Statistical Models for Predicting Stripe Rust on Winter Wheat in the Pacific Northwest

Stella Melugin Coakley, William S. Boyd, and Roland F. Line

Associate research professor and research associate, Department of Biological Sciences, University of Denver (mailing address: National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307); and research plant pathologist, Agricultural Research Service, U.S. Department of Agriculture, Pullman, WA 99164, respectively.

This research was supported in part by the Climate Dynamics Program, Division of Atmospheric Sciences, National Science Foundation Grant ATM 80-17604.

Acknowledgment is made to the National Center for Atmospheric Research, which is supported by the National Science Foundation, for supplying computer time used in this research and for research space provided by the Advanced Study Program.

Accepted for publication 18 May 1982.

ABSTRACT

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.

Statistical models developed for predicting stripe rust (caused by *Puccinia striiformis*) on winter wheat cultivars Gaines, Nugaines, and Omar at Pullman, WA, also predicted disease intensity at four other locations (Lind, Mt. Vernon, and Walla Walla, WA; and Pendleton, OR) in the Pacific Northwest when the negative degree days data used to develop the models were first standardized (NDDZ). There was no significant difference between the measured mean disease intensity and the mean disease intensity predicted by the NDDZ model at any of the sites. Positive degree days accumulated at the sites partially explained the model's small, but consistent, underprediction at Mt. Vernon and overprediction at Lind, Walla Walla, and Pendleton. The results demonstrate that a model for predicting disease intensity by using macrometeorological conditions can be useful at other locations within a synoptic weather region.

Additional key words: quantitative epidemiology, linear regression, degree days.

Stripe rust of wheat (*Triticum aestivum* L. em Thell), caused by *Puccinia striiformis* West., has severely damaged susceptible wheat cultivars during 15 of the years since 1961 in the Pacific Northwest (PNW) region of the United States. Coakley and Line (1) reported the influence of climatic variation on the frequency and severity of stripe rust epidemics. They proposed a model for predicting stripe rust on the winter wheat cultivars Gaines and Omar based on either accumulated winter temperatures (negative degree days [NDD]) or accumulated spring temperatures (positive degree days [PDD]) at Pullman, WA. In Gaines, 73% of the variation in disease index was explained by NDD (1). In subsequent studies (2), disease predictions were improved by adding NDD and PDD and correcting for the growth stage at which disease observations were made. When the sum of NDD and PDD together with a growth index were used as independent variables in a multiple regression analysis, 91% of the variation in disease intensity was explained (2).

The proposed models (1,2) show the influence of meteorological conditions on stripe rust at Pullman, but these were not tested for application in other areas of the PNW.

The objective of this research was to develop and test methods of using the Pullman models at other sites in the PNW that had different local meteorological conditions although similar synoptic meteorology.

MATERIALS AND METHODS

Data on rust intensity, stage of plant growth, and maximum and minimum temperature at Lind, Mt. Vernon, Pullman, and Walla Walla, WA, and Pendleton, OR, were used. All five locations had different local meteorological conditions because of differences in geography and topography (Table 1).

Data on disease intensity and stage of growth for plants of cultivars Omar, Gaines, and Nugaines were collected at the five sites (Table 2). The disease data were collected at growth stages 7 (milk) or 8 (dough) (4). Data collected at Walla Walla during the 1977 growing season were not included in this study because of the influence of irrigation during that year.

Daily maximum and minimum temperatures for 1968–1979 were obtained from the National Climatic Center, Asheville, NC, for the five locations. NDD and PDD were calculated from the daily average Celsius temperatures ($\bar{T}$) by using a 7 C base as described (1).

$$\text{If } \bar{T} < 7\text{ C then NDD} = \Sigma\,|\,\bar{T} - 7\text{ C}\,|.$$
$$\text{If } \bar{T} > 7\text{ C then PDD} = \Sigma(\bar{T} - 7\text{ C}).$$

Differences in NDD and PDD at the five locations are summarized in Table 3.

Correlation coefficients were calculated between NDD and PDD values for Pullman and NDD and PDD values accumulated for various beginning dates and durations of time at the other locations. The time periods varied from 7 days to 91 days, but always in 7-day intervals. For example, for a time period of 28 days and a starting date of 1 October, NDD were accumulated from 1

TABLE 1. Location, elevation, and 30-yr normal[a] temperatures at wheat stripe rust locations in the Pacific Northwest of the United States

Location	Longitude (°N)	Latitude (°W)	Elevation (m)	Temperature (C)
Washington				
Pullman	46°46′	117°12′	775	8.6
Lind	47°00′	118°35′	497	9.9
Mt. Vernon	48°26′	122°23′	4	9.9
Walla Walla	46°02′	118°20′	289	12.2
Oregon				
Pendleton	45°41′	118°51′	455	11.3

[a]Climatological normals based on the period 1941–1970 except at Mt. Vernon, which was based on 1961–1976 data, since the data required for a 30-yr normal do not exist for that location.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. § 1734 solely to indicate this fact.

0031-949X/82/12153904/$03.00/0

TABLE 2. Stripe rust disease intensity index[a] by year and wheat cultivar with mean ($\bar{x}$) of disease data for Pullman (PW), Lind (LW), Mt. Vernon (MVW), and Walla Walla (WWW) in Washington[b] and Pendleton (PO) in Oregon[b]

	Gaines					Nugaines					Omar				
Year	PW	LW	MVW	WWW	PO	PW	LW	MVW	WWW	PO	PW	LW	MVW	WWW	PO
1968	5.5	···	···	···	···	5.0	···	···	···	···	···	···	···	···	···
1969	3.0	···	···	···	···	3.0	···	···	···	···	4.0	···	···	···	···
1970	4.0	···	···	···	3.0	3.0	···	···	···	3.5	6.0	···	···	···	7.0
1971	5.5	···	···	3.0	3.5	5.0	···	···	2.0	2.5	8.0	···	···	5.0	7.5
1972	4.0	···	···	···	···	3.0	···	···	···	···	5.0	···	···	···	···
1973	4.0	···	···	···	1.0	3.0	···	···	···	1.0	7.0	···	···	···	1.0
1974	4.0	···	···	4.0	···	3.0	···	···	3.0	···	5.0	···	···	6.0	···
1975	7.0	7.0	···	4.0	···	7.0	5.0	···	4.0	···	8.0	7.0	···	7.5	···
1976	7.0	3.0	7.0	6.0	5.0	7.0	3.5	6.0	5.0	3.0	8.0	7.0	8.5	···	6.0
1977	3.0	···	···	···	···	3.0	···	···	···	···	3.0	···	···	···	···
1978	6.0	4.0	···	···	···	4.0	3.0	···	···	···	9.0	4.0	···	···	···
1979	1.0	1.0	3.0	1.0	1.0	1.0	1.0	2.5	1.0	1.0	2.0	1.0	7.5	1.0	1.0
$\bar{x}$	4.5	3.8	5.0	3.6	2.7	3.9	3.1	4.3	3.0	2.2	5.9	4.8	8.0	4.9	4.5

[a]The disease intensity data collected at growth stages 7 and 8 were converted to a 0–9 scale disease intensity index (DI): 0 = 0% disease, 1 = <1%, 2 = 1–5%, 3 = 6–20%, 4 = 21–40%, 5 = 41–60%, 6 = 61–80%, 7 = 81–95%, 8 = 96–99%, and 9 = >99%. Missing data were not collected or were not available for the specified growth stages.

[b]The final letter in each acronym indicates the state in which the plantings were located.

TABLE 3. Means ($\bar{x}$) and standard deviations (s) of negative degree days (NDD) and positive degree days (PDD) for locations in the Pacific Northwest for 1968–1979. Average temperature ($\bar{T}$) is for time over which NDD and PDD accumulated, respectively

	NDD[a]			PDD[b]		
Location	$\bar{x}$	s	$\bar{T}$	$\bar{x}$	s	$\bar{T}$
Washington						
Pullman	554	132	−1.9	429	68	11.6
Lind	542	148	−1.7	607	73	13.4
Mt. Vernon	234	88	3.6	444	49	11.8
Walla Walla	375	142	1.5	784	75	15.6
Oregon						
Pendleton	409	166	0.8	713	61	14.7

[a]Accumulated from 1 December to 31 January.

[b]Accumulated from 1 April to 30 June.

TABLE 4. Time periods with highest correlation between negative degree days (NDD) at Pullman accumulated from 1 December to 31 January and NDD at each other station for 12 yr, 1968–1979

Location	Beginning date	Duration (days)	NDD	r[a]
Washington				
Pullman	1 Dec	62	552	···
Lind	6 Dec	56	501	0.9752
Mt. Vernon	15 Nov	77	264	0.8577
Walla Walla	6 Dec	70	416	0.9563
Oregon				
Pendleton	6 Dec	63	426	0.9571

[a]Correlation coefficients (r) were significant, $P < 0.001$.

October for 28 days, then from 8 October for 28 days, then from 15 October for 28 days, etc., until a given time period was completed.

Because the lengths of the most highly correlated time periods and, therefore, the magnitude of NDD and PDD, were different for each location, it was necessary to standardize the degree-day data for each location in order to use the data in the Pullman model. Standardized values (z) were calculated by subtracting the 1968–1979 mean ($\bar{x}$) for that location from each yearly value (x_i) and dividing by the standard deviation of the mean(s) as described by Davis (3):

$$z = (x_i - \bar{x})/s \qquad (1)$$

For example, at Pullman the 1968–1979 NDD mean ($\bar{x}$) was 554 with a standard deviation (s) of 132 (Table 3) and NDD for 1970 (x_i) was 481. The standardized value (z) of NDD for Pullman (NDDZ) in 1970 was calculated as −0.55. Standardized data for 1968–1979 at each location have a mean of zero and a standard deviation of one. Trends in NDDZ can be compared directly among locations where NDD were accumulated for different time periods.

The standardized NDDZ and PDDZ data at each station were used in the regression equations developed for Pullman, and disease predictions were made for the three cultivars for each year that disease data were available.

RESULTS AND DISCUSSION

Validation of the Pullman negative degree day model. NDD that accumulated from 1 December 1968 to 31 January 1979 at Pullman were most highly correlated with the NDD accumulated during 56, 63, 70, and 77 days at Lind, Pendleton, Walla Walla, and Mt. Vernon, respectively. The models developed at Pullman could not be directly applied at the other locations because of the differences in the time periods over which NDD were accumulated (Table 4). The NDD models at Pullman were reformulated by using the standardized NDDZ data for Pullman; this does not change the predictions made by the Pullman models (1). Regression equations expressing the relationship between stripe rust intensity index (Table 2) and NDDZ for cultivars Gaines, Nugaines, and Omar at Pullman were as follows with $\hat{y}$ = predicted disease intensity index and x = NDDZ.

$$\text{Gaines: } \hat{y} = 4.5000 - 1.5473x \quad r^2 = 0.7624 \qquad (2)$$

$$\text{Nugaines: } \hat{y} = 3.9167 - 1.3394x \quad r^2 = 0.5644 \qquad (3)$$

$$\text{Omar: } \hat{y} = 5.9091 - 1.8317x \quad r^2 = 0.6341 \qquad (4)$$

These regression equations (equations 2, 3, and 4) for Pullman were used with the standardized NDDZ data from each of the other sites to predict the disease intensities at those locations only for years that disease data were available to allow verification of the predictions. The results for Gaines and a summary of the results for Nugaines and Omar are shown in Table 5.

TABLE 5. Stripe rust disease intensity index (DI)[a], predicted disease intensity ($\hat{y}$), and residuals (DI–$\hat{y}$) for Gaines wheat by location[b] with the means ($\bar{x}$) for all cultivars

Year and cultivar	Lind			Mt. Vernon			Walla Walla			Pendleton		
	DI	$\hat{y}$	DI-$\hat{y}$	DI	$\hat{y}$	DI-$\hat{y}$	DI	$\hat{y}$	DI-$\hat{y}$	DI	$\hat{y}$	DI-$\hat{y}$
1968	...	...	...	...	...	...	...	...	...	...	...	...
1969	...	...	...	...	...	...	...	...	...	...	...	...
1970	...	...	...	...	...	...	...	...	...	3.0	5.2	−2.2
1971	...	...	...	...	...	...	3.0	6.2	−3.2	3.5	5.8	−2.3
1972	...	...	...	...	...	...	...	...	...	...	...	...
1973	...	...	...	...	...	...	...	...	...	1.0	3.5	−2.5
1974	...	...	...	...	...	...	4.0	5.3	−1.3	...	...	...
1975	7.0	5.5	1.5	...	...	...	4.0	5.7	−1.7	...	...	...
1976	3.0	5.6	−2.6	7.0	5.9	1.1	6.0	5.7	0.3	5.0	5.9	−0.9
1977	...	...	...	...	...	...	...	...	...	...	...	...
1978	4.0	5.3	−1.3	...	...	...	...	...	...	...	...	...
1979	1.0	0.6	0.4	3.0	2.0	1.0	1.0	1.1	−0.1	1.0	0.9	0.1
Means: ($\bar{x}$)												
Gaines	3.8	4.3	−0.5	5.0	4.0	1.1	3.6	4.8	−1.2	2.7	4.3	−1.6
Nugaines	3.1	3.7	−0.6	4.3	3.4	0.8	3.0	4.2	−1.2	2.2	3.7	−1.5
Omar	4.8	5.6	−0.9	8.0	5.3	2.7	4.9	6.0	−1.1	4.5	5.6	−1.1

[a] Disease intensity index is on a 0–9 scale (see Table 2 footnote); missing data were not available for the growth stages used to develop the Pullman models.
[b] Lind, Mt. Vernon, and Walla Walla, WA; and Pendleton, OR.

TABLE 6. Partial explanation of NDDZ residuals for wheat cultivar Gaines by regression analysis[a]

Locations	1968–1979 Mean of PDD accumulated 1 Apr to 30 Jun	1968–1979 mean DI	Difference of DI from Pullman	Predicted difference of DI from Pullman	NDDZ residuals
Washington					
Pullman	429	4.5	0.0	0.2	0.0
Lind	607	3.8	−0.7	−0.6	−0.5
Mt. Vernon	444	5.0	0.5	0.1	1.1
Walla Walla	784	3.6	−0.9	−1.5	−1.2
Oregon					
Pendleton	713	2.7	−1.8	−1.1	−1.6

[a] The analysis determined that $\hat{y} = 2.186 - 0.0046\,x$. ($r^2 = 0.70$), in which y = the differences in mean disease intensity index (DI) between Pullman and locations and x = the mean positive degree days (PDD) for 1968–1979 at each location.

t-Tests between the actual and predicted disease indices give a measure of how well the Pullman model predicted disease at the other locations. There was no significant difference between the mean disease index and the mean predicted disease index at any of the locations for any of the three cultivars.

Examining the residuals (deviations of data from the regression line) gives a better indication than the t-test of how well the model predicted disease (Table 5). Predictions were the worst for Walla Walla in 1971 and Lind in 1976 when actual disease was in the moderate range and the disease predicted was in the severe range (see Fig. 2 in reference 1). The model's average residuals for all years for Gaines at Mt. Vernon, Lind, Walla Walla, and Pendleton were 1.1, −0.5, −1.2, and −1.6, respectively, indicating underprediction at Mt. Vernon and overprediction of disease index at the other locations.

Relationship of positive degree days. We know that the higher the spring temperature (high PDD values), the lower the disease index (1,2). The Pullman PDD models were not applicable at Lind, Walla Walla, and Pendleton because PDD values at those sites were so much greater than those used to formulate the Pullman model. Mean PDD for 1968–1979 at Lind, Walla Walla, and Pendleton (Table 3) were significantly higher ($P < 0.001$) than the mean at Pullman. We hypothesized, therefore, that the overprediction of the Pullman NDDZ model (Table 5) at Lind, Walla Walla, and Pendleton was in part due to the accumulation of PDD at those locations. PDD values and mean disease index data for the five stations were used to test this hypothesis. A regression model relating PDD to average deviation of the disease index from Pullman was formulated:

$$\hat{y} = 2.186 - 0.0046x \qquad (5)$$
$$r^2 = 0.70$$

The independent variable (x) was the 1968–1979 mean of PDD accumulated from 1 April to 30 June at each location. The dependent variable (y) was the difference in the 1968–1979 mean disease index at Pullman from that at each of the locations. The predicted deviations ($\hat{y}$) corresponded closely to the residuals of the NDDZ model predictions (Table 6). The t-test revealed no significant differences between the mean residuals of the NDDZ predictions at the four locations and the predicted mean difference in disease at Pullman based on equation 5. Therefore, most of the disease severity not explained by the NDDZ models was explained by differences in average PDD due to location.

Biological implications of analysis. Standardizing NDD made the Pullman models portable, even though the mean NDD at the locations varied greatly (Table 3). This implies that low temperatures limit the disease down to a threshold temperature, but that temperatures lower than that have no further effect. To test this interpretation, correlations between disease index and frequency of days below a given threshold were determined. The time period and threshold temperature were varied to find the highest correlations. At Pullman, disease intensity was more highly correlated with the frequency of days (FD) less than −6 C (FD < −6 C) than with NDD (S. M. Coakley and W. S. Boyd, *unpublished*). The same trend was found at the other locations for which the threshold temperature varied ± 3 C.

Regression models based on FD < −6 C at Pullman were applied at the other locations. The models yielded similar results to the NDDZ models, but were more complex because the FD < −6 C was a count and had to be transformed twice to compare time periods of different lengths. Predictions based upon these models were not sufficiently better than the NDDZ models to compensate for their greater complexity.

Our results demonstrate that a model for predicting disease intensity using macrometeorological measurements can be useful at other locations within a synoptic weather region. The methodology for application of the Pullman models at other locations was appropriate for stripe rust in the PNW and could be useful for application of similar models developed for other diseases in other locations.

LITERATURE CITED

1. Coakley, S. M., and Line, R. F. 1981. Quantitative relationships between climatic variables and stripe rust epidemics on winter wheat. Phytopathology 71:461-467.
2. Coakley, S. M., and Line, R. F. 1981. Climatic variables that control development of stripe rust disease on winter wheat. Clim. Change 3:303-315.
3. Davis, J. C. 1973. Statistics and data analysis in geology. John Wiley & Sons, New York. 550 pp.
4. Zadoks, J. S., and Konzak, C. F. 1974. A decimal code for the growth stages of cereals. Eucarpia Bull. 7. 12 pp.