

# Modeling of Relationships Between Weather and *Septoria tritici* Epidemics on Winter Wheat: A Critical Approach

S. Pietravalle, M. W. Shaw, S. R. Parker, and F. van den Bosch

First and fourth authors: Rothamsted Research, Rothamsted, Harpenden, Hertfordshire AL5 2JQ, UK; second author: Department of Agricultural Botany, School of Plant Science, The University of Reading, Reading RG6 6AS, UK; and third author: ADAS High Mowthorpe, Duggleby, Malton, North Yorkshire YO17 8BP, UK.

Current address of S. R. Parker: CSL, Sand Hutton, York, North Yorkshire YO41 1LZ, UK.

Accepted for publication 16 May 2003.

## ABSTRACT

Pietravalle, S., Shaw, M. W., Parker, S. R., and van den Bosch, F. 2003. Modeling of relationships between weather and *Septoria tritici* epidemics on winter wheat: A critical approach. *Phytopathology* 93:1329-1339.

Two models for predicting *Septoria tritici* on winter wheat (cv. Riband) were developed using a program based on an iterative search of correlations between disease severity and weather. Data from four consecutive cropping seasons (1993/94 until 1996/97) at nine sites throughout England were used. A qualitative model predicted the presence or absence of *Septoria tritici* (at a 5% severity threshold within the top three leaf layers) using winter temperature (January/February) and wind speed to about the first node detectable growth stage. For sites above the

disease threshold, a quantitative model predicted severity of *Septoria tritici* using rainfall during stem elongation. A test statistic was derived to test the validity of the iterative search used to obtain both models. This statistic was used in combination with bootstrap analyses in which the search program was rerun using weather data from previous years, therefore uncorrelated with the disease data, to investigate how likely correlations such as the ones found in our models would have been in the absence of genuine relationships.

*Additional keywords:* binary data, data mining, discriminant analysis, Window Pane.

When developing weather-based predictive models for plant diseases, two main approaches can be followed. First, mechanistic models have been used to improve understanding of the spread of disease (20). These models are generally based on physical properties of the observed phenomenon and therefore mirror it quite closely. However, they may sometimes be very complex, involving many parameters, and may be of limited use to growers to predict disease severity. Alternatively, empirical models have been derived from data mining. Their main advantage is that they are usually simpler to use at a field scale.

In 1982, a program (Window Pane) using a data mining algorithm was developed (7) in the United States. Because of its large range of possible applications and the fairly simple theory behind it, it has been used by some authors to derive predictive models in plant pathology. Its range of applications has included arable crops (5,8,9,14), papaya (28), and banana (6). Similar data mining algorithms have been developed and used for disease (16,17) and pests (34). However, the statistical validity of the data mining method used by this type of statistical program has not been addressed adequately. Its validity has been criticized because a potential problem of using an exhaustive iterative search is that it is likely to detect some significant but spurious relationships. Therefore, such models are often seen as unreliable and are not being used in the field.

In the United Kingdom, and more generally in Europe, there has been overapplication of fungicide to control *Septoria tritici* for many years (19). The top three leaf layers contribute a large per-

centage of the photosynthate accumulated in kernels, and growers therefore aim to protect these leaves from disease. Fungicide applications at first node detectable (growth stage [GS] 31) (41) and flag leaf ligule just visible (GS 39) provide good control on the upper three leaves. However, depending on disease risk, the appropriate fungicide dose applied at these timings ranges from zero to the maximum permitted. In the United Kingdom, fungicide doses do not appear to be adjusted in response to risk. Because growers need early indication of the risk of *Septoria tritici* in order to purchase products and prioritize farm activity, any model likely to be widely adopted would need to provide accurate predictions of *Septoria tritici* prior to first node detectable (GS 31).

This paper has two main objectives. First, we develop an improved version of the algorithm in order to look for predictive models. In past studies, disease severity has been studied as a continuous variable (between 0 and 100%). However, severe occurrences of disease are relatively rare in wheat (18) so observations taken from crops are commonly skewed toward low disease severity. As a result, the number of degrees of freedom of the regression is artificially increased and so may be the goodness of the fit. We therefore developed a new approach to this problem by introducing a two-step analysis. It first predicts the occurrence of an epidemic (qualitative approach), and when an epidemic occurs, it predicts its severity (quantitative approach). We also address the problem of the validity of the method using bootstrap techniques and explain why results from Window Pane are not straightforward to analyze. The second objective of this work is to improve predictive precision for *Septoria tritici* which is the most important foliar disease of winter wheat in the United Kingdom (18). In the past, research focused on understanding the summer spread of splash-dispersed conidiospores, but recent work (15,29) has highlighted the importance of the winter epidemic. In this work, we confirm this pattern and discuss the splash-driven part of the epidemic.

Corresponding author: S. Pietravalle  
E-mail address: stephane.pietravalle@bbsrc.ac.uk

## MATERIALS AND METHODS

**Crop and site details.** Because it was the most commonly grown wheat cultivar in the United Kingdom at the time of the work (20 to 30% of wheat area throughout the United Kingdom), observations used in this study were obtained from plots (2 × 18 m) of cv. Riband. It is a cultivar highly susceptible to *Septoria tritici*; good levels of resistance were not available for commercial cultivars at the time of the study. The observations were made as part of a multisite experiment conducted in four crop seasons (1993/94 to 1996/97) throughout England (Fig. 1). All field sites had at least a 1-year break from cereals. Plots were sown by a plot drill between early October and early November according to sites. The average of four replicate plots was used as a measure of the disease at the site. Weeds and insect pests were controlled with pesticides, in accordance with normal farm practice, but no fungicides were applied. Data from four combinations of sites and years were missing. First, no trial was run at Wye in 1996/97 because of poor weather conditions. Second, problems of labeling occurred when collecting the data at Bridgets in 1993/94 and made them unusable. Finally, poor establishment of the crop meant that trials had to be abandoned at Arthur Rickwood in 1995/96 and Gleadthorpe in 1996/97. A fifth combination of site and year (High Mowthorpe in 1993/94) did not have leaf size measurements. This meant that disease severity, based on leaf area, could not be calculated at that site.

**Meteorological data.** Eight meteorological variables were collected daily at the experimental sites during the experiment. They were maximum (Tmax; °C) and minimum (Tmin; °C) temperature above the canopy (1.25 m), total rainfall (Rain; mm), average wind speed (Wind; m s<sup>-1</sup>), maximum (Hmax; %) and minimum (Hmin; %) relative humidity above the canopy, total solar radiation (Rad; mJ m<sup>-2</sup> s<sup>-1</sup>), and surface (leaf) wetness (Wet; %). Surface wetness was measured using two sensors (SWS Sensors; Delta-T Devices Ltd., Burwell, Cambridge, UK). The sensitive part of these sensors consists of three carbon electrodes separated by waterproof resin. The surface of the sensing area is ridged to increase sensitivity. Records were taken as “wet” or “dry,” and leaf wetness was defined as the proportion of time the sensor was recorded as wet (using 1-min intervals). Two further weather variables were derived from these records, average daily temperature (Tmean; °C) and average daily relative humidity above the canopy (Hmean; %), both were calculated as half the sum of the daily minimum and maximum values. Some weather data were missing due to logger/sensor failures. To alleviate this problem, minimum and maximum temperatures, rainfall, and wind speed from the nearest station to each site were used. At most sites, this station was located on the same farm. Relative humidity, leaf wetness, and solar radiation could not be substituted by this method so artificial neural network techniques were used to estimate missing values (M. Taylor, *personal communication*). A total of 49% of relative humidity, 41% of leaf wetness, and 46% of solar radiation observations were estimated in this way. A huge variability however was observed, with the proportion of missing data varying from 27% (radiation at Starcross) up to 100% (radiation at Gleadthorpe). We decided to use weather functions derived from these three weather variables in the analysis, recognizing that results from them would need to be interpreted with caution.

**Disease data.** Optimum fungicide dose depends on the level of untreated disease. Because cv. Riband is highly susceptible to *Septoria tritici* and has very low potential for disease escape, it makes the epidemics largely weather-driven and the setting of the inoculum pool during winter crucial. Practical experience shows that disease observation at first node detectable (GS 31) is not a good predictor of the severity of later epidemics, but disease observation at medium milk (GS 75) is a good predictor. At GS 75, disease severity was assessed on 10 randomly sampled shoots

per plot on all leaf layers with an average of 25% or greater green leaf area. Symptoms of *Septoria tritici* were estimated as the area covered by pycnidia and all associated senescence. Measurements of the phenology of the plant (leaf size, growth stage, and leaf area index), crop density (number of fertile tillers per square meter), percent green leaf area, and green leaf area index were also made. Growth stage and percent green leaf area also were measured on 10 different plants and averaged for each block. Because leaf size is subject to very little variability and measurements are very time-consuming, leaf width and length were measured on only two plants per plot and averaged. All sampling was destructive. Leaf area index and green leaf area index were calculated from the above measurements. The absolute areas of a subsample of the assessed leaves were measured by the method of Bryson et al. (4). These measurements, combined with fertile shoot counts, allowed green area and disease area to be expressed as index values (the dimensionless planar area, expressed per unit of ground area occupied).

Because the top three leaves contribute most to wheat yield (35,39), final disease severity was calculated on these top leaves. Disease severity within the top three leaves was then calculated as the percentage of leaf area with symptoms, taking into account the differences of leaf size between leaf layers:

$$\text{Sev}_{\text{Total}} = \frac{\sum_{i=1}^3 \text{Sev}_i \times \text{Area}_i}{\sum_{i=1}^3 \text{Area}_i} \quad (1)$$

where  $\text{Sev}_i$  is the percentage of leaf  $i$  with symptoms and  $\text{Area}_i$  is the area of leaf  $i$ .

**Algorithm description: Window Pane.** A Genstat 6.1 algorithm (Genstat, VSN International Ltd., Hemel, Hempstead, UK), based on the original idea of Coakley et al. (9), was developed to find the weather functions that were most related to the observed disease severity. The basis of this program is to examine iteratively the correlation with disease of a range of functions (46 weather functions listed in Table 1) of the eight original weather variables (Table 1) at different periods of time and for different lengths of time (referred to as a window) (Fig. 2). In previous implementations of the Window Pane algorithm, the windows to calculate all the weather functions were defined with a starting date and a window length (9). This approach was inappropriate for our data because disease observation dates were different in the years and sites that were analyzed. Coakley et al. (9) solved this problem by readjusting all years and sites so that the observation date (e.g., at GS 75) was on the same adjusted Julian Day across all sites and years. In our work, we used a time lag (the number of days between the disease observation date and the start of the window) to define the windows (Fig. 2) and keep dates unchanged. This change made interpretation of the results from the program easier and made other generalizations that we tried easier to implement. For any weather function, the program searched for the window in which it was the most highly related to disease severity (the “optimum window”) using two nested loops. To start the process, a maximum time lag between the disease observation date (GS 75) and the start of the window was fixed by the user. In the inner loop, the window length was allowed to vary between two limits (the maximum and the minimum window lengths) according to a step length. Both the maximum and minimum window lengths as well as the step length were set by the user. For each window thus defined, the value of the weather function considered was calculated for each site × year combination and correlation between the derived vector and the observed disease severity was calculated. For each time lag, the maximum absolute value of this correlation across window lengths was found. Then, in the outer loop, this process was repeated for a series of time lags decremented in steps set by the user, starting from the maximum time lag, until the minimum time lag (also set by the

user) was reached. At each repeat of the outer loop, the entire procedure described in the inner loop was run. Then, the entire process, as defined by the inner and outer loops, was repeated for the remaining weather functions.

**Disease data as a continuous variable.** When locating the best disease-weather relationships in time, we first had to find a balance between the precision of the analysis and the large number of windows that could be searched. In order to do so, we started with runs using 5-day steps for both the time lag and window length. If the correlation at the optimum time lag and window length for a given weather function was significant at  $P = 0.01$ , a more accurate run of Window Pane was done around the initial optimum window in order to refine the results from that first search. Given a first best time lag ( $T_1$ ) and window length ( $L_1$ ), the second run used 1-day steps and tested lags varying from  $T_1 + 5$  down to  $T_1 - 5$ , and within each lag, window lengths from  $L_1 + 5$  down to  $L_1 - 5$ . The final output of this first version of the modified Window Pane therefore consisted of tables with weather functions, time lags, optimum window lengths for each time lag, correlations, and  $P$  values testing for the individual significance of the correlations (Table 2).

**Disease data as a binary variable.** The original version of Window Pane could not be used to analyze binary data. However, disease is often simply quantified as present or absent. Window Pane was therefore amended to allow the analysis of binary data. Although the iterative process on which the algorithm is based remained unchanged, the measure of association between disease and weather differs.

The vector of observed severities was first converted into a binary vector representing site  $\times$  year combinations with more (1) or less (0) than a defined threshold level of disease (e.g., severity greater or equal to 5% within the top three leaf layers). Then, for any weather function  $X$  (e.g.,  $T_{\min\text{Nod},7}$ ) and any delay, the optimum value,  $x_0$ , to separate diseased site  $\times$  years from healthy

site  $\times$  years was found by minimizing the misclassification rate  $m$  over the  $n$  site  $\times$  year combinations available:

$$m = \frac{\min \left( \sum_{i=1}^n |Y_i - X_i^{(-)}|, \sum_{i=1}^n |Y_i - X_i^{(+)}| \right)}{n} \quad (2)$$

where  $Y_i$  is the observed severity (0 or 1) for site  $\times$  year combination  $i$

$$X_i^{(-)} = \begin{cases} 0 & \text{if } X_i > x_0 \\ 1 & \text{otherwise} \end{cases}$$

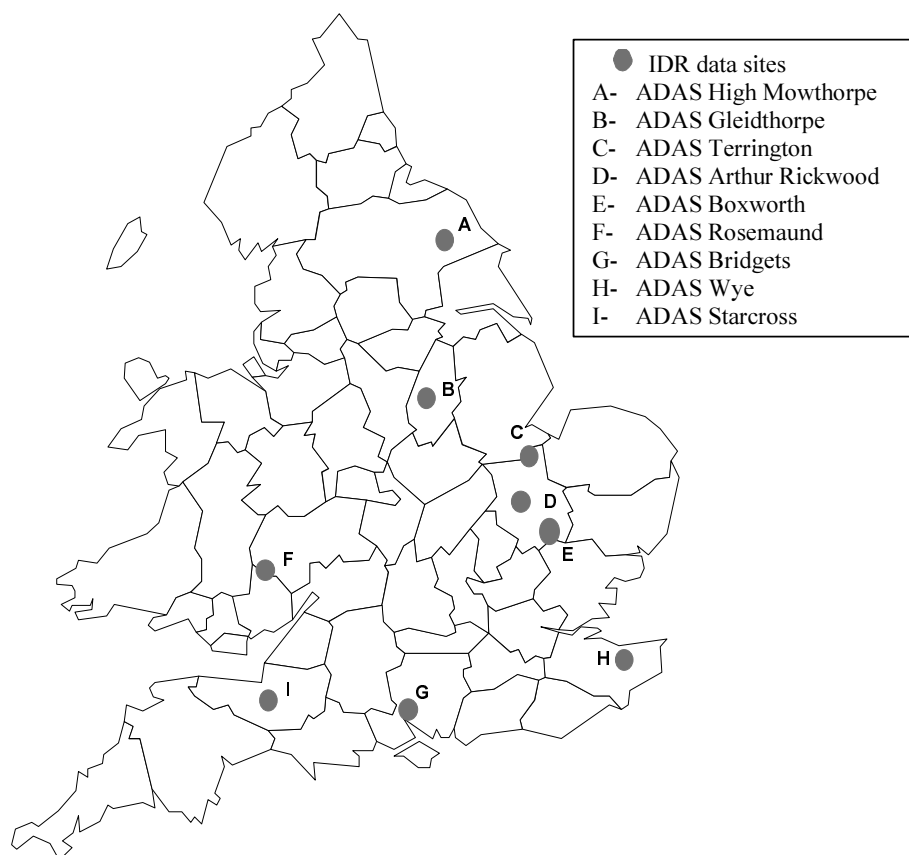
$$X_i^{(+)} = \begin{cases} 0 & \text{if } X_i \leq x_0 \\ 1 & \text{otherwise} \end{cases}$$

At each site  $\times$  year combination  $i$ , if one of the two terms in the right side of equation 2 is equal to 1, the other is equal to 0 by definition of  $X_i^{(-)}$ ,  $X_i^{(+)}$ , and  $Y_i$ . Therefore, when summing over all  $n$  site  $\times$  year combinations, if one of the two terms to minimize in the right side of equation 2 is greater than  $\text{int}(n/2)$ , the other is smaller. Thus, the misclassification rate  $m$  cannot exceed  $\text{int}(n/2)/n$ . The null hypothesis used to test the significance of the misclassification was therefore set to

$$H_0: m = \frac{\text{int}\left(\frac{n}{2}\right)}{n} \quad H_1: m < \frac{\text{int}\left(\frac{n}{2}\right)}{n} \quad (3)$$

and an approximate one-tailed  $Z$  test for comparing two proportions (21) was used.

**Bootstrap analysis.** As with all statistical questions, two problems face analyses such as the one used by Window Pane. First, chance relationships may be found and, because many possible relationships are investigated, there may be many of these. Second,



**Fig. 1.** Experimental sites in England used from 1994 until 1997 for assessing the severity of *Septoria tritici* blotch of winter wheat.

true relationships may be attributed to chance. We tried to investigate the difficulties these might cause the Window Pane methodology by testing it with data containing or lacking known relationships. To do so, bootstrap analyses were made to assess the validity of the correlations and misclassifications found by Window Pane. The idea driving these analyses was to use weather variables clearly unrelated to the disease data and rerun the Window Pane program. To do so, we kept the same disease data at each site  $\times$  year combination and reattributed new weather data to each combination. Then, Window Pane was used to calculate correlations or misclassifications and was tested using a given statistic. In order to keep realistic autocorrelations within the explanatory variables but lose the possible relationship between them and the disease variable, a new set of meteorological variables was derived from data collected at the same sites in six new years (1969/70, 1972/73, 1974/75, 1975/76, 1976/77, and 1977/78).

**The test statistic.** In the output spreadsheets from Window Pane runs, there are often a few time lags around the optimum window where the correlation is still highly significant ( $P \leq 0.01$ ). Because of the autocorrelation within many weather variables (e.g., temperature), it would not be surprising to find several such high correlations around the optimum. For instance, if  $Tmin_{Nod;7}$  [135,60] (where 135 is the line lag and 60 is the window length) is highly correlated with disease, it is very likely that  $Tmin_{Nod;7}$  [130,55] will be correlated with disease as well, even if both correlations were spurious. However, high correlations extending over a long range of lags (Table 2) are likely to be less frequent and may be a better indication of genuine relationships than the  $P$

value alone. Thus, the statistic  $N$  chosen when running each bootstrap analysis was the number of individually significant correlations or misclassifications (at a 1% level) in time lags around the optimum window (defined with a time lag  $T$ ) with a 5-day step (i.e.,  $T \pm 5, 10, 15, \dots$ ). When deriving any model, we therefore favor weather functions for which the  $N$  statistic is large.

**The two types of bootstrap.** We needed to define the way to reattribute the new weather data to each site  $\times$  year combination. Two methods were used.

The more stringent method was to keep both the spatial and temporal correlations: each site is given new weather data from a different year from the same site and each new year is reattributed the same way across sites. Because there were 6 years from which to choose an ordered subset of 4 years of weather data (the number of years the experiment was run), the total number of possibilities was more restricted in this case. It was the number of permutations of 4 years taken from the 6 years, that is  $P_6^4 = 6!/(6-4)! = 360$ . Selecting subsets by this method implies that the new years used are not necessarily consecutive. However, weather variables were not correlated between years, so this was not a problem.

The less stringent method was to take  $n$  site  $\times$  year combinations (where  $n$  is the total number of site  $\times$  year combinations in the data set) at random within the additional weather data set. Here, no account was taken of spatial correlation (i.e., one site might always be colder than another) or temporal correlation (i.e., 1 year might be more rainy than another at every site). In other words, the hypothesis was to test the statistic in the case where the only condition about the explanatory variables is their autocorrelation. In order to have two comparable analyses, a random subset of 360 allocations of disease data to site  $\times$  years was used in this case.

**Complete and partial bootstraps.** In addition to choosing the randomization method for the bootstrap analysis, we needed to choose which weather functions to include. Two different approaches were used.

First, the analysis was restricted to the weather functions that were found significant and kept within the final models. Second, all weather functions used in the runs with the original data set were included.

Because the new years were not part of the original project, only rain and minimum, maximum, and average temperature were available for this analysis, and the bootstrap was therefore restricted. However, as explained later, this was not a major problem. In order to compare them later, all four bootstrap possibilities (described previously) were tried with all three variants of the disease data: first, correlations using the entire quantitative data set; second, misclassification using the qualitative data set; and third, correlations using the quantitative data with only site  $\times$  years with severe damage.

**Deriving disease-weather relationships and validating the models.** When disease data was a continuous variable, knowledge

TABLE 1. Description of the weather functions studied in the analysis

Weather function	Description
Andd7	Accumulation of negative degree-days below 7°C
Andd14	Accumulation of negative degree-days below 14°C
Apdd7	Accumulation of positive degree-days above 7°C
Apdd14	Accumulation of positive degree-days above 14°C
Hmax <sub>Avg</sub>	Average maximum relative humidity (%)
Hmax <sub>Cnod;99</sub>	Number of consecutive days with maximum relative humidity of >99%
Hmax <sub>Nod;99</sub>	Number of days with maximum relative humidity of >99%
Hmean <sub>Avg</sub>	Average mean relative humidity (%)
Hmin <sub>Avg</sub>	Average minimum relative humidity (%)
Hmin <sub>Cnod;60</sub>	Number of consecutive days with minimum relative humidity of <60%
Hmin <sub>Nod;60</sub>	Number of days with minimum relative humidity of <60%
Rad <sub>Avg</sub>	Average solar radiation (MJ m <sup>-2</sup> s <sup>-1</sup> )
Rain <sub>Avg</sub>	Average rainfall (mm)
Rain <sub>Cnod;0-</sub>	Number of consecutive days without rain
Rain <sub>Cnod;0+</sub>	Number of consecutive days with rain of >0 mm
Rain <sub>Cnod;1</sub>	Number of consecutive days with rain of >1 mm
Rain <sub>Nod;0</sub>	Number of days with rain
Rain <sub>Nod;0.1</sub>	Number of days with rain of >0.1 mm
Rain <sub>Nod;1</sub>	Number of days with rain of >1 mm
Rain <sub>Nod;i</sub>	Number of days with rain of $\geq i$ mm ( $i = 1 \dots 9$ )
Tmax <sub>Avg</sub>	Average maximum temperature (°C)
Tmax <sub>Cnod;25</sub>	Number of consecutive days with maximum temperature of >25°C
Tmax <sub>Nod;25</sub>	Number of days with maximum temperature of >25°C
Tmean <sub>Avg</sub>	Average mean temperature (°C)
Tmin <sub>Avg</sub>	Average minimum temperature (°C)
Tmin <sub>Cnod;i</sub>	Number of consecutive days with minimum temperature of < $i$ °C ( $i = -2, 0, 7$ )
Tmin <sub>Cnod;7+</sub>	Number of consecutive days with minimum temperature of $\geq 7$ °C
Tmin <sub>Nod;i</sub>	Number of days with minimum temperature of < $i$ °C ( $i = -2, 0, 7$ )
Wet <sub>Avg</sub>	Average leaf wetness (%)
Wet <sub>Cnod;20</sub>	Number of consecutive days with leaf wetness of <20%
Wet <sub>Cnod;80</sub>	Number of consecutive days with leaf wetness of >80%
Wet <sub>Nod;20</sub>	Number of days with leaf wetness of <20%
Wet <sub>Nod;80</sub>	Number of days with leaf wetness of >80%
Wind <sub>Avg</sub>	Average wind speed (m s <sup>-1</sup> )

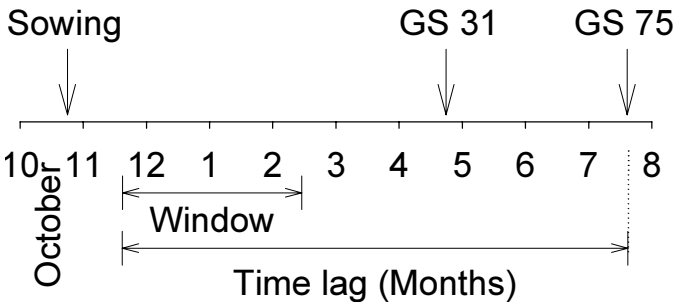


Fig. 2. Example of a window (from middle November until middle February) for a given time lag preceding the mid-milk growth stage (GS) of winter wheat (GS 75).

of the optimum window for correlation between each weather function and disease was used to derive linear models involving several weather functions. A systematic approach was used with the Genstat procedure Rsearch. This calculates several statistics, including the adjusted  $R^2$ , Mallows'  $C_p$  (24), and the Akaike information criterion (1,2), and allows the user to systematically choose between different subsets of explanatory variables for multiple linear regression. The PRESS statistic and variance inflation factor (VIF) were calculated for all submodels as additional tests. The  $C_p$  statistic compares nonnested models. It is defined as

$$C_p = \frac{RSS_p}{\hat{\sigma}^2} - n + 2p$$

where  $RSS_p$  is the residual sum of squares of the submodel considered (with  $p$  parameters),  $\hat{\sigma}^2$  is an unbiased estimate of the variance of the full model, and  $n$  is the total number of observa-

tions. Mallows (24) showed that models without significant lack of fit lead to points around the line  $C_p = p$ , hence providing a first guideline for submodel selection. The Akaike information criterion consists in minimizing  $-2(L_p - p)$ , where  $L_p$  is the log-likelihood of the submodel considered (with  $p$  parameters). The PRESS statistic for one model with  $p$  parameters is defined by

$$PRESS = \sum_{i=1}^n (Y_i - \hat{Y}_{ip})^2$$

where  $\hat{Y}_{ip}$  is the estimate of  $Y_i$  from the model derived using the original data set out of which the  $i$ th observation of the response and explanatory variables were deleted. As discussed by Draper and Smith (12), the choice of the submodel is then guided by low PRESS combined with a reasonably small number of explanatory variables. One of the reasons to keep the number of explanatory variables low is multicollinearity, an indicator of which is given by the VIF. It is defined (26) by

$$VIF = \max_{i=1 \dots p} \left( \frac{1}{1 - R_i^2} \right)$$

where  $R_i^2$  is the multiple correlation coefficient of  $X_i$  on the remaining predictors in the model and  $X_i$  is the  $i$ th of the  $p$  predictors of the model (3).

When disease data was a binary variable, after the best window and lag had been found for each weather function, we needed, as for the continuous data set, a way to combine a subset of weather functions to produce an optimum prediction. Discriminant analysis was used to derive a rule to classify each site  $\times$  year as diseased or not diseased according to the value of discriminant functions (linear combinations of the initial weather variables). As with the multiple linear regression, it was important to include only useful variables in this analysis. An  $F$  test (25) was used on the most general model to test and discard the least important variables given all the other variables. This test was then repeated

TABLE 2. Example of a part of the output from a Window Pane run

Weather function	Time lag (days)	Window length (days)	Correlation	<i>P</i> value
TminNod;7	165	60	-0.41	0.02
TminNod;7	160	55	-0.46	0.01
TminNod;7	155	50	-0.48	0.006
TminNod;7	150	45	-0.48	0.006
TminNod;7	145	65	-0.59	0.001
TminNod;7	140	65	-0.67	<0.001
TminNod;7	135	60	-0.70	<0.001
TminNod;7	130	55	-0.69	<0.001
TminNod;7	125	50	-0.63	<0.001
TminNod;7	120	45	-0.57	0.001
TminNod;7	115	40	-0.55	0.001
TminNod;7	110	35	-0.50	0.004
TminNod;7	105	30	-0.35	0.06
TminNod;7	100	60	-0.28	0.12

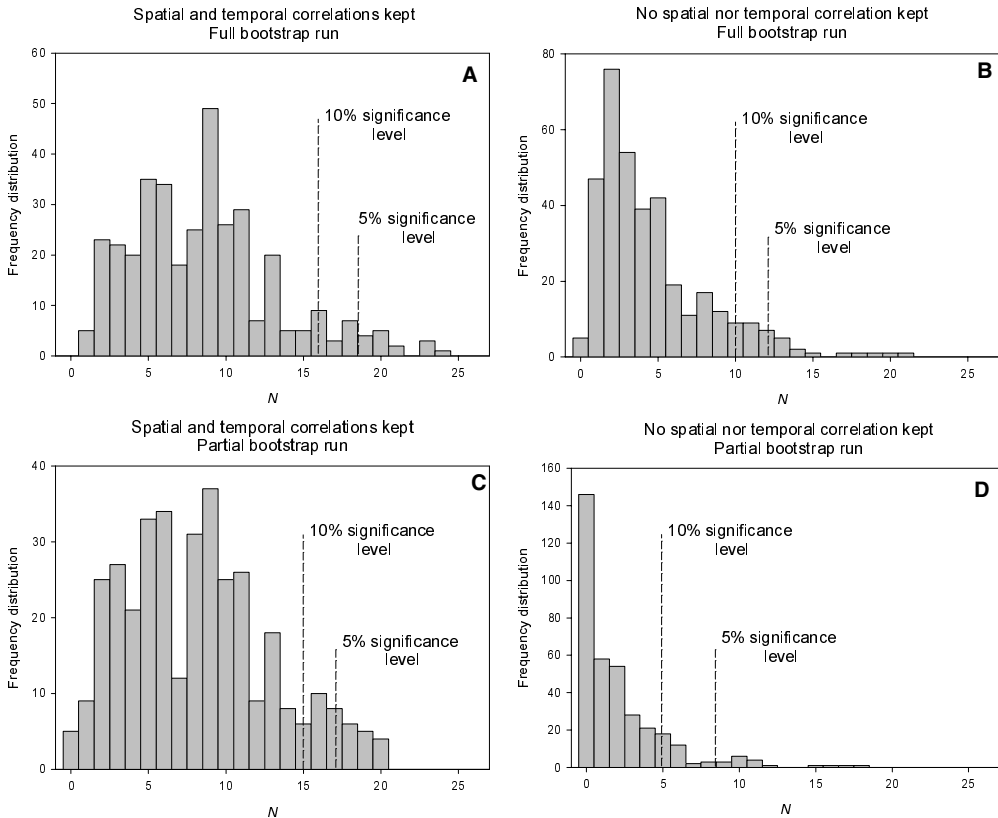


Fig. 3. Frequency distribution of the  $N$  statistic (number of consecutive significant correlations,  $P < 0.01$ , in windows around the optimum one) using continuous disease severity for the entire data set.

on the new model with one less variable until no further variables could be dropped without a significant increase in misclassification.

## RESULTS

**Method testing.** Before deriving the different predictive models from our data set, we present the results of the bootstrap analyses for each of the three types of data described previously. The first bootstrap analysis used quantitative data from all site  $\times$  year combinations. The frequency distributions of the statistic  $N$  when spatial and temporal correlations were kept and discarded, respectively, are shown in Figure 3A and B. In these first two cases, Window Pane was run using the entire set of meteorological variables available (full bootstrap). On the other hand, the distributions obtained when only the two weather functions from the final model (equation 4) that were available in the new meteorological data set (partial bootstrap) are shown in Figure 3C and D, and 0.9 and 0.95 quantiles are shown on each frequency distribution plot. Results from the analysis of binary data using misclassification and from the linear regression restricted to severely infected sites and years are shown in Figures 4 and 5.

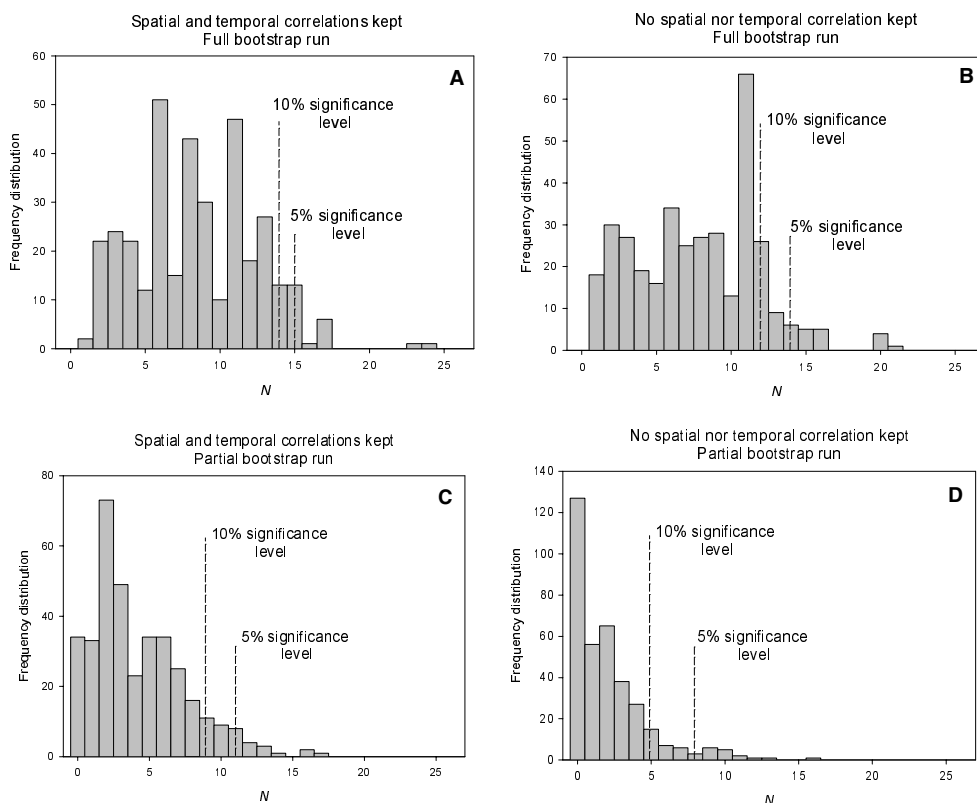
**Multiple linear regression.** All 31 combinations of sites and years were taken together, and disease severity was analyzed as a continuous variable. Initially, time lags varied between 225 (late November) and 70 days (late April), and within each of those, window length varied between 65 and 15 days. Although disease activity may be very important in later windows, we decided to restrict the analysis to windows early enough to allow the model to be used as a predictive tool. For some discrete variables (e.g., the number of consecutive days with minimum temperature greater or equal to 7°C), sometimes there was more than one optimum window. To resolve this, we chose the earliest window because one of the aims of our study was to derive predictive models to be used by farmers to optimize their use of fungicides, and the

longest time lag would provide farmers the greatest opportunity to adjust treatment decisions. When there was more than one optimum window length for a lag, the longest one was kept in order to allow the explanatory variable the maximum range of values. In order to reduce the number of possible explanatory variables to choose from, only the nine variables with absolute correlation strictly greater than or equal to 0.7 ( $P < 0.001$  with this sample size) were kept (Table 3).

The statistics described previously were used to choose among the different submodels. When more than four explanatory variables were included in any model, its VIF was high (larger than 5), indicating strong multicollinearity (27). The Genstat procedure Research therefore was used to compare all possible subsets of one up to four explanatory variables. All four statistics used (adjusted  $R^2$ , Mallows'  $C_p$ , Akaike information criteria, and PRESS statistic) agreed on the same "best" model (Fig. 6).  $F$  tests were used on nested models, and no further explanatory variable was found to significantly improve the regression. As a result, the following model was kept as a first predictor of *Septoria tritici* severity at GS 75:

$$S_{GS75} = 4.37(1.01)\text{Wet}_{\text{Nod};80}[62,17] + 3.12(1.48)\text{Tmin}_{\text{Cnod};7+}[145,70] + 2.78(1.26)\text{Rain}_{\text{Nod};9}[81,61] - 6.4(2.7) \quad \text{Adjusted } R^2 = 0.77 \quad (4)$$

where standard errors of the estimates are in parentheses and optimum time lags and window lengths are within square brackets, respectively. The observed growth stages for each of the three windows were stem elongation (approximately GS 30/39) for  $\text{Wet}_{\text{Nod};80}$ , seedling growth and early tillering (approximately GS 12/26) for  $\text{Tmin}_{\text{Cnod};7+}$ , and tillering until inflorescence (approximately GS 25/59) for  $\text{Rain}_{\text{Nod};9}$ . Because of the slight difference between sites and years as well as the frequency of the measurements, all three growth stages are approximate. The  $N$  statistics for the three weather functions kept in the final model (equation 4) were 14 ( $\text{Wet}_{\text{Nod};80}$ ), 17 ( $\text{Tmin}_{\text{Cnod};7+}$ ), and 11 ( $\text{Rain}_{\text{Nod};9}$ ). Figure 3



**Fig. 4.** Frequency distribution of the  $N$  statistic (number of consecutive significant, at a 1% level, misclassifications in windows around the optimum one) for binary disease data.

shows that, with the more stringent bootstrap, none of the  $N$  statistics were significant at a 5% level and only one was significant at a 10% level (Fig. 3A). Even when using the two bootstraps with intermediate type I and II errors, not all three  $N$  statistics were significant (Fig. 3B and C). With a full bootstrap when no spatial or temporal correlations were kept, two  $N$  statistics were significant at a 5% level and all three were significant at a 10% level. When both spatial and temporal correlations were kept but with a partial bootstrap, one  $N$  statistic was significant at a 5% level and one at a 10% level. Finally, Figure 3D shows that all three  $N$  statistics were significant at a 5% level in the less stringent bootstrap (omitting spatial and temporal correlations).

**Binary response and discriminant analysis.** Although the previous model explained much of the variability of severity of *Septoria tritici* at GS 75, it can easily be argued that the goodness-of-fit of this model was overestimated because of the cluster of points corresponding to years and sites with very low disease, which artificially increases the number of degrees of freedom. The weakness of this first model was confirmed by the bootstrap analysis. We showed that the  $N$  statistic for  $\text{Rain}_{\text{Nod};9}$  was significant only at a 5% level when using the less stringent bootstrap, therefore raising doubt on the validity of such a relationship. Further, it is more useful to determine the conditions driving the initial onset and progress of an epidemic. We therefore decided to first treat the severity as a binary variable. To do so, a threshold of 5% severity within the top three leaf layers was used so that sites and years with less than 5% severity at GS 75 were considered nonepidemic. The same time lags, window lengths, and steps were used as described previously. Then, the search was refined using 1-day steps and the 12 weather variables that lead to a misclassification of less than or equal to five site  $\times$  year combinations ( $m \leq 0.16$ ) were kept for further analysis (Table 4). Discriminant analysis was then used to derive a model from these weather variables. However, there was large variability in the  $N$  statistic among these weather variables (Table 4). Therefore, the bootstrap

was used to select the variables to include in the first model during the discriminant analysis and to discard those that were very likely to have been significant by chance. The frequency distribution of the  $N$  statistic when bootstrapping all the meteorological data is shown in Figure 4A and B. Although the frequency distribution of the  $N$  statistic was shifted slightly to the right when spatial and temporal correlations were kept (Fig. 4A), 0.5, 0.75, and 0.95 quantiles were similar (8, 11, and 15 compared with 8, 11, and 14, respectively, when keeping and ignoring spatial and temporal correlations). Consequently, when the discriminant analysis was done, a weather function was ignored when the  $N$  statistic was smaller than 11 (25% significance level from the bootstrap analysis; Fig. 4). Therefore, only 4 of the initial 12 variables were kept (average wind speed [78,25], average leaf wetness [57,22], number of days with minimum temperature less or equal to 7°C [190,70], and number of consecutive days with minimum temperature less or equal to 7°C [190,70]). In this study, the discriminant analysis was restricted to two groups (diseased

TABLE 3. The nine “best” weather functions with their respective optimum windows (correlation greater than or equal to 0.7,  $P < 0.001$  for single correlation) and the statistic  $N$  (number of windows with a significant [ $P < 0.01$ ] correlation around the optimum one)<sup>a</sup>

Weather function	Time lag	Window length	Correlation	$N$
Rain <sub>Avg</sub>	74	24	0.73	13
Wet <sub>Nod;80</sub>	62	17	0.80	14
Rain <sub>Cnod;1</sub>	56	7	0.79	13
Tmin <sub>Cnod;7+</sub>	145	70	0.76	17
Wet <sub>Cnod;80</sub>	66	22	0.78	13
Rain <sub>Nod;6</sub>	74	70	0.72	7
Rain <sub>Nod;7</sub>	80	70	0.73	14
Rain <sub>Nod;8</sub>	80	70	0.71	13
Rain <sub>Nod;9</sub>	81	61	0.71	11

<sup>a</sup> Data includes all years and sites.

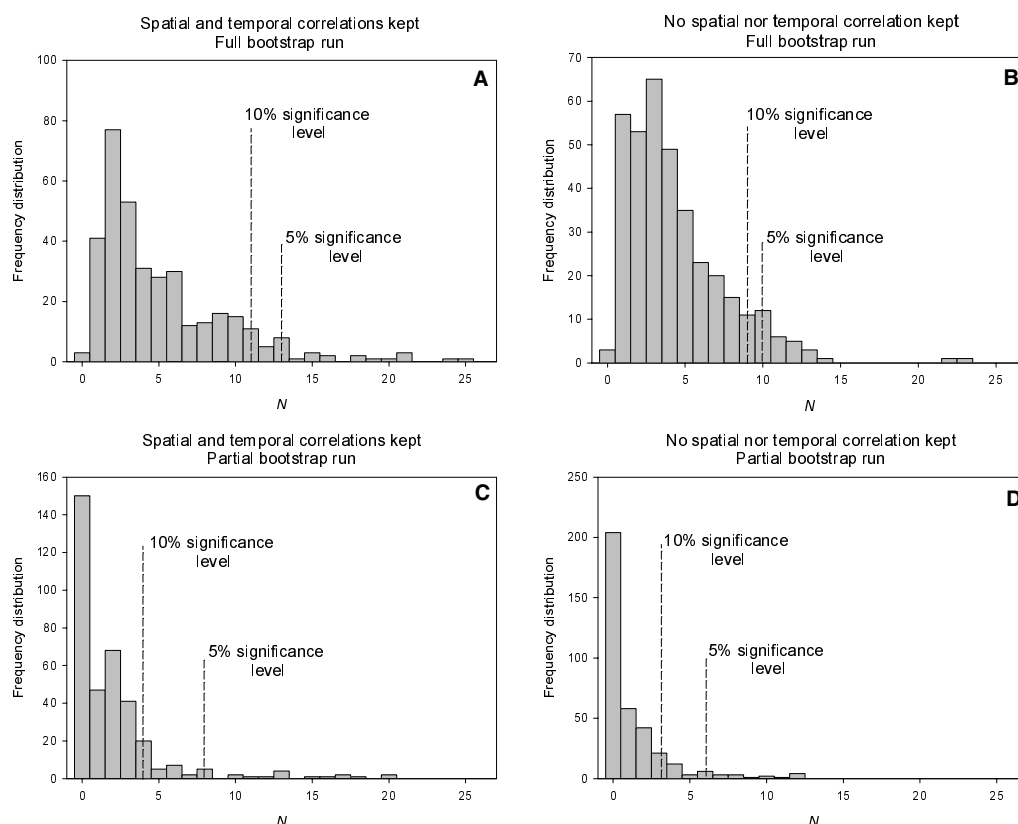


Fig. 5. Frequency distribution of the  $N$  statistic (number of consecutive significant, at a 1% level, correlations in windows around the optimum one) using continuous disease severity for sites and years with observed disease severity of *Septoria tritici* blotch greater than or equal to 5%.

sites,  $n_1 = 15$  and nondiseased sites,  $n_2 = 16$ ). Two variables ( $Tmin_{Cnod;7}$  and  $Wet_{Avg}$ ) were dropped from the initial discriminant function without any significant increase in misclassification. The final discriminant function is as follows: classify site<sub>*i*</sub> × year<sub>*j*</sub> as nondiseased if

$$h(x) = 2.7Wind_{Avg}[78,25] + 0.41Tmin_{Nod;7}[190,70] - 33.7 > 0 \tag{5}$$

and diseased otherwise (Fig. 7).

The two *F* statistics for the inclusion of these variables in the model were 7.9 ( $P = 0.01$ ) and 4.9 ( $P = 0.04$ ). Similar to equation 4, we can translate the lags into approximate growth stages corresponding to these windows. These were late tillering and early stem elongation (approximately GS 25/35) for  $Wind_{Avg}$  and seedling growth and early tillering (approximately GS 11/25) for  $Tmin_{Nod;7}$ . The *N* statistics (equation 5) were 25 for  $Wind_{Avg}$  and 14 for  $Tmin_{Nod;7}$ . Figure 4 shows that, judged against any of the bootstraps, the *N* statistic of both weather variables was significant at a 10% level. The *N* statistic for temperature was not significant at the 5% level with the more stringent bootstrap (Fig. 4A).

**A quantitative approach for high disease sites × years.** The previous analysis provided a qualitative predictor of *Septoria tritici* severity by distinguishing between severely and non-severely infected sites × years. The next logical problem to be solved was to predict disease severity more accurately in the case

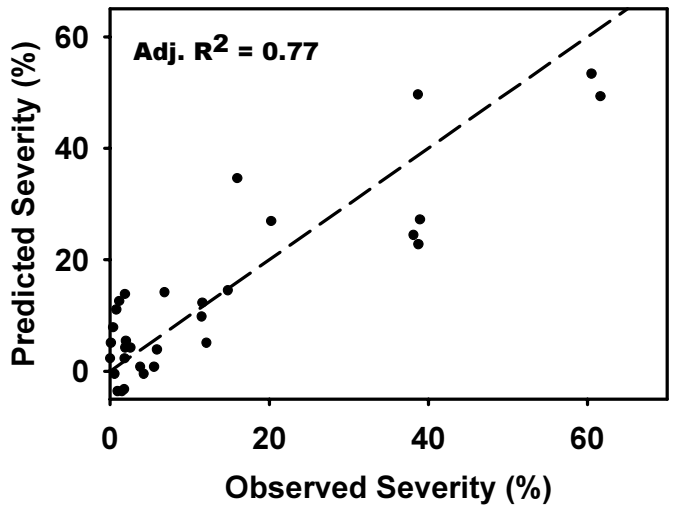


Fig. 6. Severity of *Septoria tritici* blotch measured at the mid-milk growth stage (GS) of wheat (GS 75) and predicted using the “best” multiple linear regression over all sites and years.

TABLE 4. The 12 “best” weather functions with their respective optimum windows (misclassification smaller than or equal to 0.16,  $P < 0.001$  for misclassification) and the statistic *N* (number of windows with a significant [ $P < 0.01$ ] misclassification around the optimum one)<sup>a</sup>

Weather function	Time lag	Window length	Misclassification	<i>N</i>
Rain <sub>Avg</sub>	225	26	0.16	3
Wind <sub>Avg</sub>	78	25	0.10	25
Tmax <sub>Avg</sub>	89	26	0.16	9
Wet <sub>Avg</sub>	57	22	0.16	13
Tmin <sub>Nod;7</sub>	190	70	0.10	14
Wet <sub>Nod;80</sub>	66	58	0.13	7
Tmin <sub>Cnod;7</sub>	190	70	0.10	14
Rain <sub>Nod;4</sub>	130	21	0.16	3
Rain <sub>Nod;5</sub>	135	45	0.06	9
Rain <sub>Nod;6</sub>	143	54	0.13	8
Rain <sub>Nod;7</sub>	151	28	0.16	9
Apdd7	187	67	0.16	4

<sup>a</sup> Data includes all years and sites.

of diseased years. To do so, the same multiple linear regression analysis as that used for all sites was repeated with the 15 site × year combinations in which observed severity was greater than 5%. There were potentially more explanatory variables than degrees of freedom. Therefore, when two very similar and highly correlated weather variables (e.g.,  $Rain_{Nod;8}$  and  $Rain_{Nod;9}$ ) were highly correlated with disease severity, only the most correlated one was kept, leading to 11 possible explanatory variables to choose from (Table 5). As described previously, models with more than four explanatory variables were not analyzed further because of multicollinearity. A complete search of all subsets with four or less variables was done. The adjusted  $R^2$  tended to select models with many variables (four or five), for which the explanatory variables were highly correlated. Both the Akaike information criterion and Mallows’ *C<sub>p</sub>* agreed on a simple linear regression (on  $Rain_{Nod;9}$ ) and the PRESS statistic favored a model with two explanatory variables ( $Tmean_{Avg}$  and  $Rain_{Nod;9}$ ). When *F* tests were used on nested models, no significant ( $P > 0.2$ ) improvement could be made by adding further weather variables to the simple linear regression and the perfected model was therefore as follows (Fig. 8):

$$S_{GS75} = 13.8(1.5)Rain_{Nod;9}[90,37] + 6.1(2.8) \text{ Adj. } R^2 = 0.86 \tag{6}$$

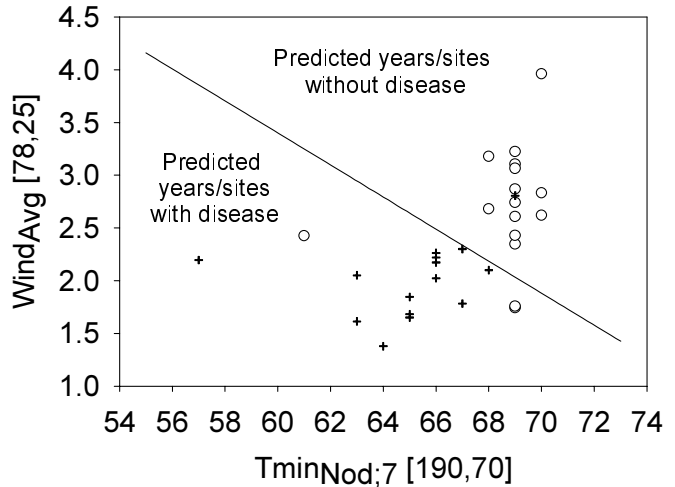


Fig. 7. Discriminant function classifying sites as infected or noninfected according to the average wind speed from middle April until middle May ( $Wind_{Avg}[78,25]$ ) that is after first node detectable (growth stage [GS] 31) and the number of cold days (with minimum temperature below 7°C) from early January until early March ( $Tmin_{Nod;7}[190,70]$ ). Circles represent the observed noninfected years and sites and crosses represent the observed infected years and sites.

TABLE 5. The 11 “best” weather functions with their respective optimum windows (correlation greater than or equal to 0.7) and the statistic *N* (number of windows with a significant [ $P < 0.01$ ] correlation around the optimum one)<sup>a</sup>

Weather function	Time lag	Window length	Correlation	<i>P</i> value	<i>N</i>
Rain <sub>Avg</sub>	89	35	0.85	<0.001	11
Tmean <sub>Avg</sub>	67	47	0.74	0.001	1
Wet <sub>Avg</sub>	90	45	0.83	<0.001	15
Rain <sub>Cnod;0+</sub>	62	15	0.77	0.001	4
Rain <sub>Nod;1</sub>	87	42	0.88	<0.001	12
Rain <sub>Cnod;1</sub>	95	50	0.88	<0.001	13
Tmin <sub>Cnod;7+</sub>	145	70	0.72	0.003	6
Wet <sub>Cnod;80</sub>	99	54	0.78	0.001	15
Wet <sub>Cnod;20</sub>	162	13	0.74	0.002	1
Rain <sub>Nod;9</sub>	90	37	0.93	<0.001	12
Rain <sub>Nod;0.1</sub>	65	63	0.80	<0.001	3

<sup>a</sup> Data were restricted to sites and years with observed severity greater than or equal to 5%.



where standard errors are in parentheses and time lag and window length are in square brackets, respectively. In this case, the corresponding growth stages were tillering until early stem elongation (approximately GS 20/35). The  $N$  statistic for  $\text{Rain}_{\text{Nod};9}$  in equation 6 was 12. This was significant at a 10% level for all the bootstrap methods and was significant at a 5% level with all except the most stringent bootstrap (Fig. 5A).

### DISCUSSION

**Window Pane and the bootstrap.** Although Rispe et al. (34) mentioned the problem of spurious correlations and used Goldwin's correlogram to try and solve it, previous studies generally have relied entirely on individual  $P$  values to determine the significance and validity of Window Pane results. Generally, the main criticism made of analyses using this method concerns the large number of possible relationships screened. It has been argued (36) that because of the iterative process used to look for correlations, some would always be found significantly large, even in the absence of genuine relationships.

First and foremost, it has to be emphasized that the analysis and interpretation of the results obtained from Window Pane are not straightforward and clear-cut. Figures 3, 4, and 5 clearly show that significant correlations (at a 1% level) are extremely likely to occur. This however does not constitute proof that the Window Pane methodology is invalid but rather emphasizes the limitations of analyses that only use  $P$  values appropriate for testing single, pre-determined relationships. Instead, a statistic was necessary to test and distinguish relationships likely to be spurious from those likely to be genuine in the large set of possible relationships in the data. On the one hand, because *Septoria tritici* infection events happen over a long duration, it is likely that the relationship between a weather variable and disease severity at host maturity is representative of infection events that happen over a long period of time. While a few consecutive time lags with high correlations may be expected in the case of spurious relationships because of the autocorrelation of many weather variables, they will not be as long as those for genuine relationships. This distinction made us choose the number of significant correlations or misclassifications (at a 1% level) in time lags around the optimum window (with a 5-day step) as a test statistic. This statistic is only one of the many that could have been chosen and its usefulness is to some extent justified by the bootstrap results. However, it obviously predisposes to finding relationships operating over extended periods and

its use will tend to make it hard to identify brief specific periods in which key biological processes might occur.

**The bootstrap, a nonunique technique.** As described previously, four different methods of bootstrapping were used in this study. We shall now discuss the advantages and disadvantages of each of these methods.

**Keeping and dropping spatial correlation.** It is important to account for spatial correlation (using weather data from the same sites as the associated disease data) when meteorological factors are not the only ones to affect disease severity. For example, differences in soil type are important to crop physiology, which can have an indirect effect on disease progress. However, by keeping spatial correlation, we are more likely to dismiss genuine correlation because disease severity may be correlated with site across years. To illustrate this, let us assume that a disease is entirely controlled by temperature. Because of the north/south gradient in temperature in the United Kingdom, bootstrapping by keeping spatial correlation would consist of running Window Pane several times on very similar data sets, giving each time very similar outputs; the relationship found between temperature and disease would be rejected as spurious. To summarize these first two possible ways to bootstrap, keeping spatial correlation reduces type I error, i.e.,  $H_0$  (there is no relationship between weather and disease) is rejected when it is true. A type I error would lead us to conclude that there is a relationship between the weather factors and the disease severity when such a relationship does not exist. In contrast, dropping spatial correlation reduces type II error, i.e.,  $H_0$  is accepted when it is false. A type II error would make us reject the existence of a genuine relationship between weather factors and disease severity.

**Full and partial bootstrap.** When including as many weather variables as possible (full bootstrap), the main disadvantage is that differences between those variables (e.g., in terms of autocorrelation, temperature is more autocorrelated than rain) are ignored even though different values for judging the significance of the  $N$  statistic might be appropriate. It is also clear (Figs. 3 to 5) that the more weather variables that are included in the bootstrap analysis, the larger the  $N$  statistic. Read as such, this point could easily be seen as supporting the assertion that adding enough weather functions would lead to significant results. This would in fact be true only if adding new independent weather functions. In practice, this will never be the case because the number of weather variables available is limited and weather functions will frequently turn out to be correlated with one another, leading to similar results from Window Pane analyses (for instance,  $\text{Tmin}_{\text{Nod};7}$  and  $\text{Tmin}_{\text{Cnod};7}$  in the previous section). The choice of how many variables to bootstrap depends on the balance found acceptable between types of error. A full bootstrap emphasizes reduced type I error at the expense of type II, whereas a partial bootstrap does the opposite. Figure 9 summarizes the four possible bootstrap methods used. There is a hierarchy between these four ways of bootstrapping evident by the consistent ordering of the 5 or 10%

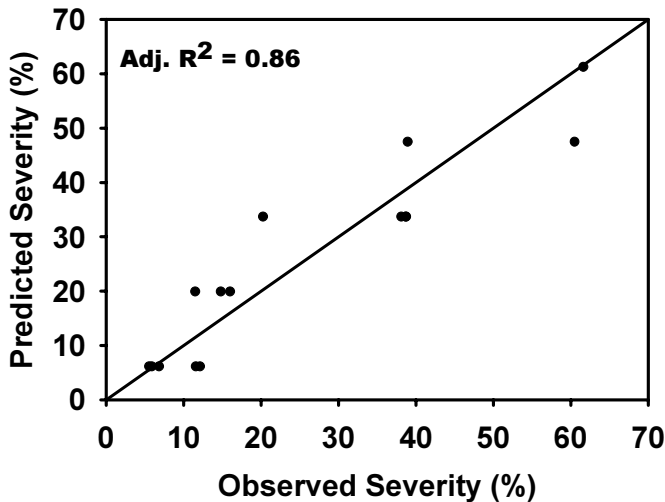


Fig. 8. Severity of *Septoria tritici* blotch measured at the mid-milk growth stage (GS 75) and predicted using the "best" multiple linear regression over sites and years where the observed severity was greater than or equal to 5%.

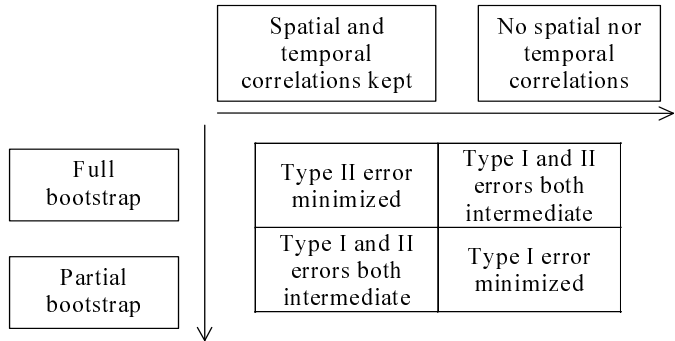


Fig. 9. Summary of the bootstrap analyses done with their respective effect on type I and II errors.

significance level across them (Figs. 3 to 5). The choice of the bootstrap method therefore depends on what one wants to achieve and is consequently related to the type of data that are analyzed.

The Window Pane technique proved very useful to derive predictive models, both qualitative and quantitative, for *Septoria tritici* on winter wheat. Further, we have shown that a bootstrap analysis can establish a context to help distinguish likely spurious from likely genuine correlations or misclassifications.

**Window Pane and field data.** Three of the four bootstrap analyses suggested that the most general model easily could have occurred by chance. The performance of the qualitative and restricted quantitative models stood out more clearly from chance. Field observations of disease are highly variable and severity scoring may be inaccurate, with large differences between observers (31). This will contribute to noise in relationships with weather deduced from the data and therefore increase type I error. It may be more sensible to accept the possibility of dubious relationships but not discard any relationship likely to explain the variability in disease observed in the field. Accepting only the relationships that are strong enough to be absolutely reliable (and which would therefore only rarely happen in field experiments) could prove too extreme a solution.

**Measuring the epidemic.** Before starting the Window Pane analysis, it was necessary to decide which leaf layers to include for the assessment of disease severity and at what growth stage to consider severity final. Although Shaw and Royle (37) found that disease severity on leaf three had no significant influence on yield loss, other experiments (13,39,42) found a significant positive correlation between the percent leaf area affected by *Septoria tritici* on all three upper leaf layers and the percent yield loss. Because an aim of this study was to derive a practical model for predicting disease severity on winter wheat in order to minimize yield loss, final disease severity was therefore measured across the top three leaves. Also, disease assessments become less reliable for leaf senescence over 30 to 40% (31). Therefore, including more leaf layers to assess severity of *Septoria tritici* (9,17) would increase the effect of problems in distinguishing senescence associated with the pathogen from natural senescence.

The choice of growth stage to measure severity has been widely discussed in the literature. Eyal and Ziv (13) compared correlations between percent yield loss and percent severity at three growth stages (GS 73/75, 83, and 87) and fitted the best linear regression at GS 73/75 (caryopsis at three quarter of its final stage). King et al. (22) also favored assessments at GS 75 (milk development to medium milk) because dry matter accumulation by grains peaks at this time. Disease severity at later growth stages is strongly correlated to severity during grain fill. There is therefore no physiologically specific growth stage. However, since we use the top three leaf layers (yield forming leaves) to summarize the epidemic, it is important to choose a particular growth stage during grain filling (e.g., GS 75) when reliable measurements can still be made, and severity at this stage is in practice well correlated with integrated measures of disease (37,39).

***Septoria tritici* epidemic as a two-step process.** The results of this study indicate the merit of a two-step process in the prediction of *Septoria tritici* severity. It is interesting to consider whether this has a biological basis. The use of a binary variable to classify disease severity in two categories (epidemic/no epidemic) suggested that preconditions for severe epidemics are a warm winter (January until early March; GS 11/25) and a calm spring (middle April until middle May; GS 25/35). At sites where a severe epidemic was recorded, the actual extent of severity was then best predicted by the frequency of heavy rainfalls (more than 9 mm a day) between April and middle May (GS 20/35).

**Temperature.** Infection and development of *Septoria tritici* are much slower at low temperatures (33). Coakley et al. (9), using 12 years of data, found that the frequency of consecutive days with temperature below 7°C at the end of tillering and start of stem

elongation was negatively correlated with *Septoria tritici* severity. More recently, Parker et al. (30) showed, using a more limited data set, a negative correlation between *Septoria tritici* severity on leaf layer two and the frequency of days with temperature below -2°C from early November until middle December. Parker et al. (30) also found a negative, but less significant, correlation with the frequency of days with temperature below 7°C. In a study based on a slightly different approach, Gladders et al. (15) found that frequent days with temperatures below -2°C in November reduced disease the following summer but their analysis did not include the 7°C threshold. Although our analysis did include three temperature thresholds (-2°C, 0, and 7°C), the greatest negative correlations were obtained for 7°C. It is therefore possible that, for the development of the winter epidemic, infection conditions (7°C) are more important than the base temperature of the pathogen development (-2°C). A recent study (29) suggests that ascospores may be important only in initiating the epidemic. Inoculum build-up in the crop prior to GS 31 would then be greatly dependent on the interaction between the crop and the pathogen as mediated by temperature. This part of the epidemic has previously been ignored and requires further attention.

**Wind.** Unlike temperature and rain, wind has not previously been shown to have a significant effect on epidemics of *Septoria tritici*. Equation 5 shows a negative correlation between wind and the presence of the epidemic. This negative correlation immediately rules out the possibility that wind increases disease development through dispersal of spores as leaves rub against one another. It may, in fact, be due to the influence of wind on leaf wetness. Although accurate predictions of dew duration use energy balance equations and depend on many variables (23,32), some multiple linear models using simple meteorological variables have been used in the past. For instance, Crowe et al. (10) derived a predictive model for leaf wetness duration and showed a negative correlation between leaf wetness duration and wind speed. Similarly, although wind speed did not appear in their simplest models, Djurle et al. (11) found a positive correlation between the number of hours with low wind and leaf wetness duration. On this basis, however, it is surprising that leaf wetness measured on-site was not shown to be a significant variable in the binary model. Two explanations may be suggested. First, leaf wetness measurements were among the least complete series and gaps had to be filled in using the neural network technique. In fact, none of the functions derived from any of the three weather variables with many values interpolated by neural networks are present in either the qualitative and quantitative models. Second, and more interestingly, measurements made by the sensors may not reflect the wetness conditions the pathogen requires to infect and develop; wind may actually reflect the internal water conditions of the plant better. For instance, in totally still air with 50% relative humidity above the canopy, conditions may be wetter inside stomata and on the leaf surface than in 70% relative humidity with high wind.

**Rain.** Previous work in temperate climates consistently has shown that after GS 31, disease development is driven by rain rather than by temperature (17,38,40). Coakley et al. (9) showed a negative correlation between severity of *Septoria tritici* and the frequency of consecutive days without rain between late March and early May, but this has not been confirmed by any further study. Gladders et al. (15) used high risk *Septoria* periods (39), which are a way of measuring important rain events derived from Tyldesley and Thompson (40), and found that their frequency in May and June was correlated with *Septoria tritici* severity on the upper two leaves at GS 73 to 75. Hansen et al. (17) pooled data from *Septoria tritici* and *Stagonospora nodorum* (*Septoria* spp.) and found a good correlation between *Septoria* severity around 15 July and the frequency of days with more than 1 mm of rain during a 1-month period starting at GS 32.

In summary, although our results confirm that *Septoria tritici* epidemics are dependent on rain throughout the growing season,

they also suggest the existence and importance of winter limits to the epidemic. As we have shown, the first part of the epidemic is driven by temperature, which may determine the number of disease cycles that occur in the winter, and therefore the level of disease before stem elongation (30). We also showed that, provided the right conditions for infection have been met during the winter, rain at the start of stem elongation is critical. Infection events around GS 30/31 (start of stem elongation) can cause young leaves to carry inoculum toward the final three leaf layers, which are crucial to grain yield.

## ACKNOWLEDGMENTS

This work was funded by the UK Department for Environment, Food and Rural Affairs. Rothamsted Research receives Grant-Aided support from the Biotechnology and Biological Sciences Research Council of the UK. We thank M. Taylor (Central Science Laboratory, York) for estimating missing weather data with neural networks.

## LITERATURE CITED

- Akaike, H. 1969. Fitting autoregressive models for prediction. *Ann. Inst. Stat. Math.* 21:243-247.
- Akaike, H. 1970. Statistical predictor identification. *Ann. Inst. Stat. Math.* 22:203-217.
- Belsley, D. A., Kuh, E., and Welsch, R. E. 1980. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. John Wiley & Sons, New York.
- Bryson, R. J., Paveley, N. D., Clark, W. S., Sylvester-Bradley, R., and Scott, R. K. 1997. Use of in-field measurements of green leaf area and incident radiation to estimate the effects of yellow rust epidemics on the yield of winter wheat. *Eur. J. Agron.* 7:53-62.
- Calvero, S. B., Coakley, S. M., and Teng, P. S. 1996. Development of empirical forecasting models for rice blast based on weather factors. *Plant Pathol.* 45:667-678.
- Chuang, T. Y., and Jeger, M. J. 1987. Predicting the rate of development of Black Sigatoka (*Mycosphaerella fijiensis* var. *difformis*) disease in southern Taiwan. *Phytopathology* 77:1542-1547.
- Coakley, S. M., Boyd, W. S., and Line, R. F. 1982. Statistical models for predicting stripe rust on winter wheat in the Pacific Northwest. *Phytopathology* 72:1539-1542.
- Coakley, S. M., and Line, R. F. 1982. Prediction of stripe rust epidemics on winter wheat using statistical models. *Phytopathology* 72:1006.
- Coakley, S. M., McDaniel, L. R., and Shaner, G. 1985. Model for predicting severity of *Septoria tritici* blotch on winter wheat. *Phytopathology* 75:1245-1251.
- Crowe, M. J., Coakley, S. M., and Emge, R. G. 1978. Forecasting dew duration at Pendleton, Oregon, using simple weather observations. *J. Appl. Meteorol.* 17:1482-1487.
- Djurle, A., Ekblom, B., and Yuen, J. E. 1996. The relationship of leaf wetness duration and disease progress of glume blotch, caused by *Stagonospora nodorum*, in winter wheat to standard weather data. *Eur. J. Plant Pathol.* 102:9-20.
- Draper, N. R., and Smith, H. 1966. *Applied Regression Analysis*. John Wiley, New York.
- Eyal, Z., and Ziv, O. 1974. The relationship between epidemics of *Septoria* leaf blotch and yield losses in spring wheat. *Phytopathology* 64:1385-1389.
- Francel, L. J., Madden, L. V., Rowe, R. C., and Riedel, R. M. 1990. Correlation of growing season environmental variables and the effect of early dying on potato yield. *Phytopathology* 80:425-432.
- Gladders, P., Paveley, N. D., Barrie, I. A., Hardwick, N. V., Hims, M. J., Langton, S., and Taylor, M. C. 2001. Agronomic and meteorological factors affecting the severity of leaf blotch caused by *Mycosphaerella graminicola* in commercial wheat crops in England. *Ann. Appl. Biol.* 138:301-311.
- Goldwin, G. K. 1982. A technique for studying the association between components of the weather and horticultural parameters. *Sci. Hortic.* 16:101-107.
- Hansen, J. G., Secher, B. J. M., Jorgensen, L. N., and Welling, B. 1994. Thresholds for control of *Septoria* spp. in winter wheat based on precipitation and growth stage. *Plant Pathol.* 43:183-189.
- Hardwick, N. V., Jones, D. R., and Slough, J. E. 2001. Factors affecting diseases of winter wheat in England and Wales, 1989-98. *Plant Pathol.* 50:453-462.
- Hardwick, N. V., Turner, J. A., Slough, J. E., Elcock, S. J., Jones, D. R., and Gladders, P. 2002. Oilseed rape and cereal diseases—How are farmers responding to their control? Pages 903-910 in: *Proc. Br. Crop Prot. Conf. Pests Dis.*, Vol. 2.
- Huber, L., Madden, L. V., and Fitt, B. D. L. 1998. Rain splash and spore dispersal: A physical perspective. Pages 348-370 in: *The Epidemiology of Plant Diseases*. J. D. Gareth, ed. Kluwer Publishers, Dordrecht, the Netherlands.
- Kanji, G. K. 1993. *100 Statistical Tests*. SAGE Publication, London.
- King, J. E., Cook, R. J., and Melville, S. C. 1983. A review of *Septoria* diseases of wheat and barley. *Ann. Appl. Biol.* 103:345-373.
- Luo, W. H., and Goudriaan, J. 2000. Dew formation on rice under varying durations of nocturnal radiative loss. *Agric. For. Meteorol.* 104:303-313.
- Mallows, C. L. 1973. Some comments on Cp. *Technometrics* 15:661-675.
- Mardia, K. V., Kent, J. T., and Bibby, J. M. 1979. *Multivariate Analysis*. Academic Press, London.
- Marquart, D. W. 1970. Generalized inverses, Ridge regression, biased linear estimation, and nonlinear estimation. *Technometrics* 12:591-612.
- Montgomery, D. C., and Peck, E. A. 1982. *Introduction to Linear Regression Analysis*. John Wiley & Sons, New York.
- Mora-Aguilera, G., Nieto-Angel, D., Teliz, D., and Campbell, C. L. 1993. Development of a prediction model for papaya ringspot in Veracruz, Mexico. *Plant Dis.* 77:1205-1211.
- Parker, S. R., and Lovell, D. J. 2001. Quantifying the benefits of seed treatment for foliar disease control. Pages 181-188 in: *Proc. Seed Treatment: Challenges & Opportunities*, 2001. British Crop Protection Council, Farnham, UK.
- Parker, S. R., Lovell, D. J., Royle, D. J., and Paveley, N. D. 1999. Analyzing epidemics of *Septoria tritici* for improved estimates of disease risk. Pages 96-107 in: *Septoria on Cereals: A Study of Pathosystems*. J. A. Lucas, P. Bowyerand, and H. M. Anderson, eds. CAB International, Oxford.
- Parker, S. R., Shaw, M. W., and Royle, D. J. 1995. The reliability of visual estimates of disease severity on cereal leaves. *Plant Pathol.* 44:856-864.
- Pedro, M. J., and Gillespie, T. J. 1982. Estimating dew duration. 1. Utilizing micrometeorological data. *Agric. Meteorol.* 25:283-296.
- Renfro, B. L., and Young, H. C. J. 1956. Techniques for studying varietal response to *Septoria* leaf blotch of wheat. *Phytopathology* 46:23-24.
- Rispe, C., Hulle, M., Gauthier, J. P., Pierre, J. S., and Harrington, R. 1998. Effect of climate on the proportion of males in the autumn flight of the aphid *Rhopalosiphum padi* L. (Hom., Aphididae). *J. Appl. Entomol.-Z. Angew. Entomol.* 122:129-136.
- Seck, M., Roelfs, A. P., and Teng, P. S. 1991. Influence of leaf position on yield loss caused by wheat leaf rust in single tillers. *Crop Prot.* 10:222-228.
- Shaw, M. W. 2002. Epidemic modelling and disease forecasting. Pages 252-265 in: *Plant Pathologists Pocketbook*. 3rd ed. J. M. Waller, J. M. Lenné, and S. J. Waller, eds. CAB International, Wallingford, UK.
- Shaw, M. W., and Royle, D. J. 1989. Estimation and validation of a function describing the rate at which *Mycosphaerella Graminicola* causes yield loss in winter wheat. *Ann. Appl. Biol.* 115:425-442.
- Shaw, M. W., and Royle, D. J. 1993. Factors determining the severity of epidemics of *Mycosphaerella graminicola* (*Septoria tritici*) on winter wheat in the UK. *Plant Pathol.* 42:882-899.
- Thomas, M. R., Cook, R. J., and King, J. E. 1989. Factors affecting development of *Septoria tritici* in winter wheat and its effect on yield. *Plant Pathol.* 38:246-257.
- Tyldesley, J. B., and Thompson, N. 1980. Forecasting *Septoria nodorum* on winter wheat in England and Wales. *Plant Pathol.* 29:9-20.
- Zadoks, J. C., Chang, T. T., and Konzak, C. F. 1974. A decimal code for the growth stages of cereals. *Weed Res.* 14:415-421.
- Ziv, O., and Eyal, Z. 1978. Assessment of yield component losses caused in plants of spring wheat cultivars by selected isolates of *Septoria tritici*. *Phytopathology* 68:791-796.