Quantifying Wheat Yield Across the Field as a Function of Wheat Streak Mosaic Intensity: A State Space Approach

F. Workneh, D. C. Jones, and C. M. Rush

Texas AgriLife Research, Amarillo 79106.
Accepted for publication 1 December 2008.

ABSTRACT

Workneh, F., Jones, D. C., and Rush, C. M. 2009. Quantifying wheat yield across the field as a function of wheat streak mosaic intensity: A state space approach. *Phytopathology* 99:432-440.

Wheat streak mosaic virus (WSMV), vectored by the wheat curl mite *Aceria tosichella*, is one of the major limiting factors in wheat production in the Texas Panhandle. The mites are blown by wind into wheat fields from nearby volunteer wheat fields or fields supporting vegetation which harbor virulent mites. Consequently, gradients of wheat streak severity are often observed stretching from the edges of wheat fields into the center or beyond. To describe the magnitude of the spatial relationships between grain yield and wheat streak intensity across the field, studies were conducted in 2006 and 2007 in three infected fields. Wheat streak severity was quantified with reflectance measurements (remote sensing) at 555-nm wave length using a hand-held radiometer. Measurements were taken in several equally spaced 1 m² locations along a total of eight transects and grain yield was assessed from a 0.8 m² area of each location. State space analysis was used to describe the relationships in which yield data and reflectance values were used as dependent and explanatory

variables, respectively. A structural time series model was formulated as a state space model where the unobserved components were modeled explicitly. In the analysis the state of yield at current location (d) was related to the state of wheat streak intensity either at current locations (d) or lagged locations with autoregressive values of the first order ($d-1$) or greater. There were significant cross-correlations between yield and wheat streak intensity up to distances of 150 m ($P \leq 0.05$). Grain yield at the current position was significantly correlated with reflectance values at the same location and/or at lagged locations. The spatial aspects of the yield-reflectance relationships were best described by state space models with stochastic trends without slopes or deterministic trends with or without slopes. The models correctly predicted almost all of the observed yield values as a function of wheat streak intensity across the field within the 95% confidence interval. Results obtained in this study suggest that state space methodology can be a powerful tool in the study of plant disease spread as a function of other variables.

Additional keywords: Kalman filter, observation equation, state equation, state vector.

The Texas Panhandle produces approximately 45 to 50% of the state's hard red winter wheat crop each year on approximately one million hectares (2). Wheat production in this region is dual purpose for grain and cattle grazing, and *Wheat streak mosaic virus* (WSMV) vectored by the wheat curl mite *Aceria tosichella* Keifer (40) is one of the major limiting factors affecting both forage and grain yield production. Fields destined for both grazing and grain, or grazing alone, are usually planted early (late August to early September) and wheat streak is generally more severe in early than in later plantings. The virus and the mite have multiple host ranges and overwinter on volunteer wheat and other grass vegetation (9,18) between wheat crops. Severe outbreaks in early planted wheat are partly attributed to overlapping of the wheat crop with summer green vegetation hosting mites and the virus (7). The disease is consistently prevalent in the region from year to year (47) but appears to be more severe during drought years than wet years and its effect is sometimes masked by drought conditions in dry land productions. In the 2005 and 2006 growing season, 98% of the 130 diseased samples submitted by growers to the Texas High Plains Plant Diagnostic Laboratory in Amarillo for identification tested positive for WSMV.

Even though there are germplasm lines that can confer genetic resistance to the pathogen (36), currently they are not commercially widely available. Plants infected with WSMV exhibit symp-

toms ranging from chlorotic streaking and mosaic to complete chlorosis of leaves and stunting of the plants. Disease development often begins at the edges of fields facing nearby volunteer wheat fields or grasslands which harbor the vector (43). As the season progresses a disease severity gradient often develops frequently affecting over half the field. Knowledge of the disease spread across the field may provide new insight into the epidemiology of the disease and how it affects yield. Our primary objectives were to (i) quantify the spatial relationships between wheat streak severity and grain yield across the field along the disease gradient, and (ii) determine whether the spatial relationships could be described with state space modeling. A brief overview of state space methodology as we related it to this study is presented.

MATERIALS AND METHODS

Data collection. In the spring of 2006 a 100 ha wheat field in Deaf Smith County, infected with WSMV, was studied using four transects each 720 m in length and all oriented in the same direction. Each transect consisted of 37 sampling locations which were 20 m apart. At each location the severity of wheat streak in 1 m² (5 rows in width) was quantified with reflectance measurements (remote sensing) using a hyperspectral, hand-held radiometer at 555-nm wave length at Feekes 7 to 8 growth-stage scale (25). In a preliminary study, measurements taken at this reflectance band showed a high correlation with visual WSMV severity assessments. In addition, disease severity was visually assessed using a 1 to 4 scale where 1 is healthy green leaves and 4 is leaves completely chlorotic and plants stunted. Diseased leaf samples

Corresponding author: F. Workneh; E-mail address: f-workneh@tamu.edu

doi:10.1094/PHYTO-99-4-0432

© 2009 The American Phytopathological Society

from each location were collected and tested by double-antibody sandwich enzyme-linked immunosorbent assay (DAS-ELISA) to verify that the symptoms were caused by WSMV. The first row of plants from each location was harvested at heading for biomass determination (not used in this particular study) and grain yield was determined from the remaining four rows (0.8 m²).

The study was repeated in 2007 in two fields, one in Lamb County (approximately 24.3 ha) and the other in Deaf Smith County (approximately 13.8 ha). In each field the study was conducted using two transects oriented in the same direction. Transects in the field in Lamb County comprised of 60 locations each at 2.5 m intervals and the other field had 50 locations at 10 m intervals. The lengths of the intervals among the locations were mainly dictated by the sizes of the areas affected. The 2007 field in Deaf Smith County was a dry land field (rain fed), while the other two were under center-pivot irrigation systems.

State space modeling. State space is a class of time series which can be represented either as a univariate or a multivariate form (8,11). Since its original introduction by Kalman (23) in the field of engineering, the state space methodology has been extended in use to many areas of research (6,16,26,30,31,37, 38,42,45,46). The basic state space form, known as the state equation, is expressed as

$$Z_d = \Phi Z_{d-1} + \varepsilon_d \quad (1)$$

where Z_d is unobserved state vector, and Φ is the state coefficient (also known as transition matrix), and ε_d is a white noise disturbance assumed to be normally and independently distributed with zero mean and variance σ^2 . The above expression can also be formulated as $Z_{d+1} = \Phi Z_d + \varepsilon_d$. The state space model comprises another component known as the observation equation,

$$Y_d = M_d Z_d + \eta_d \quad (2)$$

in which the unobserved true state vector (Z_d) is embedded, where Y_d is the measurement vector related to the state vector through the coefficient M_d and η_d is the measurement noise assumed to be independent of ε_d . The current estimate or the predicted state vector (Z_d^{pr}) is updated (Z_d^{up}) as Z_d moves through space or time (and new observations are brought in) using the error ($Y_d - M_d Z_d^{pr}$) associated with predicting Y_d . The updating is accomplished using the Kalman filter as

$$Z_d^{up} = Z_d^{pr} + K_d(Y_d - M_d Z_d^{pr}) \quad (3)$$

where K_d stands for the Kalman gain (39). The Kalman filter and details of its processes in filtering and smoothing during updating is a broad subject by itself (13,15,41) and, therefore, its description is beyond the scope of this article.

The classical time series methodology, as formulated by Box and Jenkins (3), requires the series to be stationary (the means and variances have to be constant over time or space), which means that time series data which consist of unobserved components such as trends, seasonal effects or cycles have to be detrended or differenced before analysis. This requirement has been regarded as a limiting factor in inclusion of unobserved components of interest in time series analysis including state space (10,11) and has been implicated as one of the primary causes of convergence failures in state space analysis (1,4). However, nonstationary series have been used in state space modeling (30) and it has been suggested that series stationarity is not a prerequisite in state space analysis (32). As indicated earlier, wheat streak severity often declines from the edges of the field (where infection first begins) towards the center or the opposite end, depending on the magnitude of the spread. Consequently, there were obvious trends (nonstationary) in all of our data, and

analysis of trends in disease intensity across the field was one of our primary interests. Thus, rather than differencing, we opted to retain the trends and treat the data as a structural time series in which the trend component can be decomposed and explicitly modeled (15,19).

In structural time series framework, a model may be expressed as a sum total of different components as shown below,

$$Y_d = \mu_d + \gamma_d + X_d + \varepsilon_d \quad (4)$$

where Y_d is the observed series of interest at position d , μ_d is a trend, γ_d is a seasonal effect, X_d is an explanatory variable, and ε_d is the irregular component (15,19). The trend consists of *level* and *slope* and may be thought of as a moving intercept (allowed to change over time) but the slope may or may not be present (8,10). The series components can be formulated in a state space format either as a stochastic or deterministic form. For instance, the trend components, level (μ_d) and slope (v_d), and a regression effect (β_d) for an explanatory variable (X_d) can be specified in the observation equation in a stochastic form as in equation 5,

$$\begin{aligned} Y_d &= \mu_d + \beta_d X_d + \eta_d \\ \mu_d &= \mu_{d-1} + v_d + \omega_d \\ v_d &= v_{d-1} + \xi_d \\ \beta_d &= \beta_{d-1} + \zeta_d \end{aligned} \quad (5)$$

where η_d , ω_d , ξ_d , and ζ_d are assumed to be uncorrelated white noise disturbances with zero means and unit variances (8,10). In a deterministic formulation, the disturbance terms are fixed to zero. The components can be similarly specified in the state equation. However, the outcome can be different for the different forms of specifications (8).

Analyses. To determine the spatial relationships between wheat streak and grain yield, cross correlation analyses were conducted for each transect using SAS software (SAS Institute Inc., Cary, NC). Structural time series analysis in a state space form was conducted using STAMP 8 software (structural time series analyzer, modeler, and predictor; Timberlake Consultants, Union, NJ). Yield data were square-root transformed before analysis to stabilize the variances. The series trend components (level and slope) and the explanatory variable (wheat streak severity as quantified by reflectance measurements) were specified in the state space model as deterministic or stochastic form in the observation equation or in the state equation. Model selection was based on residual diagnostic tests and goodness of fit statistics.

State space models must satisfy the underlying assumptions that the disturbances (irregulars) are serially independent and normally distributed (8). In addition, the variances have to be constant across the length of the series. The basic yardstick for measuring the goodness of fit in structural time series modeling is prediction error variance (PEV). For a model to be acceptable, the ratio of prediction error variance and the square of the error mean deviation (MD) should approximate unity (8,15). The traditional coefficient of determination (R^2) does not play an important role in time series unless the series is stationary because relatively high values can be achieved simply from the existence of trends in the series. A more restrictive measure (for data with trends) is an R^2 based on comparison of PEV and the variance of the first difference (R^2_D) (8,15). Another important goodness of fit measure in selecting among models is the Akaike information criteria (AIC) for which a smaller value is an indication of a better fit. The AIC is important for two reasons. First, it follows an objective comparison of nonnested models, and secondly, it rewards parsimony in model fitting so that all other things being equal it always gives precedence to simpler models over more complex ones.

Fig. 1. Grain yield and wheat streak intensity levels (as measured by reflectance at 555 nm) in relation to distance from the edge of the field in four transects (A, B, C, and D) of the field sampled in 2006.

Fig. 2. Grain yield and wheat streak intensity levels (measured by reflectance at 555 nm) in relation to distances from the edge of the fields in the first (transects A and B) and 2nd (transects C and D) fields sampled in 2007.

RESULTS

Almost all of wheat samples collected during the 2-year period from locations along the transects tested positive for WSMV with the percentage of positive samples varying among transects between 90 and 100%. Correlations between ranks of the visual disease scores for individual locations and reflectance values ranged between 0.66 and 0.90 ($P < 0.01$) for the 2006 data and between 0.76 and 0.92 ($P < 0.01$) for the 2007 data. In both years reflectance values decreased with increasing distances from the edges of the fields while grain yield values increased with distance (Figs. 1 and 2). Trends in reflectance and yield values were similar among transects within a field but appeared different among fields. In cross-correlation analysis, reflectance values were strongly correlated with grain yield at zero lag (which is analogous to the classical correlation), with coefficients ranging from -0.78 to -0.90 in 2006 and from -0.72 to -0.91 in 2007 ($P < 0.05$) for both fields and all transects (Fig. 3A to H). For the 2006 field, grain yield and reflectance values significantly ($P \leq 0.05$) spatially cross-correlated up to distances ranging from 40 to 80 m in one direction (positive) and up to 40 to 120 m in the other direction (negative) (Fig. 3A to D). For the 2007 fields significant ($P \leq 0.05$) relationships were observed up to 40 to 120 m in one direction and 20 to 90 m in the other direction (Fig. 3E to H). The cross correlations appeared to be somewhat less symmetrical in some fields and transects than in others, especially for the 2007 samples.

Across all transects and fields reflectance values ranged between 1 and 18 nm from healthy to severely infected area of the field, respectively, while yield values in 0.8 m^2 locations ranged from 0.2 to 325 g, reversely. For all transects the square root transformation put the yield data within similar distribution ranges of the reflectance values, which resulted in a better normalization of the data than any other attempted transformations. Histograms (overlaid with normal density curves), QQ plots (quantile-quantile plots for standardized residuals), and correlograms (depicting autocorrelation functions) of the series for different models are presented in Figure 4. The residual correlograms show no organized serial correlations with most values within the 95% confidence interval. In the QQ plots, observed values (dots) approximated the normal density scores (lines). Overall, the residual plots exhibited no obvious deviations that can be considered significant departures from assumptions stated earlier. Thus, to avoid redundancy, since data from all transects showed satisfactory diagnostic features in terms of the underlying assumptions, only results from the first transect from each field are presented (Fig. 4A to C, D to F, and G to I for 2006 and for 2007 first and second fields, respectively).

State space modeling. Based on whether the models were stochastic or deterministic in terms of level and slope specifications, and the absence or presence of the slope, six models for each transect were compared using the goodness of fit statistics described earlier and the best model was selected for each transect. For the selected models the PEV/MD² ratios for all transects were close to unity (Table 1), suggesting satisfactory goodness of fit. The R^2 values (ranging from 0.66 to 0.94) also were generally high. However, values of the more restrictive parameter, R^2_D , were overall low ranging from 0.23 to 0.57. Across all transects the selected models had AIC values of ≤ 1 .

During model specifications, if the slopes were not significant, they were removed from the models and considered absent and the candidate models then would contain only the level aspect of the trend. For the 2006 data for all four transects the best models were models with stochastic levels without slopes (Table 1). However, transect data from each of the 2007 fields did not behave as uniformly as those from the 2006 field in terms of the outcome of model specifications. For the first field and the first transect the best model was a model with stochastic level containing a non-

significant slope as in the 2006 data. In contrast, data from the second transect were better explained by a model with deterministic level and deterministic slope. The two transects in the second field also behaved differently from each other and from the other transects as well. For this second field, data from the first transect were best described by a model with a stochastic level and a deterministic slope, while a model with a deterministic level, but without a slope, best fitted the second transect.

As expected, time series data from all transects for both years showed highly significant trends, with level values ranging from 14.26 to 22.04 ($P < 0.0001$; Table 2) but the slopes of the trends were significant only for data from two of the transects of the 2007 fields. The second transect in the first field and the first transect in the second had significant slope (both deterministic) values of 0.14 and 0.22, respectively ($P < 0.05$; Table 2). For the 2006 data, current values (X_d) of the explanatory variable, wheat streak (reflectance), were highly significantly related to yield at the same position ($P < 0.01$). However, none of the lagged values (X_{d-1}) of wheat streak were significantly related to yield at position d (Table 3). In the 2007 data, in the first field, in both transects, reflectance values at position $d-1$ were significantly related to yield at position d ($P < 0.05$). However, unlike the first field, the two transects in the second field had mixed results. In the first transect, only current values of wheat streak were significantly related to yield at position d , while in the second, grain yield at position d was significantly related to wheat streak intensity at both positions of d and $d-1$ (Table 3). The general stochastic yield model (from Tables 2 and 3) as a function of wheat streak intensity then can be expressed as yield at position

Fig. 3. Cross-correlations between grain yield and wheat streak intensity (as measured by reflectance at 555 nm) in four transects (A, B, C, and D) sampled in 2006, and for the first (transects E and F) and second (transects G and H) fields sampled in 2007. Dotted lines indicate significance level ($\alpha = 0.05$).

d = level – wheat streak intensity at position d and/or $d-1$ + irregular (η).

Results of the grain yield predictions using the state space models are presented in Figures 5 to 8. Overall a great majority of the data points fell within the 95% confidence interval of the predicted values indicating a high level predictive ability of the models. This is especially true of the first and the fourth transects in which nearly 100% of the data points were within the 95% confidence interval for the field sampled in 2006 (Figs. 5A and 6B). The predictive ability of the models was much more evident in the residuals corresponding to each prediction graph where the prediction errors were within ± 2 standard errors around zero. However, an exception to the high predictive ability of the models was depicted in the second transect of the first field in the 2007 samples in which quite a number data points fell outside of the confidence interval (towards the end of the transect) indicative of a less accurate prediction (Fig. 7B). This was also evident in the corresponding residual plot where there were some obvious deviations from zero (Fig. 7D).

DISCUSSION

Wheat streak is a major production problem in the Texas Panhandle especially in wheat planted in proximity of fields with volunteer wheat or other grass vegetation which harbor the vector. Two of the fields studied in this project were adjacent to fallow fields, cropped with wheat the previous year and had volunteer wheat with obvious wheat streak symptoms, while the third one was located adjacent to conservation reserve program grass land. Over 90% of the samples collected from the sampling locations in these fields tested positive for WSMV. In all of the fields, wheat streak intensity declined with increasing distance from the edges of the fields while grain yield increased. The strong negative correlations exhibited at zero-lag indicate that the observed relationship between wheat streak and grain yield was not due to some underlying factors (unaccounted for) related to increase in distance from the edge but due primarily to the fundamental negative relationship between the two. The strength and the extent of cross-correlations between the two suggested that state space

Fig. 4. Histograms plus estimated densities (normality — and t test ----), quantile-quantile (QQ) plots, and correlograms of prediction errors, respectively, for samples collected in 2006 (A, B, and C, respectively) and for the first (D, E, and F, respectively) and second (G, H, and I, respectively) fields sampled in 2007.

TABLE 1. Goodness of fit statistics for wheat streak series state space models in fields sampled in 2006 and 2007

Year	Field	Transect	Level/slope	PEV ^a /MD ²	R ²	R ² _D ^b	AIC ^c
2006	1	1	Stochastic/absent	1.05	0.85	0.49	0.42
		2	Stochastic/absent	1.11	0.88	0.53	-0.01
		3	Stochastic/absent	1.12	0.74	0.57	0.59
		4	Stochastic/absent	1.13	0.66	0.48	0.61
2007	1	1	Stochastic/absent	1.09	0.94	0.23	0.54
		2	Deterministic/deterministic	1.06	0.90	0.31	1.00
	2	1	Stochastic/deterministic	1.00	0.90	0.57	0.83
		2	Deterministic/absent	0.98	0.90	0.37	0.90

^a PEV/MD² stands for ratio of prediction error variance to the square of mean deviation.

^b R²_D is the value of PEV relative to the variance of the first difference.

^c Akaike information criteria.

analysis could be used to describe their relationships across the field using either the current reflectance values (X_d) or autoregressive values of the first order (X_{d-1}) or greater.

The use of the state space model provided a good description of trends in grain yield across the field as a function of wheat streak

intensity. In most of the transect grain yield gradients as affected by wheat streak, were best described with state space models with highly significant stochastic trends in which the slopes were absent. Specifications of the explanatory variable in the observation equation (equation 5) as a fixed effect provided the best fit

TABLE 2. Estimates of trend components (level and slope) and their significance levels for selected state space models for samples collected in 2006 and 2007

Year	Field	Transect	Level	Slope	<i>t</i> value	<i>P</i>
2006	1	1	15.19	Absent	19.96	<0.00001
		2	14.26	Absent	23.04	<0.00001
		3	17.10	Absent	11.51	<0.00001
		4	18.26	Absent	18.578	<0.00001
2007	1	1	17.44	Absent	17.21	<0.00001
		2	18.56	0.22	25.43/10.48 ^a	<0.00001/0.00001
	2	1	21.44	0.14	14.84/2.86	<0.00001/0.00624
		2	22.04	Absent	31.04	<0.00001

^a Numbers following the backslash are values for the slope.

TABLE 3. Estimates of coefficients of wheat streak intensities (reflectance) at current (d) and lagged ($d-1$) locations along transects and their significance levels for selected state space models in the 2006 and 2007 sampled fields

Year	Field	Transect	WSM coefficient ^a		<i>t</i> value		<i>P</i>	
			X_d	X_{d-1}	X_d	X_{d-1}	X_d	X_{d-1}
2006	1	1	-0.81	NS ^b	-4.80	NS	0.00003	NS
		2	-1.23	NS	-9.65	NS	0.00000	NS
		3	-0.63	NS	-3.59	NS	0.00107	NS
		4	-0.62	NS	-6.45	NS	0.00000	NS
2007	1	1	NS	-0.38	NS	-2.17	NS	0.03452
		2	NS	-0.40	NS	-2.55	NS	0.01343
	2	1	-0.98	NS	-4.53	NS	0.00004	NS
		2	-0.73	-0.74	-3.24	-3.33	0.00220	0.00173

^a WSM indicates wheat streak mosaic.

^b NS stands for not significant.

Fig. 5. Observed and predicted grain yield (g) modeled as a function of wheat streak intensity (reflectance at 555 nm) across the field in the first and second transects (A and B, respectively) and their respective residuals (C and D) sampled in 2006. Dotted lines in C and D represent ± 2 standard deviations.

Fig. 6. Observed and predicted grain yield (g) modeled as a function of wheat streak intensity (reflectance at 555 nm) across the field in the third and fourth transects (**A** and **B**, respectively) and their respective residuals (**C** and **D**) sampled in 2006. Dotted lines in **C** and **D** represent ± 2 standard deviations.

Fig. 7. Observed and predicted grain yield (g) modeled as a function of wheat streak intensity (reflectance at 555 nm) across the first field in the first and second transects (**A** and **B**, respectively) and their respective residuals (**C** and **D**) sampled in 2007. Dotted lines in **C** and **D** represent ± 2 standard deviations.

in all transects but specifications in the state equation option uniformly showed no effect across all transects. The state space output results were uniform among transects in two of the three fields studied in that wheat streak intensity either at position d or $d-1$ was significantly associated with grain yield at position d . However, observed variations in trends and coefficients of the explanatory variable among fields and transects may be attributed to differences in intensity of the wheat streak. Differences in time of infection and densities of the incoming mite populations may have played a significant role in differentially affecting the trends in the spatial aspects of the disease spread across the fields. In addition, some of the data points outside of the 95% confidence interval were pointed out by the software as definite outliers and intervention analysis (where one can model outliers separately from the rest of the data) showed significant improvement in the models' fit in absence of the outliers (data not shown). Although there were sporadic weed infestations and patches of bare ground across all transects around the sampling locations that might have served as background contaminants to the wheat streak reflectance readings, we did not find sufficient grounds that would warrant intervention analysis.

Formulation of structural time series as a state space model enabled us to model the trends separately without the need for differencing. A preliminary attempt at data analysis, using either differenced or nondifferenced data, was less successful in model fitting, probably because of the strong trends in the data (data not shown). This may have resulted because of the fact that differencing may not entirely eliminate the trend components and, consequently, the expected level of stationarity may not have been achieved (15).

Even though state space modeling has its inception in aerospace engineering, its use has been widely accepted in a variety of

research fields including econometrics, agronomy and soil sciences, and water sciences (6,16,26,30,31,37,38,42,45,46). The model's ability to update future predictions based on current prediction errors, using recursive filtering, has made its applications very attractive for use in multidisciplinary research. In addition, unlike the univariate classical time series models (e.g., ARIMA models), state space models can be expressed naturally in multivariate forms in which explanatory variables can be incorporated as factors influencing the series trends and evolution over time or space. In the present study, the effect of wheat streak on wheat yield was effectively described by state space models across the field. The state of grain yield at position d was significantly related to the state of wheat streak at the same position and/or at position $d-1$. The PEV/MD² ratios approximated unity suggesting a good fit (15). Even though R^2_D values were generally low, it is worth mentioning here that it can take negative values as an indication of a much less satisfactory goodness of fit (24).

Results obtained in this study suggest that state space methodology can be a very powerful tool in the study of plant disease epidemics, especially in monitoring trends in disease spread over time or space (across the field) as a function of other associated variables. The methodology also has a potential to be extended to cover assessments of diseases spread across regions or for studies of local focal expansions from a single point with or without explanatory variables.

Various models have been used in the past to describe plant disease spread both in space and time including its effect on yield (5,12,14,20,21,22,27,28,29,33,34,35,44). The state space model described in this article is similar, in some ways, to the autoregressive models described in some of the works cited above. However, state space is unique in its use of recursive filtering to update future predictions using current prediction errors as new

Fig. 8. Observed and predicted grain yield (g) modeled as a function of wheat streak intensity (reflectance at 555 nm) across the second field in the first and second transects (**A** and **B**, respectively) and their respective residuals (**C** and **D**) sampled in 2007. Dotted lines in **C** and **D** represent ± 2 standard deviations.

data are added. In addition, because of its multivariate nature, explanatory variables can be accounted for in the prediction. Further coverage of its theoretical background and applications are provided in some of the citations provided (8,10,11,15,17,37).

ACKNOWLEDGMENTS

We thank C. Childers, J. Arthur, J. Kirby, and L. Brown for their technical assistance. We are grateful to the wheat growers who allowed us to use their fields for the project.

LITERATURE CITED

- Anonymous. 1999. SAS/ETS User's Guide, Version 8. SAS Institute, Cary, NC.
- Anonymous. 2006. Texas Agricultural Statistics. Texas Agricultural Statistics Service, Texas Department of Agriculture.
- Box, G. E. P., Jenkins, G. M., and Reinsel, G. C. 1994. Page 598 in: Time Series Analysis, Forecasting and Control. 3rd ed. Prentice Hall, Englewood Cliffs, NJ.
- Brocklebank, J. C., and Dicky, D. A. 2003. SAS for Forecasting Time Series. 2nd ed. SAS Institute, Cary, NC.
- Campbell, C. L., and Madden, L. V. 1990. Introduction to Plant Disease Epidemiology. John Wiley & Sons, New York.
- Cassel, D. K., Wendroth, O., and Nielsen, D. R. 2000. Assessing spatial variability in agricultural experiment station field: Opportunities arising from spatial dependence. *Agron. J.* 92:706-714.
- Christian, M. L., and Willis, W. G. 1993. Survival of *Wheat streak mosaic virus* in grass hosts in Kansas from wheat harvest to fall wheat emergence. *Plant Dis.* 77:239-242.
- Commandeur, J. J. F., and Koopman, S. J. 2007. Introduction to State Space Time Series Analysis. Oxford University Press, Oxford, UK.
- Connin, R. V. 1956. The host range of the wheat curl mite, vector of wheat streak mosaic. *J. Econ. Entomol.* 49:1-4.
- Durbin, J. 2004. Introduction to state space time series analysis. Pages 3-25 in: State Space and Unobserved Component Models: Theory and Applications. A. Harvey, S. J. Koopman, and N. Shephard, eds. Cambridge University Press, Cambridge, MA.
- Durbin, J., and Koopman, S. J. 2001. Time Series Analysis by State Space Methods. Oxford University Press, Oxford, UK.
- Gregory, P. H. 1968. Interpreting plant disease dispersal gradients. *Annu. Rev. Phytopathol.* 6:189-212.
- Eubank, R. L. 2006. A Kalman Filter Primer. Chapman and Hall/CRC Press, Boca Raton, FL.
- Gottwald, T. R., Reynolds, K. M., Campbell, C. L., and Timmer, L. W. 1992. Spatial and spatiotemporal autocorrelation analysis of citrus canker in citrus nurseries and groves in Argentina. *Phytopathology* 82:843-851.
- Harvey, A. C. 1989. Forecasting, Structural Time Series Models, and the Kalman Filter. Cambridge University Press, Cambridge, MA.
- Hervey, A. C., and Durban, J. 1986. The effects of seat belt legislation on British road casualties: A case study in structural time series modeling. *J. Roy. Stat. Soc. A* 149:187-227.
- Harvey, A. C., Koopman, S. J., and Shephard, N. 2004. State Space and Unobserved Component Models: Theory and Application. Cambridge University Press, Cambridge, MA.
- Harvey, T. L., Seifers, D. L., and Martin, T. J. 2001. Host range differences between two strains of wheat curl mites (Acari: Eriophyidae). *J. Agric. Urban Entomol.* 18:35-41.
- Harvey, A. C., and Shephard, N. 1993. Structural Time Series Models. Pages 261-302 in: Handbook of Statistics. Vol. 11. G. S. Maddala, C. R. Rao, and H. D. Vinod, eds. Elsevier, New York.
- Holt, J., and Chancellor, T. C. B. 1996. Simulation modeling of the spread of rice tungro virus disease: The potential of management by rouging. *J. Appl. Ecol.* 33:927-936.
- Jeger, M. J. 1983. Analyzing epidemics in time and space. *Plant Pathol.* 32:5-11.
- Jeger, M. J., Holt, J., van den Bosch, F., and Madden, L. V. 2004. Epidemiology of insect-transmitted plant viruses: Modeling disease dynamics and control interventions. *Physiol. Entomol.* 29:291-304.
- Kalman, R. E. 1960. A new approach to linear filtering and prediction problems. *Trans. ASME J. Basic Eng.* 82:35-45.
- Koopman, S. J., Harvey, A. C., Doornik, J. A., and Shephard, N. 2007. Page 231 in: Stamp 8.0: Structural Time Series Analyzer, Modeler, and Predictor. Timberlake Consultants, London.
- Large, E. 1954. Growth stages in cereals: Illustrations of the Feekes scale. *Plant Pathol.* 3:128-129.
- Li, H., Lascano, R. J., Booker, J., Wilson, L. T., and Bronson, K. 2001. Cotton yield variability in heterogeneous soil at a landscape scale. *Soil Tillage Res.* 58:245-258.
- Madden, L. V., Hughes, G., and van den Bosch, F. 2007. The Study of Plant Disease Epidemics. American Phytopathological Society, St. Paul, MN.
- Madden, L. V., and Nutter, F. W. 1995. Modeling crop losses at the field scale. *Can. J. Plant Pathol.* 17:124-137.
- McRoberts, N., Hughes, G., and Madden, L. V. 1996. Incorporating spatial variability into simple disease progress models for crop pathogens. *Asp. Appl. Biol.* 46:75-82.
- Morkoc, F., Biggar, J. W., Nielsen, D. R., and Rolston, D. E. 1985. Analysis of soil water content and temperature using state-space approach. *Soil Sci. Soc. Am. J.* 49:798-803.
- Nielsen, D. R., and Alemi, M. H. 1989. Statistical opportunities for analyzing spatial and temporal heterogeneity of field soils. *Plant Soil* 115:285-296.
- Nielsen, D. R., and Wendroth, O. 2003. Spatial and Temporal Statistics: Sampling Field Soils and their Vegetation. Catena Verlag, Reiskirchen.
- Nutter, F. W. 1997. Quantifying the temporal dynamics of plant virus epidemics. *Crop Prot.* 16:603-618.
- Pethybridge, S. J., and Madden, L. V. 2003. Analysis of spatiotemporal dynamics of virus spread in an Australian hop garden by stochastic modeling. *Plant Dis.* 87:56-62.
- Reynolds, K. M., and Madden, L. V. 1988. Analysis of epidemics using spatio-temporal autocorrelation. *Phytopathology* 78:240-246.
- Seifers, D. L., Martin, T. J., Harvey, T. L., and Haber, S. 2007. Temperature-sensitive *Wheat streak mosaic virus* resistance identified in KS03HW12 wheat. *Plant Dis.* 91:1029-1033.
- Shumway, R. H., Azari, A. S., and Pawitan, Y. 1988. Modeling mortality fluctuations in *Lose Angles* as functions of pollution and weather effects. *Environ. Res.* 45:224-241.
- Shumway, R. H., Biggar, J. W., Morkoc, F., Bazza, M., and Nielsen, D. R. 1989. Time- and frequency-domain analysis of field observations. *Soil Sci.* 147:286-298.
- Shumway, R. H., and Stoffer, D. S. 2006. Time Series Analysis and its Applications with R Examples. 2nd ed. Springer-Verlag, New York.
- Slykuis, J. T. 1955. *Aceria tulipae* Keifer (Acarina:Eriophyidae) in relation to wheat streak mosaic. *Phytopathology* 45:116-128.
- Sorenson, H. 1985. Kalman Filtering: Theory and Application. IEEE Press, Piscataway.
- Stevenson, F. C., Knight, J. D., Wendroth, O., Kessel, V. C., and Nielsen, D. R. 2001. A comparison of two methods to predict the landscape-scale variation of crop yield. *Soil Tillage Res.* 58:163-181.
- Thomas, J. A., and Hein, G. L. 2003. Influence of volunteer wheat plant conditions on movement of the wheat curl mite, *Aceria tosichella*, in winter wheat. *Exp. Appl. Acarol.* 31:253-268.
- Thresh, J. M. 1976. Gradients in plant virus diseases. *Ann. Appl. Biol.* 82:381-406.
- Wendroth, O., Al-Omaran, A. M., Kirda, C., Reichardt, K., and Nielsen, D. R. 1992. State-space approach to spatial variability of crop yield. *Soil Sci. Soc. Am. J.* 56:801-807.
- Wendroth, O., Jurschick, P., Kersebaum, K. C., Reuter, H., van Kassel, C., and Nielsen, D. R. 2001. Identifying, understanding, and describing spatial processes in agricultural landscapes—Four case studies. *Soil Tillage Res.* 58:113-127.
- Workneh, F., Jones, D. C., Heidel, G. B., Price, J. C., and Rush, C. M. 2008. Surveys of wheat viruses in the Texas Panhandle. (Abstr.) *Phytopathology* 98(suppl.):S173.