response functions	0.30	7.78
RH907	Duration (h) RH ≥ 90%	0.26	4.08
TRH907	Duration (h) 15 ≤ T ≤ 30°C, and RH ≥ 90%	0.40	12.57
TRH807	Duration (h) 15 ≤ T ≤ 30°C, and RH ≥ 80%	0.30	8.18
PPT7	Sum of precipitation (mm)	0.23	1.82
DPPT7	Duration of precipitation (h)	0.11	1.75
Anthesis weather variables[e]			
MT10	Average daily minimum T (°C)	0.29	5.25
T1510	Duration (h) T ≥ 15°C	0.27	6.38
T153010	Duration (h) of 15 ≤ T ≤ 30°C	0.36	10.65
TR10	Sum of T response function	0.25	5.78
RH10	Average RH (%)	0.48	19.06
RHR10	Sum of hourly RH response function	0.40	13.48
TRHR10	Combination of T and RH response function	0.49	17.89
RH9010	Duration (h) RH ≥ 90%	0.41	15.69
RH8010	Duration (h) RH ≥ 80%	0.40	14.34
RH5D10	5 day moving window for average RH (%)	0.46	17.85
TRH9010	Duration (h) 15 ≤ T ≤ 30°C, and RH ≥ 90%	0.54	31.22
TRH8010	Duration (h) 15 ≤ T ≤ 30°C, and RH ≥ 80%	0.47	21.30
PPT10	Sum of precipitation (mm)	0.23	1.65
DPPT10	Duration of precipitation (h)	0.13	1.20
Interaction terms			
INT1	Interaction term = T15307 × DPPT7	0.23	4.52
INT2	Interaction term = DPPT7 × TRH9010	0.44	15.00
INT3	Interaction term = T15307 × TRH9010	0.55	40.84
INT4	Interaction term = T15307 × DPPT7 × TRH9010	0.48	20.47

[a] T = temperature, RH = relative humidity, sum of RH response = values derived from an equation developed to estimate a response of *Gibberella zeae* to relative humidity or temperature (11).

[b] Kendall correlation coefficient for weather variable given in first column and disease level coded as a binary variable (1 = epidemic; 0 = nonepidemic, or low disease).

[c] F statistic calculated by stepwise regression procedure.

[d] All variables correspond to the 7 days prior to anthesis.

[e] All variables corresponded to the 10 days following the initiation of anthesis (i.e., the 10 days starting when 50% of flowering has occurred).

DPPT7). Several of the individual variables that were highly associated with FHB epidemics (e.g., TRH8010) were not included in the multiple-variable logistic model because of their correlation with other variables in the model. Conversely, one of the variables identified by the stepwise logistic regression, DPPT7, individually had only a low correlation with epidemics. Two- and three-way interaction terms were calculated for these three variables, and all subsequent analyses and models were based on these variables and their interactions.

Box plots of predictor variables and interaction terms support the results of the correlation and stepwise analysis (Fig. 1). The greatest degree of separation of epidemic and nonepidemic location-years visible in the box plots was associated with the TRH9010 variable, and related interaction terms. When considered singly, variables summarizing pre-anthesis weather (T15307 and DPPT7) provided only a limited amount of visual class-separation relative to TRH9010.

Logistic regression models. When the three identified variables and their interaction terms were used to develop logistic regression models, 11 models were identified with prediction accuracies ranging from 62 to 84% (Table 3). Of the 11 models, 10 significantly ($P < 0.05$) predicted epidemics. Four models (A to D) correctly classified 84% of the 50 cases. One of these four models used only the temperature and humidity combination variable postanthesis, TRH9010 (model A). The other three models utilized at least one of the interaction terms. Both pre-anthesis and anthesis variables were included in models B to D, usually in the form of an interaction. Two models (models C and D) correctly classified >90% of the nonepidemics (specificity), but correctly classified <73% of the FHB epidemics (sensitivity). The other two models (A and B) had nearly equal sensitivity and specificity. Each of these four identified models (A to D) incorrectly classified eight cases, but only four were incorrectly classified by all these models (Table 4). Two of these four errors were false negatives and two were false positives.

The remaining models identified by the stepwise procedure (E to K) had prediction accuracies (percentage of correctly classified cases) between 62 and 82% (Table 3). Some of these were very complicated, with many predictor terms (E and H). This can result in biased parameter estimates and difficulty in interpretation, and extrapolation (20). The others had either lower sensitivity or lower prediction accuracy than did models A to D.

Models that used only pre-anthesis weather (I, J, and K) generally were inferior to those that used postanthesis variables of combinations of pre- and postanthesis. Although the models using either pre-anthesis precipitation or temperature (J and K) had high specificities (>80%), they had extremely low sensitivities (<34%). In contrast, the postanthesis model using one variable (A) had high sensitivity and specificity. The highest prediction accuracy for a model that used only pre-anthesis weather variables was 70% (model I). This model incorrectly classified a total of 15 of the 50 location-years. Of these 15 location-years, 8 were false negative predictions. Six of the location-years incorrectly classified by model I corresponded to cases with errors made by the identified models that used anthesis weather variables (A to D) (Table 4). However, model I did correctly classify two of the four location-years that models A to D had incorrectly classified.

Model assessment. Based on model prediction accuracy, fit, and simplicity, the models that used either TRH9010 (A), or T15307 × TRH9010 (INT3; B) variable to predict FHB epidemics were selected for further evaluation as potential FHB risk assessment models. In addition, a model using only pre-anthesis weather (I) was included in the further analysis because it had the potential to provide predictions early enough to make management decisions without forecasting the weather during anthesis.

The three logistic models (A, B, and I) were rearranged algebraically to solve for the weather variables that would give a
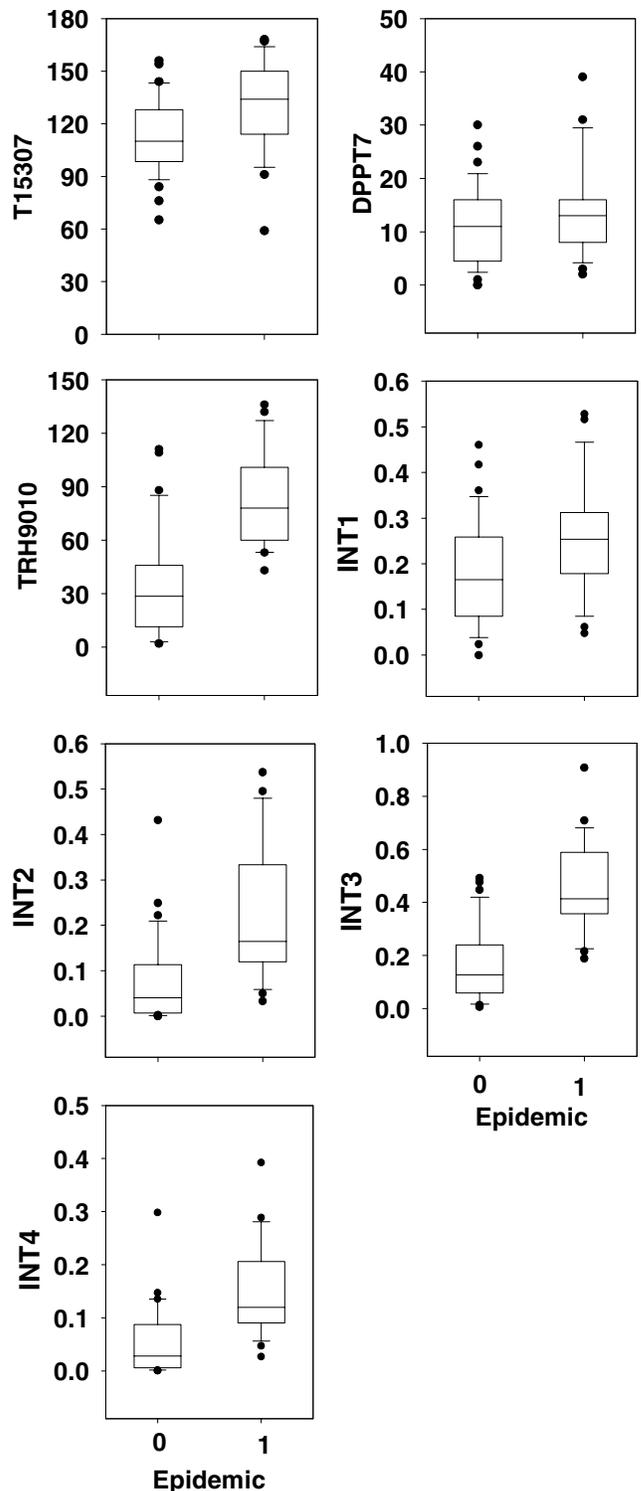
Fig. 1. Box plots of three weather variables and their interactions (products) identified by Kendall correlation and stepwise regression procedures as potential predictors of wheat Fusarium head blight epidemics. Each box represents the distribution of weather data from epidemic (coded as 1) and nonepidemic years (0). The line within the box is the median. The top and bottom lines of the box represent 25th and 75th percentile of the data. Lines extending vertically beyond the box represent the 10th and 90th percentiles, and solid circles indicate outliers. T15307 = duration (h) of $15 \le T \le 30°C$; TRH9010 = duration (h) $15 \le T \le 30°C$, and RH $\ge 90\%$; DPPT7 = duration of precipitation (h); INT.1 (interaction term) = T15307 × DPPT7; INT.2 = DPPT7 × TRH9010; INT.3 = T15307 × TRH9010; and INT.4 = T15307 × DPPT7 × TRH9010.

predicted probability of an epidemic equal to the critical value (p^*). The following equations were obtained:

$$\text{TRH9010} = \frac{Logit(p^*) - \beta_0}{\beta_1} \quad (2)$$

$$\text{TRH9010} = \frac{Logit(p^*) - \beta_0}{\beta_1 \, \text{T15307}} \quad (3)$$

$$\text{T15307} = \frac{Logit(p^*) - \beta_0 - \beta_1 \, \text{DPPT7}}{\beta_2} \quad (4)$$

with the p^* values and β estimates for model equations are given in Table 3, and all variables scaled between 0 and 1. For equation 2, a scaled TRH9010 value of >0.41 indicates an epidemic; returning to original units, this indicates an epidemic when TRH9010 is >56 h (0.41 × 136). This threshold can be represented graphically with a straight line passing through a scatter plot of epidemic and nonepidemic location-years; thus separating the two classes by respective TRH9010 values (Fig. 2A). The proximity of

each point in the scatterplot relative to this line, or decision threshold, confers the likelihood of membership in either class. Points lying near the decision threshold have a near-equal likelihood of being classified correctly into either class. Moreover, as distance from the decision threshold increases, the probability of a given point (location-year) being classified correctly into the appropriate class also increases. For example, if THRH90 = 71 h (model A), the probability of an epidemic is 51% (Table 3). Similarly, if TRH9010 = 131 h, the probability of an epidemic is 96%. These same principles apply for equations 3 and 4; however, only for these equations a combination of variables determines the prediction (Fig. 2A and B). For example, equation 3 contains the interaction term INT3 (T15307 × TRH9010), resulting in a curved decision threshold with respect to the environmental conditions both prior to (T15307) and during crop anthesis (TRH9010) (Fig. 2A). For instance, at T15307 = 105 h, TRH9010 must be at least 72 h for an epidemic to be predicted. However, at T15307 = 150 h, TRH9010 needs only to exceed 51 h for an epidemic to be predicted. In the case of equation 4, because there was no interaction term, pre-anthesis temperature and rainfall are combined to

TABLE 3. Logistic models developed for classifying location-years for Fusarium head blight epidemics, together with prediction accuracy, sensitivity, and specificity

Model	Model equation[a]	p^{*}[b]	Predicted accuracy (%)[c]	Sensitivity (%)[d]	Specificity (%)[e]
A	−3.3756 + 6.8128TRH9010	0.36	84	83	84
B	−3.7251 + 10.5097INT3	0.44	84	83	84
C	−1.0585 + −14.2400INT1 + 39.4590INT4	0.42	84	72	91
D	−1.5424 + 31.7868INT4 + −5.8085DPPT7	0.50	84	67	94
E[f]	18.5627 + −36.4659T15307 + −56.6710DPPT7 + −75.8177TRH9010 + 87.3140INT1 + 171.0INT2 + 128.3INT3 + −256.9INT4	0.44	82	78	84
F	−4.3732 + 5.5900INT2 + 9.8315INT3	0.50	82	72	88
G	−9.3117 + 9.4323T15307 + 13.4019INT2	0.52	82	67	91
H	−12.0661 + 9.2514T15307 + 4.5941DPPT7 + 7.5393TRH9010	0.50	78	72	81
I	−8.2175 + 8.4358T15307 + 4.7319DPPT7	0.50	70	56	78
J[f]	−4.2804 + 5.1343T15307	0.52	68	33	88
K[f]	−1.1646 + 1.8297DPPT7	0.44	62	17	88

[a] Logistic regression models were developed using data collected in Ohio, North Dakota, Missouri, and Kansas (n = 50). Variables are defined in Table 2. Model variables were rescaled between 0 and 1 by dividing by the maximum observed value in cases considered in this analysis. To use the equations, variables must first be placed in same scale as data used to develop the models. This can be done by dividing TRH9010, T15307, or DPPT7 by 136, 168, or 39, respectively (TRH9010/136; T15307/168; DPPT7/39). The probability of a greater χ^2 value for a likelihood ratio test for models A to H was $P < 0.0001$, and $P < 0.0015$, 0.0155, and 0.1869 for models I, J, and K, respectively.
[b] The critical predicted probability of a severe epidemic, p^*, as determined by calculating sensitivity and specificity for the full range of p values, and choosing the p with the largest summation of sensitivity and specificity.
[c] Predicted accuracy as percentage of correctly classified cases of epidemic and nonepidemics (epidemic = disease severity of ≥10%).
[d] Sensitivity, as percentage of correctly classified epidemics.
[e] Specificity, as percentage of correctly classified nonepidemic.
[f] Predictor variable not selected by the stepwise procedure, but model included for comparison.

TABLE 4. Location-years incorrectly classified by one or more models, with information on corresponding year, location, and values of some weather variables used and not used in the models

	Location/year[a]											
	1/1991	2/1997	3/1997	4/1996	4/1998	5/1996	5/1999	6/1994	7/1998	7/1993	7/1995	8/1997
FHB severity[b]	0	0	0	0	50	15	8	18	12	45	12	10
Model[c]												
Model A	x	x	x	x	…	…	x	x	…	x	x	…
Model B	x	x	x	x	…	…	x	x	…	x	…	x
Model C	x	…	x	x	…	x	…	x	x	x	x	…
Model D	…	…	x	x	x	x	…	x	x	x	…	x
Model I[d]	…	…	…	x	x	x	x	…	…	x	…	x
Weather variables[e]												
T15307	98	103	111	116	157	135	144	133	167	91	154	59
DPPT7	3	0	14	26	2	3	10	14	7	16	12	39
TRH9010	111	109	84	88	132	58	67	43	60	54	53	73

[a] Location of observations 1 = Hoytville, OH; 2 = Wooster, OH; 3 = Columbia, MO; 4 = Novlty, MO; 5 = Cando, ND; 6 = Carrington, ND; 7 = Fargo, ND; 8 = Langdon, ND.
[b] FHB = Fusarium head blight.
[c] Location-years incorrectly classified by the indicated model; x = error; … = correct. Model equations are provided in Table 3.
[d] Model I had an additional seven errors not included in this comparison with models A to D.
[e] T15307 = duration (h) of 15 ≤ T ≤ 30°C; DPPT7 = duration of precipitation (h); and TRH9010 = duration (h) 15 ≤ T ≤ 30°C, and RH ≥ 90%.

produce a sloped straight line that functions as the decision threshold (Fig. 2B). The lower accuracy for model I is demonstrated in Figure 2B by the multiple occurrences of open circles (epidemics) below the line and solid circles (nonepidemics) above the line.

Model validation. When the performance of models A, B, and I was evaluated with the validation data set, models A and B correctly classified 7 of the 9 location-years and incorrectly classified the same 2 location-years; both errors were false positive epidemic predictions. Model I correctly classified 6 of the 9 validation cases. Two of the errors made by model I were false negative, and the remaining error a false positive. Errors of model I did not correspond with errors made by models A and B.

DISCUSSION

In general, FHB epidemics from the 50 location-years evaluated in this analysis were associated with extended periods of favorable temperature (15 to 30°C) and extended periods of high RH (≥90%) before or during anthesis. This result is in agreement with past observations of weather conditions associated with epidemics (1,4,5,16,24,25,30).

The value of temperature and humidity variables used in classifying FHB epidemics supports the conclusions of previous research conducted in controlled environments (2,9,33). Specifically, the 56-h critical value of TRH9010 for the logistic model that used this variable alone (model A) is within the range of high RH durations (48 to 72 h) determined in experiments at constant temperatures between 15 and 30°C (2). The usefulness of the pre-anthesis temperature and precipitation variables (T15307 and DPPT7) may reflect the influence of temperature and moisture on perithecia development by *G. zeae* (9,24,33). Although prediction accuracy of models that used only variables that summarized pre-anthesis weather was lower than models that used variables that summarized weather conditions during anthesis, the 70% accuracy of model I indicates that pre-anthesis conditions have potential value as predictors of FHB epidemics. In the future, research to clarify the relationship between perithecia development and weather variables may aid in the identification of other pre-anthesis variables that could improve model prediction accuracy.

The importance of temperature and moisture variables to the prediction of FHB epidemics also is supported by other efforts to predict disease level (10,18,19), or mycotoxin contamination (12). Moreover, Hooker et al. (12) reported the importance of temperature and moisture variables during similar periods of host growth. However, these models differ considerably in the way that temperature and moisture are represented. For example, the models proposed by Hooker et al. (12) represent moisture in the form of rain as opposed to RH, and use different minimum and maximum temperature thresholds than proposed in this report. Direct comparison between the prediction models for FHB and associated mycotoxins were limited because one or more of the variables (i.e., inoculum level or deoxynivalenol content) needed for the different modeling efforts where not available in the existing data sets. Future validation analysis will need to address comparisons in model performance.

Logistic regression models. This analysis identified four logistic regression models with 84% prediction accuracy (Table 3; models A to D). Given that the models used data from only 17-day windows, and were applied to different cropping systems and wheat types, this prediction accuracy was considered high. Estimates of model accuracy with a validation data set (78%) were within 6% of estimates of model accuracy with the model generating data sets. Based on model prediction accuracy and high level of sensitivity (percentage of correctly classified epidemics), models A and B were selected for further evaluation.

An error analysis can be useful for understanding model properties. Errors of models A and B were associated with location-years

when severity level was near (±2%) the preselected 10% field disease severity threshold, or when weather variables that likely contributed to the epidemic were not considered by the models (Table 4). For example, the 1993 Fargo location-year was incorrectly classified as a low disease year by the models. In this year, the 10-day time period during anthesis was associated with 54 h of favorable temperatures and RH as defined by the TRH9010 variable. This value is very close to the 56-h threshold of TRH9010 established for model A (Fig. 2). Moreover, this location-year received 21 h of rainfall during this 10-day period, and 137 mm of rain in the 7 days following the time period considered by this model. Thus, weather conditions (i.e., frequent rains) not considered by model A likely contributed to the high disease levels observed in this year. The impact of variables not considered by the model on model errors was particularly evident with model I. Errors for this model, which used only pre-anthesis variables, most commonly corresponded with location-years that had conditions during crop anthesis favorable for infection (Table 4). For example, model I incorrectly classified the 1998 Novlty location-year as nonepidemic based on limited (2-h) duration of pre-anthesis rainfall. However, this model did not consider the 132 h of favorable temperature and humidity (TRH9010) that
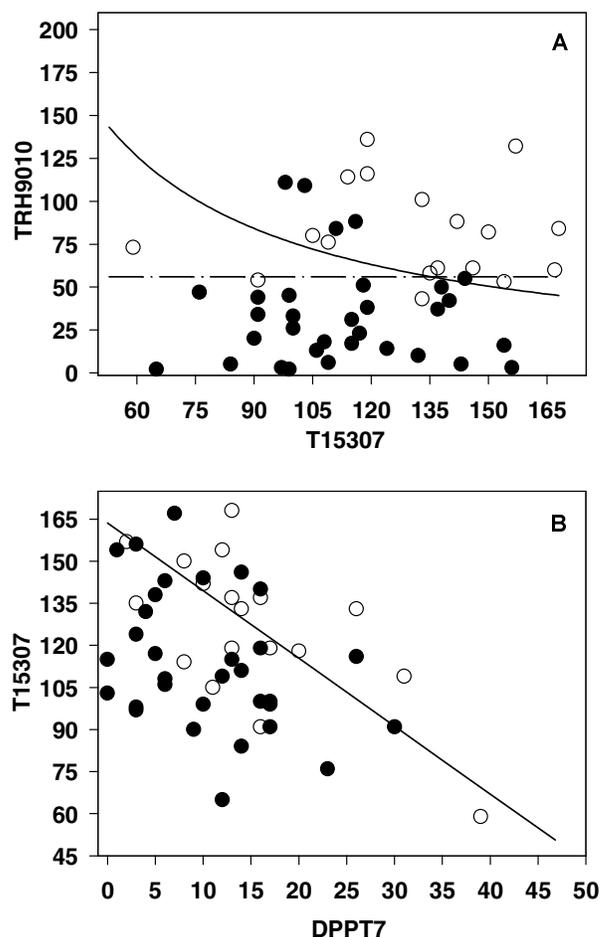
Fig. 2. Values of weather variables for the 50 location-years used in models A, B, and I; and threshold values (lines) for predicting an epidemic of Fusarium head blight based on equations 2 to 4. An observed epidemic is indicated with an open circle and a nonepidemic is indicated with a solid circle. **A,** Solid line (equation 3) represents the threshold for model B, and the broken line for model A (equation 2). **B,** Line represents the threshold for model I (equation 4). T15307 = duration (h) of 15 ≤ T ≤ 30°C; TRH9010 = duration (h) 15 ≤ T ≤ 30°C, and RH ≥ 90%; DPPT7 = duration of precipitation (h). Models are provided in Table 3. The coefficients in the models of Table 3, and corresponding equations 2 to 4, are based on scaled predictor variables.

occurred during the anthesis time period in that year. The 1999 Cando location-year is an example of a model error associated with disease severity close to the severity threshold. In this case, average disease severity for the replicated plots was estimated to be 8% (Table 4), but models A, B, and I predicted an epidemic of 10% severity or greater based on recorded weather variables. Variability in disease assessment or influence of field-scale environment and inoculum levels may account in part for this error. Patterns in model errors identified with the validation cases were consistent with errors described for model development data.

Attempts to incorporate variables that had potential to improve prediction accuracy, such as precipitation during anthesis (as done by Hooker et al.) (12), did not improve the accuracy of models A and B (E. D. De Wolf, *unpublished data*). This may be due, in part, to the infrequency of these extreme weather events in the 50 location-years evaluated in this analysis. Incorporating nonsignificant terms in a regression model leads to biased and imprecise estimates of parameters (19). Thus, other variables were not added to the models.

Model application. The application of these predictive models, like many weather-driven prediction systems, will depend on the availability, resolution, and reliability of weather data. A potential limitation of both models A and B (as well as models C and D) was the dependence on weather information during anthesis when the most successful fungicide applications have been timed. While it may be possible to overcome this limitation by using forecasted weather, the uncertainty of the predicted weather variables may reduce model prediction accuracy. In contrast, the model that used only pre-anthesis variables (model I) had a 14% lower prediction accuracy than models A or B, but was free from potential limitations associated with using forecasted weather variables. However, precise forecasts of temperature and RH are not needed to use the models with TRH9010, only forecasts for extended periods of high RH with moderate temperatures. Further research is needed on forecasts of these variables in wheat-growing areas.

One possible way of using these models would be to combine the pre-anthesis model (I) with the two models that use weather conditions after the initiation of crop anthesis (models A and B), and make successive predictions in a given year. For example, a preliminary prediction of the probability of an FHB epidemic could be made at the start of the flowering period (Feeks' GS 10.5.1) based on observed weather and model I, or possibly based on model I in combination with forecasted weather and models A and B. Disease prediction could be updated with models A and B as more real-time weather information became available during the anthesis period. A final prediction then could be made at the end of the 17-day period for the probability of an FHB epidemic provided by all three models. The model predictions can be influenced by unusually wet conditions outside the 17-day time period, or precipitation conditions not considered by the model; therefore, prediction will never be 100% accurate. Thus, predictions will be subject to uncertainty about weather conditions in subsequent time periods. The advantages to such an approach include disease predictions with maximum available observed data at anthesis, when a fungicide application might be warranted, and then shifting priorities to predicting *Fusarium* spp.-damaged grain and risk of mycotoxin contamination to allow for the establishment of grain handling priorities.

The current level of model accuracy represents a considerable advancement toward predicting FHB epidemics in North America considering that data from only a 17-day window are utilized. However, it may be possible to improve model accuracy through the incorporation of precipitation variables during the anthesis time period (12). Precipitation variables considered in the present analysis did not improve model accuracy; therefore, it is probable that novel representations of this information will need to be evaluated for possible contribution to model accuracy. Given the importance of crop residues in pathogen survival (14,34), it also may be possible to improve model performance through the addition of variables containing information about local inoculum sources. This information was unavailable for this analysis, but factors such as the presence of crop residue and residue type (i.e., corn, wheat, or soybean) may serve as valuable predictors of disease at a local level when coupled with critical weather variables. The presence of inoculum from residue may be related to some of the prediction errors with the models. The identified models should be of use in predicting the risk of FHB epidemics in the regions for which they were developed. As with all disease prediction models, it is advisable to evaluate model performance before application in other wheat production regions.

LITERATURE CITED

1. Adams, J. F. 1921. Observations on wheat scab in Pennsylvania and its pathological histology. Phytopathology 11:115-125.
2. Andersen, A. L. 1948. The development of *Gibberella zeae* head blight of wheat. Phytopathology 38:595-611.
3. Anderson, J. A. 2001. Variety development and uniform nurseries: Progress in FHB resistance in hard spring wheat. Page 225 in: Proc. 2001 Natl. Fusarium Head Blight Forum, Erlanger, KY.
4. Arthur, J. C. 1891. Wheat scab. Purdue Univ. Agric. Exp. Stn. Bull. No. 36 Vol. II.
5. Atanasoff, D. 1920. Fusarium-blight (scab) of wheat and other cereals. J. Agric. Res. 20:1-32.
6. Bai, G.-H., and Shaner, G. 1991. Effect of moist period on response of wheat cultivars to infection by *Fusarium graminearum*. (Abstr.) Phytopathology 81:1145.
7. Bai, G.-H., and Shaner, G. 1994. Scab of wheat prospects for control. Plant Dis. 78:760-766.
8. De Wolf, E. D., Lipps, P. E., Francl, L. J., and Madden, L. V. 1999. Role of environment and inoculum level in wheat Fusarium head blight development. Pages 87-91 in: Proc. 1999 Natl. Fusarium Head Blight Forum, Sioux Falls, SD.
9. De Wolf, E. D., Lipps, P. E., and Madden, L. V. 2000. Crop residue moisture and *Gibberella zeae* perithecia development. Page 136 in: Proc. 2000 Natl. Fusarium Head Blight Forum, Erlanger, KY.
10. Francl, L., Shaner, G., Bergstrom, G., Gilbert, J., Pedersen, W., Dill-Macky, R., Sweets, L., Corwin, B., Jin, Y., Gallenberg, D., and Wiersma, J. 1999. Daily inoculum levels of *Gibberella zeae* on wheat spikes. Plant Dis. 83:662-666.
11. Hau, B., Eisensmith, S. P., and Kranz. J. 1985. Construction of temporal models: II. Simulation of aerial epidemics. Pages 31-65 in: Advances in Plant Pathology Mathematical, Vol. 3. Mathematical Modeling of Crop Disease. C. A. Gilligan, ed. Academic Press, New York.
12. Hooker, D. C., Schaafsma, A. W., and Tamburic-Ilincic, L. 2002. Using weather variables pre- and post-heading to predict deoxynivalenol content in winter wheat. Plant Dis. 86:611-619.
13. Hosmer, D. W., and Lemeshow S. 2000. Applied Logistic Regression. 2nd ed. John Wiley & Sons, Inc., New York.
14. Khonga, E. B., and Sutton, J. C. 1988. Inoculum production and survival of *Gibberella zeae* in maize and wheat residues. Can. J. Plant Pathol. 10:232-239.
15. Lu, W. H. Chen, S. H., and Wang, Y. Z. 2001. Research on Fusarium Head Scab in China. Science Publisher, Beijing.
16. McMullen, M., Jones, R., and Gallenberg, D. 1997. Scab of wheat and barley: A re-emerging disease of devastating impact. Plant Dis. 81:1340-1348.
17. Milus, E. A., Hershman, D., and McMullen, M. 2001. Analysis of the 2001 uniform wheat fungicide and biocontrol trials across locations. Page 75 in: Proc. 2001 Natl. Fusarium Head Blight Forum, Erlanger, KY.
18. Moschini, R. C., and Fortugno, C. 1996. Predicting wheat head blight incidence using models based on meteorological factors in Pergamino, Argentina. Eur. J. Plant Pathol. 102:211-218.

19. Nakagawa, M., Wanatabe, S., Gocho, H., and Nishio, K. 1965. Studies on the occurrence forecast of *Gibberella zeae* on wheat varieties which caused by climatic factors. Bull. Tokai-Kinki Natl. Agric. Exp. Stn. 15:55-67.

20. Neter, J., Wasserman, W., and Kutner, M. H. 1989. Applied Linear Regression Models. Richard D. Irwin, Inc., Boston.

21. Nganje, W. E., Johnson, D. D., Wilson, W. W., Leistritz, F. L., Bangsund, D. A., and Tiapo, N. M. 2001. Economic impacts of Fusarium head blight in wheat and barley: 1998-2000. Agribusiness Appl. Econ. Rep. No. 464. Agric. Exp. Stn. North Dakota State University, Fargo.

22. Ohm, H. W. 2001. Variety development and uniform nurseries: Winter wheat research progress. Page 269 in: Proc. 2001 Natl. Fusarium Head Blight Forum, Erlanger, KY.

23. Parry, D. W., Jenkins, P., and McLeod, L. 1995. *Fusarium* ear blight (scab) in small grain cereals—A review. Plant Pathol. 44:207-238.

24. Paulitz, T. C. 1996. Diurnal release of ascospores by *Gibberella zeae* in inoculated wheat plots. Plant Dis. 80:674-678.

25. Pugh, G. W., Johann, H., and Dickson, J. G. 1933. Factors affecting infection of wheat heads by *Gibberella saubinetii*. J. Agric. Res. 46:771-797.

26. Selby, D., and Manns, T. F. 1909. Studies in diseases of cereals and grasses: I. New anthracnose diseases of certain cereals and grasses; II. The fungus of wheat scab as a seed and seedling parasite. Also general on small grains. Bull. Ohio Agric. Exp. Stn. No. 203. Wooster.

27. Sprent, P., and Smeeton, N. C. 2001. Applied Nonparameteric Statistical Methods, 3rd ed. Chapman & Hall, New York..

28. Stokes, M. E., Davis, C. S., and Koch, G. G. 1995. Categorical Data Analysis Using the SAS System. SAS Institute, Cary, NC.

29. Sung, J. M., and Cook, R. J. 1981. Effect of water potential on reproduction and spore germination by *Fusarium roseum* 'Graminearum,' 'Culmorum,' and 'Avenaceum'. Phytopathology 71:499-504.

30. Sutton, J. C. 1982. Epidemiology of wheat head blight and maize ear rot caused by Fusarium graminearum. Can. J. Plant Pathol. 4:195-209.

31. Teng, P. S., and Yang, X. B. 1993. Biological impact and risk assessment in plant pathology. Annu. Rev. Phytopathol. 31:495-521.

32. Tschanz, A. T., Horst, R. K., and Nelson, P. E. 1975. Ecological aspects of ascospore discharge in *Gibberella zeae*. Phytopathology 65:597-599.

33. Tschanz, A. T., Horst, R. K., and Nelson, P. E. 1976. The effect of environment on sexual reproduction of *Gibberella zeae*. Mycologia 68:327-340.

34. Windels, C. E., and Kommedahl, T. 1984. Late-season colonization and survival of *Fusarium graminearum* group II in corn stalks in Minnesota. Plant Dis. 68:791-793.

35. Zhao, S., and Yao, C. 1989. On the sea temperature prediction models of the prevailing level of wheat scab. Acta Phytopathol. Sin. 19:229-234.