Published July, 2008

RESEARCH

Patterns of Regional Yield Stability in Association with Regional Environmental Characteristics

Carol L. Williams,* Matt Liebman, Jode W. Edwards, David E. James, Jeremy W. Singer, Ray Arritt, and Daryl Herzmann

ABSTRACT

Regional-level recurring spatial patterns of yield variability are important for commercial activities, strategic agricultural planning, and public policy, but little is known about the factors contributing to their formation. An important step to improve our understanding is recognizing regional spatial patterns of yield variability in association with regional environmental characteristics. We examined the spatial distribution of county-level mean yields and CVs of mean yields of four functionally different crops—corn (*Zea mays* L.), soybean [*Glycine max* (L.) Merr.], alfalfa (*Medicago sativa*), and oat (*Avena sativa* L.)—in Iowa using Moran's Index of spatial autocorrelation. Patterns of association with 12 county-level climatic, edaphic, and topographic environmental characteristics were examined using partial least squares regression. Two distinct geographic provinces of yield stability were identified: one in the northern two-thirds of the state characterized by high mean yields and high yield constancy, and one in the southern third of the state characterized by low mean yields and low yield constancy. Among eight partial least squares regression models, which explained 50 to 81% of variation of mean yields and yield CVs, mean organic matter and mean depth to seasonally high water table had greatest relative importance to mean yields of grass crops and legume crops, respectively. Among the CV models, variables describing water availability were of greatest relative importance, with less distinct differences between grass and legume crops. Partial least squares regression is a potentially powerful tool for understanding regional yield variability.

C.L. Williams, M. Liebman, R. Arritt, and D. Herzmann, Dep. of Agronomy, Iowa State Univ., 2101 Agronomy Hall, Ames, IA 50011; J.W. Edwards, USDA-ARS, Corn Insects and Crop Genetics Research Unit, Ames, IA 50011; D.E. James and J.W. Singer, National Soil Tilth Lab., USDA-ARS Ames, IA 50011. Received 4 Jan. 2007. *Corresponding author (willico@iastate.edu).

Abbreviations: CRP, Conservation Reserve Program; GIS, geographic information systems; ISPAID, Iowa Soil Properties and Interpretation Database 7; NED, National Elevation Dataset; NWSCS, National Weather Service Cooperative Station; PLS, partial least squares regression; PRESS, predicted residual sum of squares.

An important characteristic of crop yields is variation over space and time. Yields vary as crops respond to spatial and temporal heterogeneity of the environment, changes in management, and interactions among these factors (Bakker et al., 2005; Kaspar et al., 2003; Wood et al., 2004). A rich agronomic literature addresses within-field yield variability as it relates to yield potential (Lobell and Ortiz-Monasterio, 2006), precision agriculture, and field management zones for increasing yields (Jaynes et al., 2003). At the regional scale, yield forecasting (Bannayan et al., 2007) focuses on estimation of absolute yields within a growing season by using process-based models and data on short-term changes in atmospheric and soil moisture conditions (Launay and Guérif, 2003; Stöckle et al., 2003). Yield stability, yet another focus of the yield variability literature (Mead et al., 1986; Tollenaar and Lee, 2002), includes measures of year-to-year constancy and relates to producers' concept of dependability (Berzsenyi et al., 2000; Mead et al., 1986; Tokatlidis and Koutroubas, 2004).

Published in Crop Sci. 48:1545–1559 (2008).
doi: 10.2135/cropsci2006.12.0837

677 S. Segoe Rd., Madison, WI 53711 USA

In addition to constancy, descriptions of recurrent spatial patterns of yield variability are a means to characterize yield stability. Recognition of field-level recurrent spatial patterns of yield variability is necessary for spatial management of fields by individual farmers (Jaynes et al., 2003; Kaspar et al., 2003; Schepers et al., 2004). At regional levels, recurrent spatial patterns of yield variability are important to longer-term commercial interests (Jagtap and Jones, 2002), strategic agricultural planning, and public policy formulation and application (DeWit et al., 2005; Lobell and Ortiz-Monasterio, 2006; Wassenaar et al., 1999). However, at regional scales, little is known about the interactions of multiple cultivars, multiple cropping systems, and environmental heterogeneity in pattern formation. An important first step to increasing this understanding is identification of regional recurring spatial patterns of yield variability in relation to environmental factors (Bates, 1995; Kravchenko et al., 2005).

Knowledge of deterministic relationships between yield variability and environmental heterogeneity beyond the scale of individual plants is limited (Day et al., 2003; Pausas et al., 2003). However, an expanding literature by ecologists, geographers, geologists, and climatologists addresses yield variability in association with environmental heterogeneity beyond the plant scale (Day et al., 2003; Hutchings and John, 2004; Si and Farrell, 2004; Waltman et al., 2004; Williams et al., 2008). From this literature, spatially explicit predictive (versus explanatory) modeling (e.g., Legendre et al., 2002; Lobell and Ortiz-Monasterio, 2006; Miller et al., 2007; Popp et al., 2005; Williams et al., 2008), and concepts of the association of different potential limiting factors with different spatial patterns of yield variability (Lobell and Ortiz-Monasterio, 2006) have emerged. Applied to regional scales, these approaches offer opportunities for identifying longer-term spatial patterns of yield variability in association with environmental heterogeneity and thus, a vital first step in developing hypotheses about causes of their formation (Begon et al., 1990).

Regression models are often used in analysis of vegetation–environment associations (Guisan and Zimmerman, 2000), but use of multiple environmental variables may be hampered by confounding and colinearity (Bakker et al., 2005; Helland, 1988; Wigley and Qipu, 1983). Alternative modeling approaches (Abdi, 2003; Wilson, 2007) can be used, however, to explore associations of regional-level yield variability with sets of environmental variables, particularly when assumptions necessary for ordinary multiple linear regression cannot be met. Partial least squares regression (PLS) is an approach to quantitative modeling of empirical relationships using covariance structures among strongly collinear, noisy variables (Abdi, 2003; Wold et al., 2001). Hence, use of PLS presents a promising approach for improved understanding of regional crop yield variability.

The goal of our analysis was to construct empirical models that describe recurrent spatial patterns of yield variability across a landscape in association with regional-level environmental variables, to improve understanding of the relationship between crops and regional environmental heterogeneity, and to identify potentially important drivers in regional yield variability. We hypothesized that longer-term spatial patterns of yield variability, as described by differences among regions within a landscape, were nonrandom and that these patterns were nonrandomly associated with differences in environmental conditions among regions. We also hypothesized that yield variability could be described using a limited set of environmental parameters.

We conducted geographic analysis of yield variability of corn (*Zea mays* L.), soybean [*Glycine max* (L.) Merr.], alfalfa (*Medicago sativa*), and oat (*Avena sativa* L.), four functionally different crops grown throughout the study area of Iowa. Our objectives were (i) to quantitatively describe spatial distribution of mean annual yields and the coefficient of variation of mean annual yields across the study area landscape, (ii) to quantitatively describe the spatial distribution of selected mean climatic, edaphic, and topographic conditions across the study area landscape, (iii) to quantify the degree of association between distributions of yield variability and environmental conditions, and (iv) to develop hypotheses about observed patterns. Regional-level analysis of recurrent spatial patterns of yield variability using spatially referenced data in combination with PLS modeling presents an opportunity for exploring empirical relationships between crop yields and environmental characteristics at spatial and temporal scales previously under-represented in the agronomic literature.

MATERIALS AND METHODS

Study Area

The area of this study was the state of Iowa, located between 89°30′00″ and 96°30′00″, and 40°30′00″ and 43°30′00″. Total area of Iowa is approximately 145,785 km^2, and elevation ranges from 146 to 509 m above sea level. Because of its latitude and interior continental location, Iowa's climate is characterized by distinct seasonal variation, with long, hot summers (Strahler and Strahler, 1984). Principal soil orders are mollisols, alfisols, inceptisols, and entisols (NRCS, 1999). About 89% of the land area of Iowa is in cultivation, and corn and soybean account for 93% of total land area harvested (USDA-ERS, 2007). Average annual temperature of Iowa ranges from 7.2°C in the extreme north to 11.1°C in the southeast. Average annual precipitation is 864 mm, ranging from 660 mm in the northwest to 965 mm in the southeast.

Regions

To measure and define yield variability, and to analyze associations of yield variability with environmental heterogeneity across the Iowa landscape, it was necessary to discretize the study area into spatial units (regions). State political subunits, or counties, are the smallest spatial unit for which agricultural statistics in Iowa are reported (NASS, 2005). Although these data give no clue of the geographic distribution of crop yields within counties, they provide information about the distribution of yields across the entire Iowa landscape. Iowa's 99 counties are relatively uniform in size and shape (e.g., rectangular), averaging 148,600 ha, and therefore are used as regions for spatial analysis in our study. Hence, for our analyses, $N = 99$.

Crop Yields

Crop yield can be characterized by mass per unit area (Evans and Fischer, 1999) and relative degree of constancy year to year (Mead et al., 1986). Using these two metrics, yield can be characterized as low and relatively constant year to year, low and relatively inconstant year to year, high and relatively constant year to year, or high and relatively inconstant year to year. Thus, yield stability, as defined by variability over time and space, can be measured by (i) average yield and (ii) the CV (i.e., CV%; Berzsenyi et al., 2000; Dobermann et al., 2003). These yield stability metrics can then be aggregated over larger geographic areas (Popp et al., 2005). In conventional agronomic studies, yield across years is measured by location–year data (e.g., county yield for each year within a period of years of observation). However, we hypothesized that although absolute yields have increased over the 20-yr study period (1985–2004), yield trends among counties would not be significantly different, and therefore, use of annual yield would not increase the information in the data. Hence, we used county mean yields averaged across the 20-yr study period of the four crops as dependent variables. County mean yield (hereafter, mean yield) was defined as total harvest (kg) divided by total area harvested (ha) per year, averaged across years. Four additional dependent variables were obtained from CVs of county mean yields among the 20 yr in the study for the four crops. In this study, then, yield stability across the Iowa landscape is measured as differences in mean yield among counties and differences in the CV of mean yield (CV%) among counties. We used crop yield data for the 20-yr period 1985–2004 because this period is likely representative of gains in corn yield throughout the latter half of the 20th century (Duvick and Cassman, 1999), and other agronomic crops have likely followed similar patterns of improvement. Yield data were acquired in tabular format from the National Agricultural Statistics Service (NASS, 2005). In our analyses, $N = 99$.

Environmental Variables

Based on associations of yields with environmental characteristics reported in the literature and the limits of data availability, we selected 12 environmental characteristics as potentially important to yield variability (Table 1). These 12 predictors include climatic, edaphic, and topographic attributes of the study area. To maintain a one-to-one relationship with yields and environmental characteristics, we used county-level environmental data. The 12 predictors were calculated as spatial means, with the exception of two climate variables. Variability of annual precipitation and variability of growing-season precipitation were calculated as temporal variances of spatially averaged variables.

Spatial averages of temporally permanent topographic and edaphic variables (i.e., county means) include land that is not in production. Bias in spatial averages of environmental variables may be a result. Unfortunately, the spatial distribution of crop production within counties is not reported in agricultural statistics, and land-cover data are unavailable year-to-year. Therefore, it is not possible to omit from spatial averaging of environmental variables those areas that are not in production, and it is unknown to what degree, if any, inclusion of these lands has introduced bias into the analysis. However, approximately 89% of the Iowa land surface is in agronomic production (USDA-ERS, 2007), which may limit the amount of bias unaccounted for as a result of inclusion of lands not in production.

Significant associations between crop yields and climate, including air temperature, precipitation, growing season length

Table 1. Summary of environmental parameters.

Category	Parameter	Units	Description	Source
Climatic	Mean annual precipitation	mm	Average amount of precipitation within a year	Iowa State University, Department of Agronomy (2005)
	Standard deviation of mean annual precipitation	mm		
	Mean growing season precipitation	mm	Average amount of precipitation between average date in spring with less than 20% chance of frost, and average date in fall with greater than 20% chance of frost	
	Standard deviation of mean growing season precipitation	mm		
Edaphic	Percent organic matter	%	Percentage of organic matter in tilled surface	Iowa State University (2004)
	Cation exchange capacity	$cmol_c\ kg^{-1}$	Sum of exchangeable cations held by soil	
	Plant-available water capacity	$cm\ cm^{-1}$ soil	Difference between amount of soil water at field capacity and wilting point	
	Depth to seasonally high water table	m	Depth to level of a saturated zone for 30+ days in most years	
	Percent sand	%	Percentage of sand in surface horizon	
	Permeability	$mm\ h^{-1}$	Rate which water moves down through saturated soil	
	Low value of pH range		Measure of acidity or alkalinity	
Topographic	Slope gradient	degrees	Degrees from horizontal	USGS (1999)

and heat accumulation, have been reported (DeWit et al., 2005; Wassenaar et al., 1999). Although temperature during the growing season, growing season length, and heat accumulation influence absolute yield, we posit that differences in these climate attributes would be insufficiently large across the study area to account for differences in yields among regions. Therefore, we included only precipitation variables in our models.

Precipitation predictors were derived from daily weather observations for the period 1985 to 2004, originating from 98 National Weather Service Cooperative Stations (NWSCS) evenly distributed among Iowa's counties (Iowa State University, Department of Agronomy, 2005). The NWSCS station in each county was considered representative of the county. Because Story County did not have a station, it was assigned values from the nearest station, in adjacent Boone County. Growing season was determined for each observation station as the number of days between the mean day-of-year in spring with less than 20% chance of the air temperature falling below 0°C, and mean day-of-year in fall with a greater than 20% chance of the air temperature dropping below 0°C. The amount of precipitation falling between these dates each year was determined as growing-season precipitation and was averaged across the observation period.

Significant relationships have also been found between spatial variability of crop yields and topographic attributes of elevation, slope, slope position, slope curvature, and aspect (Batchelor et al., 2002; Kravchenko and Bullock, 2000; Timlin et al., 1998; Wassenaar et al., 1999). Because our focus is the regional level (areas of tens of thousands of hectares), we did not include in our models topographic characteristics that were likely to have only local (i.e., within-field) influences, including slope curvature, aspect, and slope position. We also posited that the elevational range of Iowa was probably too small to have a significant influence on yield differences among regions, and it too was omitted from this study.

Slope was derived from the publicly available National Elevation Dataset (NED; USGS, 1999) in raster format at 30-m resolution. The NED consists of merged sets of digital elevation models at the 1:24,000 scale for the conterminous United States and provides the best currently available data. Vertical accuracy of the NED is estimated as ± 7 to 15 m (USGS, 2006). County means were calculated using the Zonal Statistics Tool with a county boundary data layer in ArcGIS 9.1 (ESRI, 2005).

Reports of significant relationships between yields and soil pH, cation exchange capacity, permeability, texture, water capacity, and depth to water table are abundant in the literature (e.g., Gish et al., 2005; Kaspar et al., 2004). These chemical and physical properties of soil may be relevant at broader spatial scales and over longer time spans as these soil attributes change relatively little compared to temporal weather conditions. These soil attributes are likely sufficiently varied across Iowa to influence regional yield differences and are therefore included in this study.

Soil predictors were derived from the publicly available Iowa Soil Properties and Interpretation Database 7.0 (ISPAID), which consists of rasterized national soil maps originating from the U.S. Natural Resources Conservation Service (Iowa State University, 2004). We used data from the ISPAID dataset that contains values for selected soil characteristics in a raster format at 100-m resolution. Values for soil characteristics are from surface layers, usually less than 19 cm from the top of soil, with the exception of depth to seasonally high water table. County means of each parameter were calculated using the Zonal Statistics Tool with a county boundary data layer in ArcGIS 9.1 (ESRI, 2005).

Statistical Models

Yield Trend

Crop and management improvements over the past century have greatly improved crop yields (Duvick and Cassman, 1999). It was hypothesized that rate of yield increase over the observation period would be equal among counties. Therefore, rate of yield increase would not explain differences in mean yields among counties and hence, not be a significant influence on longer-tern recurrent spatial patterns of yield stability. This hypothesis was tested using a year-from-base parameter with a unique identification number for each county to construct an interaction term, county *x* year-from-base.

Yield Stability

Analysis of spatial pattern of yield stability was conducted as a necessary first step in recognition of recurrent patterns and to inform hypothesis formulation. Moran's Index of spatial autocorrelation (Fortin and Dale, 2005, p. 124) was used in ArcGIS to determine whether distribution of mean yields and yield CVs among counties were randomly distributed, evenly distributed, or clustered. It is possible that low mean yields can occur in a stable system (i.e., relatively invariable; Berzsenyi et al., 2000; Mead et al., 1986). However, we hypothesized that mean yields would be negatively associated with yield CV (i.e., low yields would be associated with higher CVs). Therefore, pairwise correlation (SAS Institute, 2005) was used to test the relationship between mean yields and yield CVs.

Partial Least Squares Regression Models

Partial least squares regression is a multivariate regression method designed to build prediction models with highly correlated predictors and responses. Partial least squares is particularly well suited to complex environmental characteristics and crop responses, which can be highly collinear (Nguyen and Lee, 2006; Vågen et al., 2006). The PLS algorithm is an iterative procedure in which successive linear combinations of derived predictors, termed latent variables, are extracted so as to maximize the covariance between the linear combination of predictors and the response(s). Each successive latent variable is computed to be orthogonal to the previous latent variable and predicts the maximum proportion of variability in the response not predicted by prior latent variables (Frank and Friedman, 1993).

Partial least squares regression was used to develop prediction models of crop yield and crop CV. After data were centered and scaled, each of the eight responses (four yield responses and four CV responses) was regressed on the 12 environmental variables shown in Table 1 using PLS. Although it is possible to model simultaneously several responses in PLS, because this focus of this study was four functionally different crops, crop responses were modeled separately. We estimated the proportion of variance predicted for models containing from 1 to 12 factors. Significance of the addition of each success factor was tested using a delete-one cross validation approach in combination with van der Voet's (1994) test. The predicted residual sum of squares (PRESS) was estimated using delete-one cross-validation. Specifically, one observation at a time was removed, and the residual for the deleted observation was computed from the predicted value of that observation. Squared residuals were summed across observations to obtain the PRESS statistic. The significance of each factor was determined by comparing the PRESS statistics between the model with the smallest PRESS statistic and all models with a smaller number of factors. According to the van der Voet's (1994) procedure, the model with the smallest number of factors and a PRESS statistic that is not significantly larger than the minimum PRESS is taken as the final model. Significance of the difference in PRESS values was determined by using a Monte Carlo simulation of the differences in PRESS statistics as implemented in SAS Proc PLS (SAS Institute, 2005). We used 100,000 Monte Carlo samples to determine p-values because with 100,000 samples, p-values were relatively stable across runs.

Table 2. Crop yield summary statistics for the observation period 1985–2004.

Crop	Mean	SD	Min.	Max.	Median	Interquartile range	Skewness
Corn (kg ha^{-1})	8280.42	617.20	6540.7	9197.4	8368.0	759.1	−1.003
Soybean (kg ha^{-1})	2804.03	195.52	2315.4	3295.3	2828.0	232.4	−0.344
Alfalfa (Mg ha^{-1})	8.01	0.72	6.469	9.726	8.06	0.96	−0.211
Oat (kg ha^{-1})	2302.63	283.55	1666.1	2783.1	2372.3	397.8	−0.612

RESULTS AND DISCUSSION

Yield Trend

Summary statistics of crop yields are provided in Table 2. Mean yields of all four crops increased significantly over the period of observation (Table 3). The rate of increase computed by regression of mean yields on year-from-base for each individual county was found to be the same across all 99 counties for all crops but oat (Table 3), suggesting that use of 20-yr averages across counties is a sufficient representation of mean yields among counties for corn, soybean, and alfalfa. For the observation period, lowest mean oat yields for 75 of Iowa's 99 counties occurred in 1993, a year of unprecedented flooding (Johnson et al., 2004) in Iowa and the midwestern United States in general. Differences in oat yield trends among counties with 1993 data omitted are not significant (F = 0.87, prob. 0.79), suggesting that use of longer-term averages of oat yield provides sufficient representations among counties. Therefore, county mean oat yield, inclusive of 1993 data, is also used as a dependent variable in crop models.

Spatial Patterns of Yield Stability

Yield distributions of all crops exhibited a negative skewness (Table 2), where the number of counties with below-mode yield exceeded the number of counties with above-mode yield. Several counties had far-below mode yield (data not shown). Although Iowa possesses some of the best agricultural soils in the world (Prior, 1991), the skewing of the distribution of mean yields indicates (i) that only a few counties are "elite among the elite" and (ii) that there are differences among counties in the amount and use of "marginal" lands. While the actual amount of agronomic crop production on marginal lands over the observation period is unknown, enrollment of lands in the Conservation Reserve Program (CRP) provides a general indication of the distribution of marginal lands in Iowa. According to Secchi and Babcock (2007), the greatest amount of Iowa CRP lands is in the southern tier of counties, particularly among south-central counties. The amount of CRP in northeastern Iowa is also relatively high, but the fewest CRP lands are in central, north-central, and extreme northwestern Iowa. Regardless, visual examination of residual plots of the regressions of county mean yield on year-from-base (not shown) did not indicate evidence of serious assumption violations.

Mapped county mean yields are shown in Fig. 1. The geographic distributions of county mean yields of all four crops were significantly clustered, according to the Moran's I (Table 4). For all four crops, counties with higher mean yields occur in the northern two-thirds of the state, and counties with lower mean yields occur in the southern third of the state. Additionally, there are visible differences between the grass crops (corn and oat) and legume crops (soybean and alfalfa) in the distribution of mean yields within these two broader areas. Counties of higher mean yields of grass crops occur in central and northwestern Iowa, whereas counties of higher mean yields of legume

Table 3. Yield trends for the 99 Iowa counties and observation period 1985–2004.

Crop (R^2; RMSE)	df	F ratio	$P > F$
Corn (0.43; 1389.69)	197	6.86	<0.0001
YFB†	1	904.12	<0.0001
County	98	3.94	<0.0001
YFB × county	98	0.63	0.99
Soybean (0.26; 430.00)	197	3.24	<0.0001
YFB	1	212.96	<0.0001
County	98	4.13	<0.0001
YFB × county	98	0.22	1.00
Oat (0.26; 430.00)	197	4.51	<0.0001
YFB	1	91.34	<0.0001
County	98	5.69	<0.0001
YFB × county	98	2.50	<0.0001
Alfalfa (0.34; 1.23)	197	4.72	<0.0001
YFB	1	162.77	<0.0001
County	98	6.85	<0.0001
YFB × county	98	0.97	0.53

†YFB, year from base, the first year of the study (1985) being year 0; the second year (1986) is YFB 1, and so on.

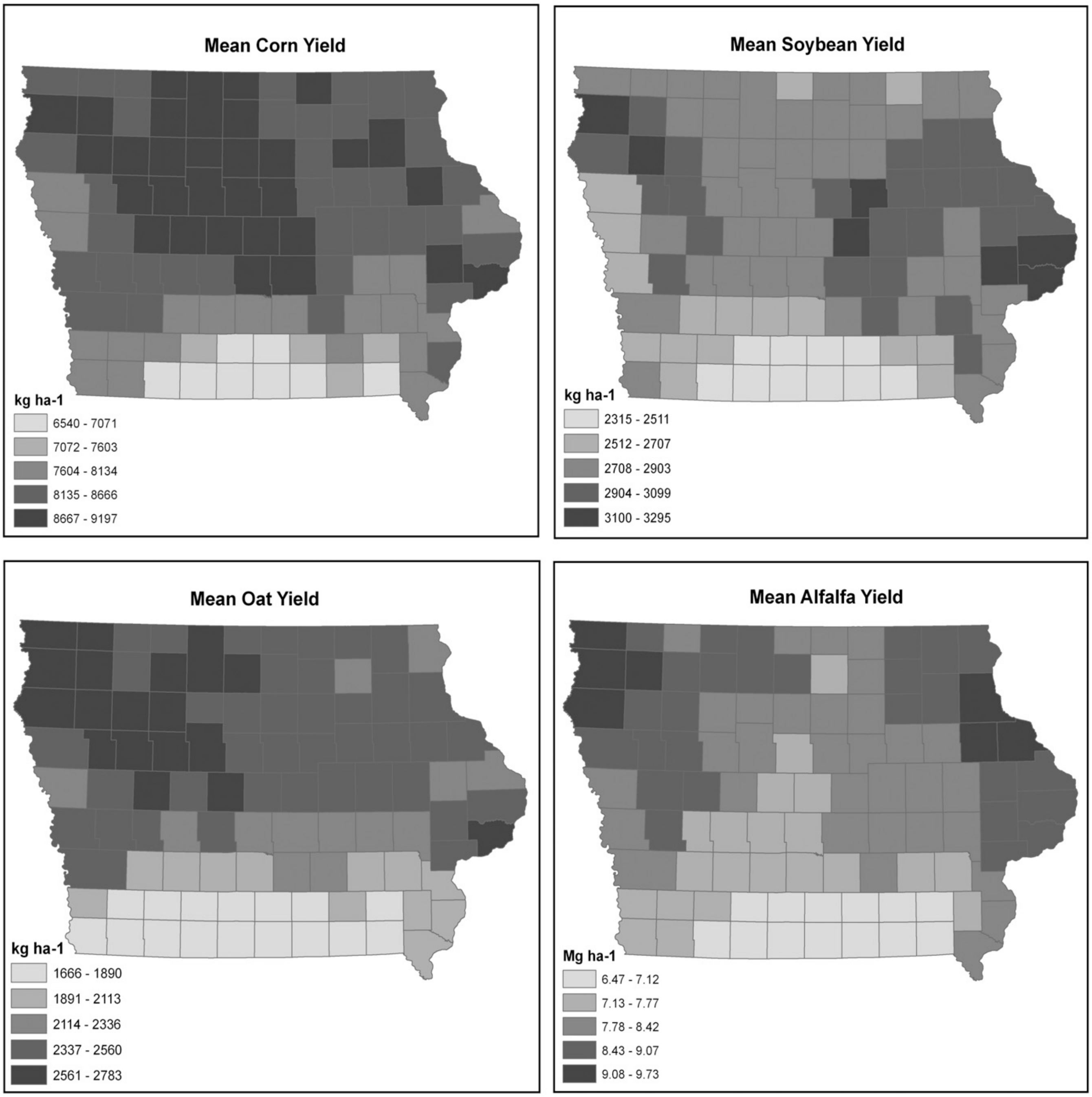

Figure 1. The distribution of mean yields (1985–2004) of four crops grown throughout the state of Iowa. The highest-yielding counties of all four crops occur in the northern two-thirds of the state, and counties with lowest mean yields occur in the southern third of the state.

crops occur in extreme northwest, western, and east-central Iowa. For all four crops, highest yield CVs occurred in counties of the southern third of the state, and lowest yield CVs occurred in counties of the northern two-thirds of the state (Fig. 2).

The general geographic pattern of association of higher yields with lower yield CVs (or lower yields with higher yield CVs) is supported by correlation analysis. For all four crops, mean yields were significantly negatively correlated with yield CVs (Table 5), similar to the findings of Tollenaar and Lee (2002), who reported an inverse relationship between mean yield and relative yield constancy in commercial maize hybrids. On the basis of mean yield and yield CV relationships, two distinct provinces of yield stability can be described: one of high and relatively constant mean yields in the northern two-thirds of the state, and the other of low and relatively inconstant mean yields in the southern third of the state.

There are substantial north-south differences in climate, soils and geomorphology in Iowa. Mean annual precipitation and mean growing season precipitation are greatest in the southeast and decrease to the northwest (Prior, 1991). Similarly, interannual variability of precipitation is greatest in the southern portions of Iowa,

Table 4. Values of the Moran's *I* test for spatial randomness.

Crop response	Corn mean yield	Corn CV	Soybean mean yield	Soybean CV	Mean oat yield	Oat CV	Alfalfa mean yield	Alfalfa CV
Moran's *I* (*P*-value)	0.21 (<0.001)	0.24 (<0.001)	0.14 (<0.001)	0.11 (<0.001)	0.25 (<0.001)	0.07 (<0.001)	0.20 (<0.001)	0.08 (<0.001)

decreasing to the northwest. Soils in the southern third of Iowa are loess-derived with greater agronomic limitations compared to the till-derived, organic matter–rich soils in the central and northern portions of the state (Prior, 1991). Topography of the southern third of the state is rolling, with shallow bedrock and limited level upland, whereas central and northern portions of Iowa are relatively level and undissected (Prior, 1991). Williams et al. (2008) describe agroecozones of Iowa by such regional differences in dominant environmental characteristics, as well as unique combinations of climatic, edaphic and topographic factors. As such, Iowa can be described as having two broad environmental provinces; one in the south characterized by relatively high amounts of precipitation

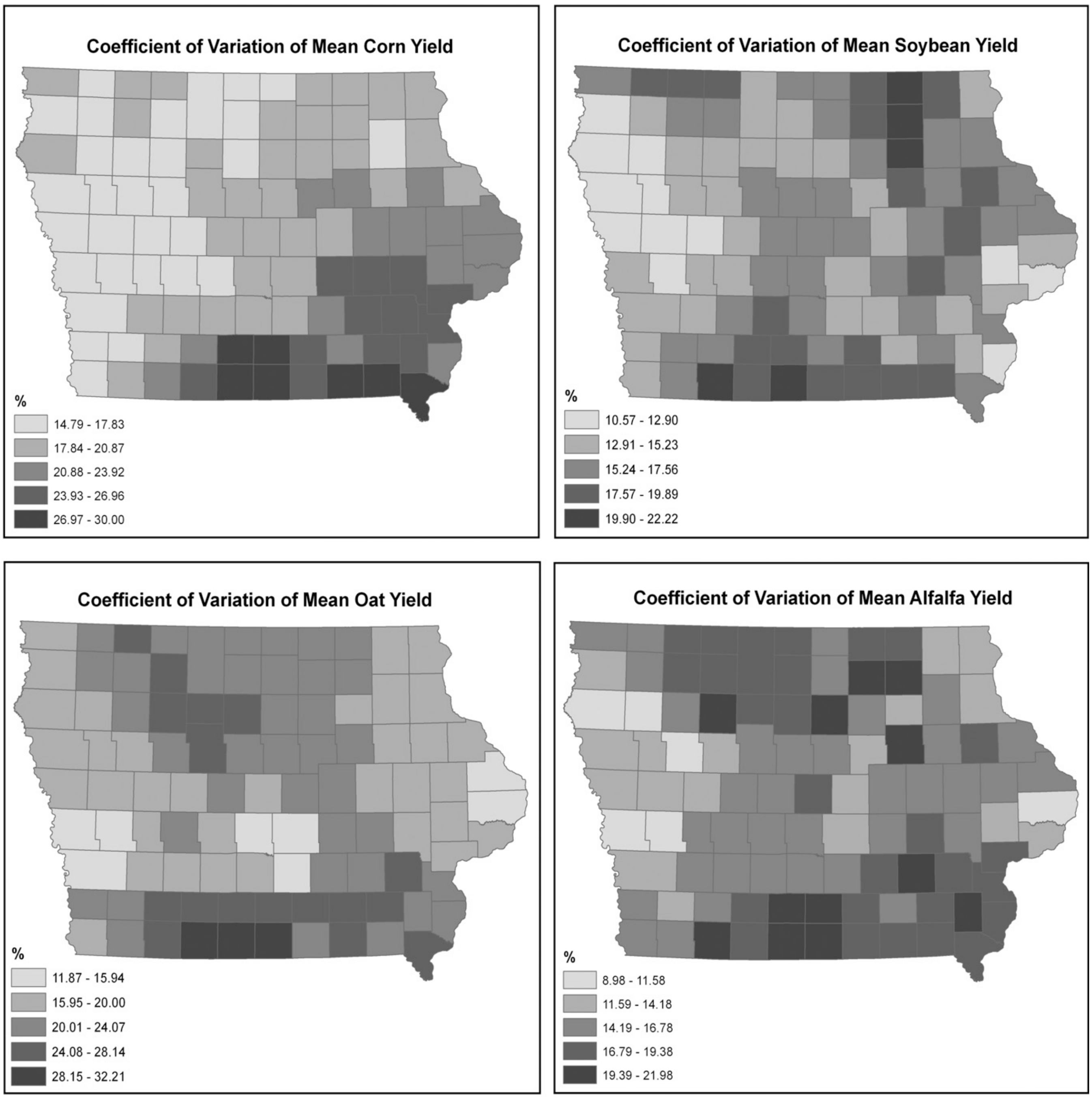

Figure 2. The distribution of yield CVs for the four crops of this study, indicating that the southern third of the state is a province of low yield constancy for all four crops.

Table 5. Correlation matrix of mean yield and yield CV.

Crop response	Mean corn yield	Corn CV	Mean soybean yield	Soybean CV	Mean alfalfa yield	Alfalfa CV	Mean oat yield	Oat CV
Mean corn yield	1							
Corn CV	–0.76***	1						
Mean soybean yield	0.78***	–0.70***	1					
Soybean CV	–0.03	–0.04	0.21*	1				
Mean alfalfa yield	0.67***	–0.65***	0.71***	0.32**	1			
Alfalfa CV	–0.34***	0.45***	–0.48***	–0.24*	–0.47***	1		
Mean oat yield	0.85***	–0.66***	0.68***	–0.01	0.76***	–0.38***	1	
Oat CV	–0.38**	0.55***	–0.52***	–0.30**	–0.49***	0.58***	–0.344**	1

*Significant at the 0.05 probability level.

**Significant at the 0.01 probability level.

***Significant at the 0.001 probability level.

and higher variability of precipitation, soils with relatively high agronomic limitations, and substantial portions of the land surface with relatively steep slopes. The other environmental province can be characterized as having lesser precipitation but relatively high constancy of precipitation, agronomically superior soils, and fewer areas of slope-limited suitability.

Models of Yield-Environment Associations

The performance of the PLS models, as measured by the coefficient of determination of the model (R^2; Nguyen and Lee, 2006) indicates that generally, a high amount of variation in crop response variables was accounted for by the environmental variables (Table 6). Parameter estimates for the PLS prediction equations are provided in Table 7. Each equation was obtained with the corresponding number of significant latent variables for each crop response. Overall, the amount of variation explained was greater among mean yield models than yield CV models. The amount of total variation explained among all eight models ranged from 50% (alfalfa yield CV) to 81% (corn mean yield). Quantile–quantile plots indicated that predicted values were within two standard deviations of observed values (i.e., plotted points approximate a 45 degree line; not shown) and thus affirm model performance as good. Less than 5% of counties fell outside the two-standard deviation range. We interpret the performance of the PLS models (e.g., high R^2 values) as indication of the overall importance of the selected environmental variables to the observed spatial patterns of yield stability of the agronomic crops of this study.

For each crop response, the optimum number of latent variables was less than seven (Table 6). Although inclusion of additional latent variables would have increased the amount of total variation taken into account, and minimized the absolute value of the PRESS, the optimal number of latent variables was determined as the number of factors after which explained variance no longer increased significantly (i.e., comparative p value > 0.05, Table 6; Geladi and Kowalski, 1986; Vågen et al., 2006). The loadings of environmental variables on the first latent variable (also known as X-loadings) are shown in Fig. 3. Magnitude of loadings (i.e., high values) correspond to maximum prediction information (Janik and Skjemstad, 1995) and is an indication of importance (Wold et al., 2001; Devillers et al., 2004; Holland et al., 2002; Nguyen and Lee, 2006). Sign of loadings indicates the direction of correlation (Wold et al., 2001). Loadings, however, are not necessarily direct indicators of functional drivers of response variables in the "soft modeling" approach of PLS (Abdi, 2003; Tobias, 2007). Although total variation of crop responses was maximized in some cases with up to seven latent variables, a large number of latent variables can make interpretation of individual predictors difficult (Wold et al., 2001). Most of the variability of all crop responses was explained by the first latent variable (Table 6); therefore, the first latent variable is used for interpretation of yield–environment associations in the following discussion.

Models of Mean Yield

For all four crops, mean growing season precipitation, the standard deviation of annual precipitation, the standard deviation of growing season precipitation, mean soil permeability, and mean slope were of moderate to high relative importance in explanation of the distribution of mean yields among counties (i.e., moderately high loadings for these variables; Fig. 3a). Loadings of relatively low magnitude were found for all four crop responses on mean annual precipitation, mean available water capacity, and mean cation exchange capacity (Fig. 3a). The magnitude of these loadings indicates that relative to other environmental characteristics, these variables were relatively less important in explanation of spatial variability of mean yield among counties.

The loadings of the remainder of the environmental variables (i.e., those with highest relative importance) indicate differences between the grass crops (corn and oat) versus the legume crops (soybean and alfalfa). Loadings for mean organic matter were very high for corn and oat but relatively low for soybean and alfalfa (Fig. 3a). Substantial

Table 6. Partial least squares regression (PLS) analysis of mean yield and yield CV, with cross-validation.

Response	Crop	Number of PLS factors	Percent XY variation accounted for (R^2)		Cross-validation	
			Current	Total	PRESS†	Comparison p
Mean yield	Corn	1	45.43	45.43	0.772	0
		2	19.98	65.42	0.672	0.0003
		3	7.62	73.05	0.613	0.0002
		4	2.09	75.14	0.598	0.0372
		5	4.22	79.36	0.552	0.0208
		6	1.52	80.88	0.512	0.0605
		7	0.59	81.47	0.502	0.1102
	Soybean	1	33.53	33.53	0.865	0
		2	16.15	49.69	0.797	0
		3	8.68	58.37	0.751	0.0273
		4	3.40	61.78	0.725	0.0358
		5	3.34	65.12	0.704	0.0817
	Oat	1	56.16	56.16	0.683	0
		2	13.25	69.42	0.595	0
		3	4.39	73.82	0.564	0.0032
		4	0.47	74.29	0.561	0.1742
	Alfalfa	1	47.58	47.58	0.756	0
		2	13.19	60.77	0.672	0
		3	4.33	65.11	0.654	0.0191
		4	2.18	67.29	0.639	0.0440
		5	1.09	68.38	0.640	0.4196
Yield CV	Corn	1	52.26	52.26	0.720	0
		2	10.22	62.48	0.672	0.0004
		3	4.72	67.20	0.640	0.0015
		4	1.52	68.73	0.626	0.0640
		5	2.62	71.35	0.606	0.1687
	Soybean	1	38.73	38.73	0.816	0
		2	9.09	47.82	0.768	0
		3	9.57	57.39	0.716	0.0044
		4	3.73	61.13	0.696	0.0101
		5	1.56	62.69	0.702	0.2078
	Oat	1	41.04	41.01	0.818	0
		2	9.19	50.23	0.776	0.0002
		3	3.14	53.37	0.771	0.1263
	Alfalfa	1	45.01	45.01	0.772	0
		2	4.90	49.91	0.752	0.0014
		3	3.52	53.43	0.741	0.1118

†PRESS, predicted residual sum of squares.

differences between the grass crops and legume crops were also observed in the loadings of mean depth to seasonally high water table (Fig. 3a). Loadings for mean percentage sand were almost identical (positive) for the grass crops and very similarly negative for the two legume crops (Fig. 3a).

The relationship of the grass crops with mean organic matter is graphically represented in visual comparison of mean yield maps (Fig. 1) with a map of mean organic matter (Fig. 4). Counties of higher mean yields of corn and oat correspond to counties of higher mean organic matter, which also happen to be counties of high mean percentage sand and mean pH (data not shown). Grasses, incapable of fixing atmospheric nitrogen like the leguminous crops, are dependent on soil organic matter (and external inputs) for nitrogen, and the occurrence of higher mean yields within areas of highest soil organic matter therefore comes as no surprise. It is also little surprise then, that the grass crops had higher loadings for soil pH compared with the leguminous crops as this soil characteristic has important influences on mineralization of nitrogen necessary for utilization by grass crops (Troeh and Thompson, 1993).

The counties of highest mean soybean and alfalfa yields were in the northwestern and northeastern portions of the state, corresponding to counties of greater mean depth to

Table 7. Partial least squares regression (PLS) parameter estimates.

Parameter	PLS model: Mean yield, Corn	Soybean	Oat	Alfalfa	Yield CV, Corn	Soybean	Oat	Alfalfa
Intercept†	5777.91	2706.84	1973.99	9.82	35.35	30.99	21.39	20.07
AP	0.60	0.30	−0.09	−0.0007	0.005	0.003	−0.001	0.0001
APSD	−3.10	−1.00	−1.77	−0.005	0.02	0.004	0.02	0.02
AWC	0.02	0.006	0.007	7.5E-07	−0.0001	−0.004	−0.0001	−0.0002
CEC	−11.80	−0.85	−0.17	−0.01	0.07	−0.05	0.24	0.02
GSP	−0.15	−0.29	−0.67	−0.002	−0.002	−0.005	0.0001	−0.0006
GSPSD	−0.98	0.19	−0.56	−0.003	0.07	0.001	0.005	0.01
OM	1.83	0.29	0.48	0.0008	−0.009	−0.0005	0.0002	0.0009
PRM	0.05	0.02	0.02	6.8E-05	−0.0002	−0.0002	−0.0003	−0.0001
pH	2.05	−0.48	0.99	0.0006	−0.03	−0.01	−0.002	−0.003
SND	−2.30	−3.95	1.09	−0.006	−0.02	0.14	0.02	0.04
SLP	−190.50	−76.86	−76.62	−0.12	0.17	0.49	0.27	−0.003
WT	1.92	0.60	0.64	0.002	−0.009	−0.003	−0.01	−0.007
Centered and scaled values								
Intercept	0	0	0	0	0	0	0	0
AP	0.06	0.10	−0.02	−0.07	0.09	0.11	−0.03	0.002
APSD	−0.15	−0.15	−0.19	−0.19	0.17	0.05	0.16	0.22
AWC	0.06	0.06	0.05	0.002	−0.07	−0.28	−0.07	−0.14
CEC	−0.07	−0.02	−0.002	−0.06	0.07	−0.07	0.22	0.03
GSP	−0.01	−0.07	−0.11	−0.12	−0.03	−0.10	0.002	−0.01
GSPSD	−0.05	0.03	−0.06	−0.12	0.13	0.02	0.04	0.13
OM	0.34	0.17	0.19	0.12	−0.25	−0.02	0.007	0.04
PRM	0.31	0.38	0.30	0.38	−0.16	−0.34	−0.28	−0.21
pH	0.15	−0.11	0.16	0.04	−0.32	−0.19	−0.02	−0.06
SND	−0.03	−0.16	0.03	−0.06	−0.05	0.44	0.04	0.12
SLP	−0.38	−0.49	−0.33	−0.21	0.06	0.24	0.09	−0.002
WT	0.52	0.52	0.38	0.40	−0.40	−0.22	−0.51	−0.41

†AP, mean annual precipitation; APSD, standard deviation of mean annual precipitation; AWC, plant-available water capacity; CEC, soil cation exchange capacity; GSP, mean growing season precipitation; GSPSD, standard deviation of mean growing season precipitation; OM, soil organic matter; PRM, soil permeability; SND, soil percentage sand; SLP, slope; WT, depth to seasonally high water table.

seasonally high water table (Fig. 4), which also happen to be counties of lower mean pH and lower mean percentage sand (data not shown). The associations of soybean and alfalfa yields with depth to water table are documented in agricultural education materials (Hall et al., 2004), as well as peer-reviewed literature (Ogunremi et al., 1981). Leguminous crops do not tolerate saturated soils (i.e., high water tables) because of effects on nodulation and fixing of atmospheric nitrogen, as well as increased incidence of fungal diseases. High soil pH has been associated with soybean disease (Sanogo and Yang, 2001) and may therefore account for lower soybean yields in central and north-central Iowa compared to mean yields of corn and oat. Similarly, potassium availability may be limited in higher soil pH, and this may be an underlying cause of lower relative yields of alfalfa in central and north-central Iowa (area of highest corn and oat yield; Peters et al., 2000).

The direction of correlation between all four mean yield responses and all seven soil variables was positive (Fig. 3). The direction of this correlation is not surprising given the published information on the relationships between yields and these variables (Kaspar et al., 2004; Kravchenko and Bullock, 2000). Likewise, the direction of correlation between all four mean yield responses and slope was negative (Fig. 3). The direction of this correlation is also not surprising given the published information on the yield-reducing effects of increased slope (Kravchenko and Bullock, 2000; Kravchenko et al., 2005; Timlin et al., 1998). However, the negative direction of correlation between all four mean yield responses and the four climatological predictor variables (Fig. 3), requires interpretation. Precipitation in Iowa is generally adequate to maintain a positive moisture balance (Widrlechner, 1999), and the geographic pattern of mean growing-season precipitation decreases from southeast to northwest (Fig. 5a) and is similar for mean annual precipitation (not shown). Mean annual precipitation and mean growing-season precipitation have significant negative bivariate correlations with all four mean crop yields (data not shown), and the relationships can be visualized through comparison of the map of mean growing season precipitation (Fig. 5a) and maps of mean crop yields (Fig. 1). However, precipitation should not be interpreted as driving yields. Instead, because precipitation is generally adequate for rainfed agriculture throughout the state, it is more likely that other variables are influencing yield. That is, the general spatial pattern of crop yields in Iowa is such that locations with higher yields are locations with better soils and, by geographic happenstance are also those locations with lower precipitation. Support for this interpretation is found in the significant negative bivariate correlations of mean growing season precipitation with mean organic matter ($r = -0.25$, $p < 0.01$), mean pH (-0.37, $p < 0.01$), and mean depth to water table ($r = -0.30$, $p < 0.01$), and similar with mean annual precipitation (not shown). As discussed above, these soil variables were highly important in the first latent variable and had positive correlation directions with mean yields.

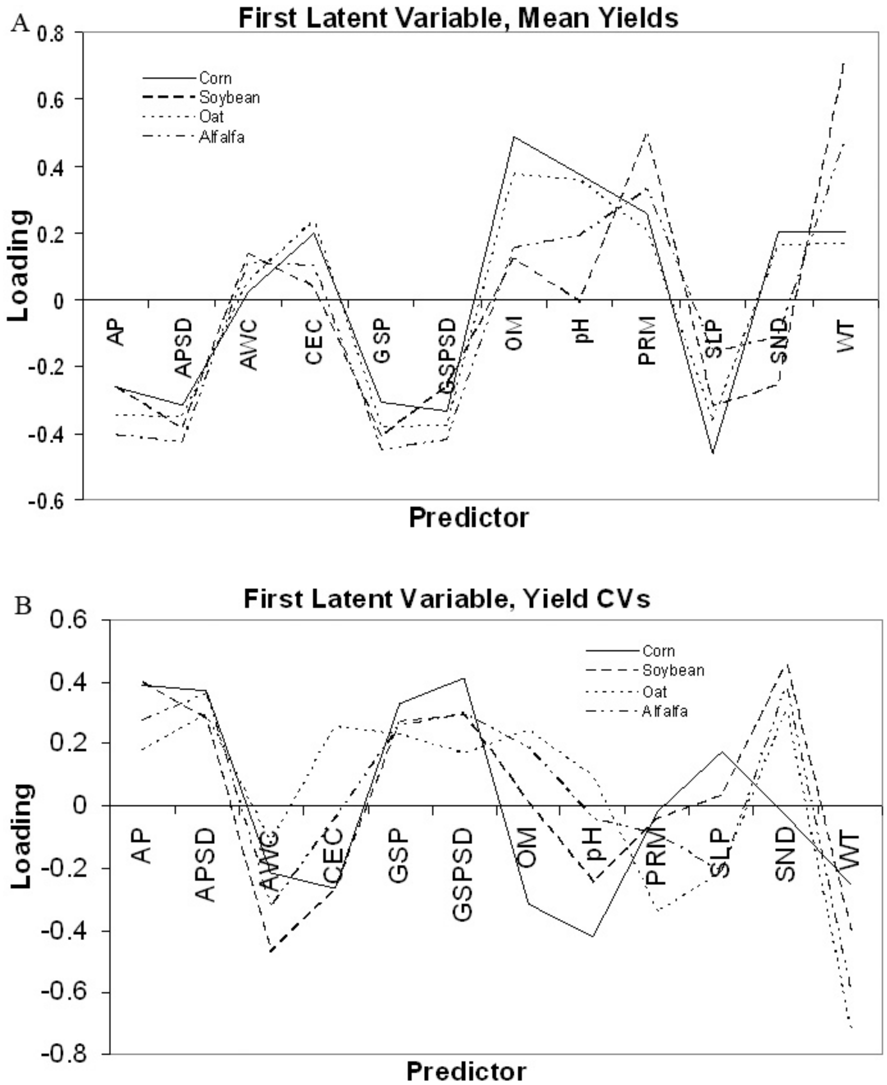

Figure 3. Loadings for the first latent variable of the crop models. (A) Differences in the relative importance of specific environmental variables in the models of mean yields indicate differences between the grass crops versus the legume crops. B) Differences in the relative importance of specific environmental variables in the yield CV models is greater between corn and oat, than between the grass crops versus legume crops. AP, mean annual precipitation; APSD, standard deviation of mean annual precipitation; AWC, mean plant available water capacity; CEC, mean cation exchange capacity; GSP, mean growing season precipitation; GSPSD, standard deviation of mean growing season precipitation; OM, mean % organic matter; pH, mean pH; PRM, mean soil permeability; SLP, mean slope; SND, mean % sand; WT, mean depth to seasonally high water table.

Models of Yield Coefficients of Variation

Unlike the mean yield models, crops did not have the same environmental variables with highest relative importance among the CV models. There were fewer instances of distinct differences between grass crops versus legume crops compared to the mean yield models. On the first latent variable, magnitudes of loadings of all four crops were similarly moderate for the standard deviation of mean annual precipitation and mean growing season precipitation (Fig. 3b), indicating moderate relative importance of these climate variables for all four crops. Small-to-moderate differences in magnitude of loadings among grass crops versus legume crops was observed for mean available water capacity, mean cation exchange capacity, mean organic matter, mean pH, and mean slope, although the signs of the loadings among the grass crops and among the legume crops were opposite for mean cation exchange capacity, mean organic matter, mean pH, and mean slope (Fig. 3b). The loadings for mean percentage sand and mean depth to seasonally high water table indicate moderate-to-high relative importance of these variables for all the crops except corn, for which they are relatively unimportant (Fig. 3b).

Differences in the magnitudes of the loadings among the CV models indicate that differences among the legume crops were minimal. The differences between the legume crops versus corn were moderate, as were the differences between the legume crops versus oat. However, the most striking pattern among the yield CV models are the differences among the grass crops. A difference of 0.24 was observed in the magnitudes of the loadings for the standard deviation of mean growing season precipitation of the grass crops (higher for corn; Fig. 3b). Water use efficiency of the C_4 crops (e.g., corn) is generally greater than that of the C_3 (e.g., oat) particularly under conditions of heat stress

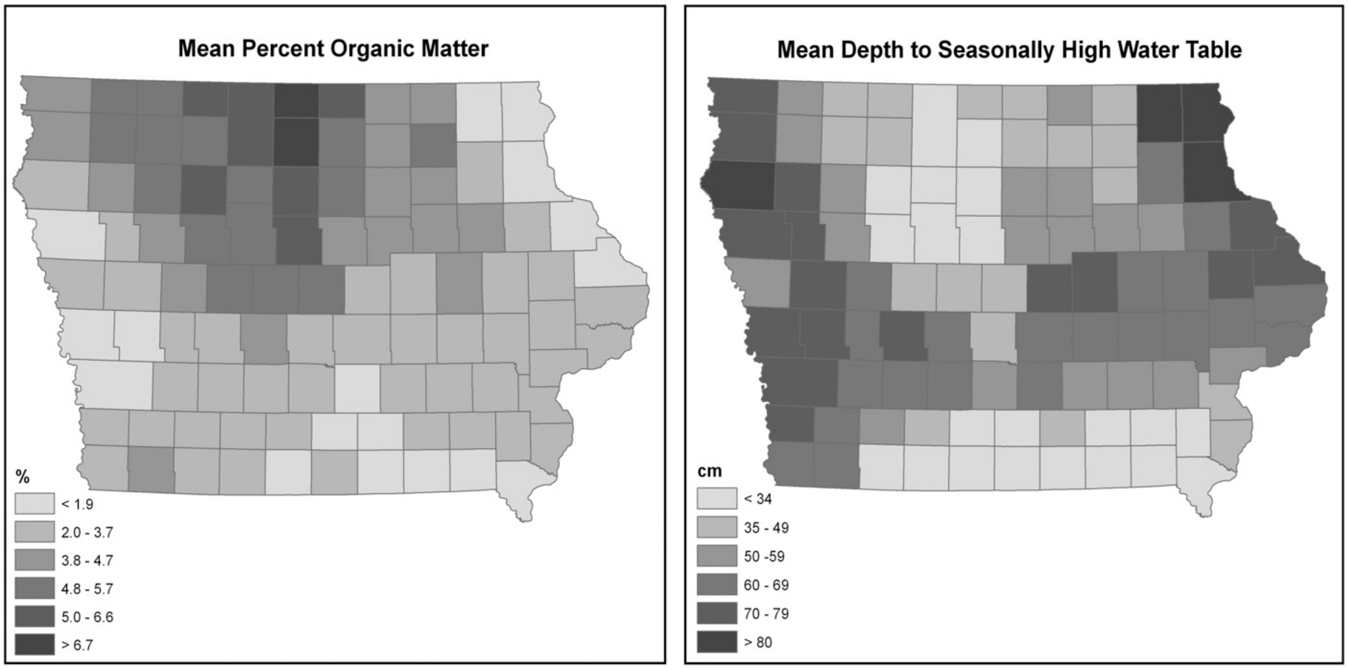

Figure 4. The distributions of mean organic matter and mean depth to seasonally high water table in Iowa. The areas of higher organic matter correspond to areas of higher grass crop yields. The areas of greater depth to seasonally high water table correspond to areas of higher legume crop yields.

(Long, 1999). However, water availability during anthesis and grain-fill are critical for corn (Classen and Shaw, 1970; NeSmith and Ritchie, 1992). Therefore, interannual variability of growing season precipitation could greatly influence interannual corn yield variability. The grass crops differed by 0.33 in the magnitudes of loadings for mean permeability (highest for oat), by 0.33 in the magnitudes of loadings for mean percentage sand (highest for oat), and by 0.47 in the magnitude of loadings for mean depth to seasonally high water table (highest for oat; Fig. 3b). That is, mean permeability, mean percentage sand, and mean depth to seasonally high water table were of moderate to high relative importance in the CV of oat and of low or very little relative importance in corn CV. Percentage sand, permeability, and depth to seasonally high water table affect soil drainage. Oat is an early crop, dependent on cool weather for high yields (Stoskopf, 1985). Well-drained soils warm sooner than less-well-drained soils, permitting earlier planting of oats and decreasing the chances of yield-reducing late-season heat (Stoskopf, 1985).

Lobell and Ortiz-Monasterio (2006) found that spatial patterns of yield contain information on the relative importance of soil and management factors in yield variability. Calvino and Sadras (1999) found variation in the interactions of soil depth and precipitation variability resulting in variation of water availability, leading to variation in soybean yields. In a test of simulation models, Riha et al. (1996) examined the influence of variability of temperature and precipitation on crop yields at locations among three soil types and found that the yield-reducing effect of increased precipitation variability was mediated by soil characteristics. This suggests that in southern Iowa, relatively high interannual variability of precipitation cannot be mediated by soil characteristics.

Regarding the sign of loadings, a pattern opposite that of the mean yield models was observed. For example, among CV models, the direction of correlation of all four crop responses with all four climatological variables was positive. Figure 5b illustrates the distribution of interannual variability of precipitation in Iowa. Visual comparison of maps of yield CVs (Fig. 2) with a map of the standard deviation of growing season precipitation (Fig. 5b) demonstrates the general spatial relationship where counties of higher interannual yield variability are counties with greater interannual variability of precipitation. Indeed, the bivariate relationship of all four CV responses with all four climatological variables is positive (data not shown). Simultaneously, counties with higher standard deviation of precipitation are counties with lower soil organic matter (Fig. 4), lower cation exchange capacity, lower permeability and lower plant-available water capacity (data not shown). These findings are interpreted as an indication that in locations where interannual variability of precipitation is relatively high, soil quality is inadequate to compensate (e.g., store moisture), so that in years of low precipitation, moisture stress leads to reduced yield and therefore, increased yield CV.

Applications

The use of a 20-yr data set for an entire state has provided a means for identifying subregions of distinctly different character with potentially important relevance to agriculture, particularly crop breeding. A primary strategy for overcoming lower yields of corn in southern Iowa has been an effort

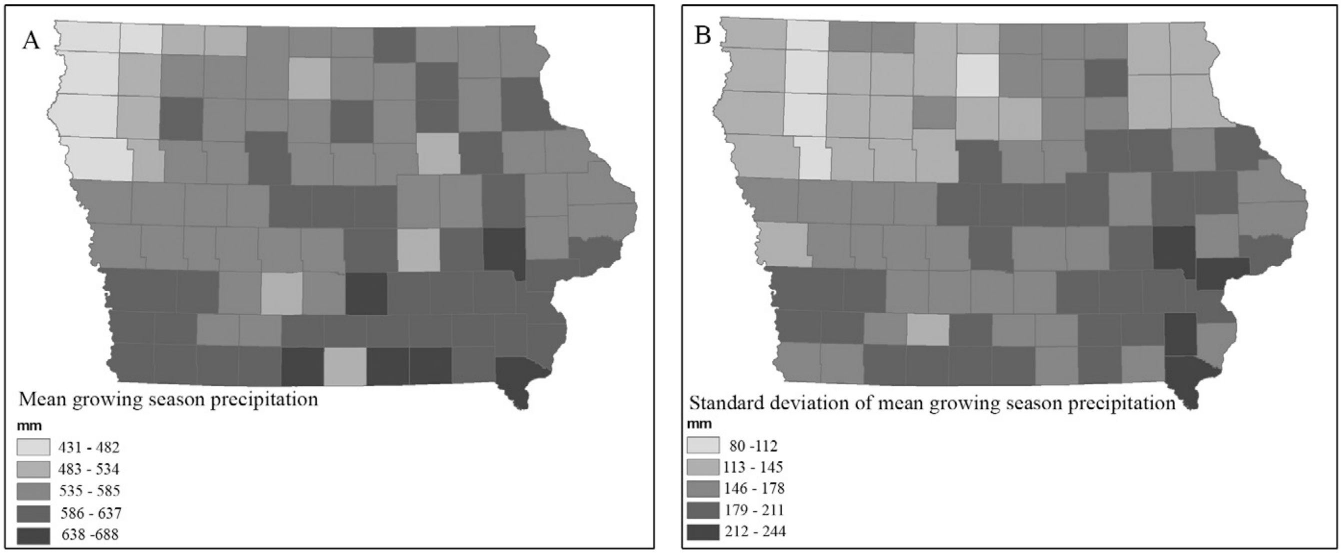

Figure 5. The distributions of mean growing season precipitation and the standard deviation of mean growing season precipitation in Iowa. (A) Mean growing season precipitation decreases from the southeast to the northwest and is similar to the distribution of mean annual precipitation ($r = 0.86$, $p < 0.001$). (B) The SD of mean growing season precipitation is an indication of interannual variability of growing season precipitation and is similar to the pattern of distribution of the SD of mean annual precipitation ($r = 0.81$, $p < 0.001$).

to increase yield potential, but at the cost of yield constancy (Zhisheng Qing, personal communication, 2007). Deployment of high yield–low constancy varieties in combination with relatively high inconstancy of precipitation and lower soil quality in southern Iowa may be a major factor in the perpetuation of lower yields in the area. Alternatively, breeding to increase yield constancy could improve the likelihood of increased mean yields for the area.

The ability to identify distinct subregions within a large landscape is potentially important to issues of risk, including assessment, management, and mitigation. Decision making by individual land managers could be improved by understanding the longer-term patterns of yield stability of their subregion, potentially resulting in reduced losses over time, especially with the availability of crop varieties specifically suitable for their subregion. Likewise, state and federal policy formulation for addressing resource limitations could be geographically targeted, potentially reducing costs while increasing resource protection (Akyurek and Okalp, 2006). Lastly, by identifying yield subregions, future research efforts, particularly field-based experiments for exploring deterministic relationships between yield stability patterns and specific drivers, can be more efficiently planned and implemented. Indeed, it is this that the present study aims to inform.

CONCLUSIONS

Use of PLS, while common in chemical science, is relatively rare in the ecological sciences. However, it is gaining recognition as a useful tool in ecological analysis because of the method's ability to analyze strongly collinear variables, as is typical of ecological data. In addition to providing support for previous studies finding that broad-scale environmental heterogeneity is a key element in crop yield variability beyond the field level, use of PLS in this study, in combination with spatial analyses, permitted the identification of yield-stability regions, as well as differences in environmental associations among crop functional types. However, the full potential of PLS regression in increasing knowledge of crop–environment relationships is probably not adequately understood. Future research should explore the further potentials for application of PLS in agronomic and agroecological studies.

Our study demonstrates how relatively inexpensive, publically available data can be used to address agroecological questions and produce important results. As the quality and accessibility of such data continue to increase, researchers should find increased opportunities for using these data in increasingly robust ways. Identifying crop–environment regions makes it possible to conduct subsequent research that incorporates environmental differences among regions as a fixed rather than a random effect. This possibility should greatly encourage researchers using traditional agronomic (i.e., plot) experiments to more fully consider location and distribution of plots in, for example, crop breeding and crop introduction studies. Using such an approach, future agronomic and agroecological research could contribute important information to an expanding literature on crop variability beyond the field level.

Acknowledgments

Phillip Dixon, Department of Statistics, Iowa State University, is gratefully acknowledged for his contributions. Special thanks to Robin Gomez and Michael Cruse, and the helpful comments of anonymous reviewers. This research was funded by the Agricultural Systems Initiative of the College of Agriculture and Life Sciences, Iowa State University.

References

Abdi, H. 2003. Partial least squares (PLS) regression. p. 792–795. *In* M. Lewis-Beck, A. Bryman and T. Futing (ed.) Encyclopedia of social sciences research methods. Sage, Thousand Oaks, CA.

Akyurek, Z., and K. Okalp. 2006. A fuzzy-based tool for spatial reasoning: A case study on soil erosion hazard prediction. Available at http://www.spatial-accuracy.org/2006/PDF/Akyurek2006accuracy.pdf (verified 23 Apr. 2008). *In* M. Caetano and M. Painho (ed.) Proc. 7th Int. Symp. on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, Lisbon, Portugal. 5–7 July 2006. Instituto Geográfico Português, Lisbon, Portugal.

Bakker, M.M., G. Govers, F. Ewert, M. Rounsevell, and R. Jones. 2005. Variability in regional wheat yields as a function of climate, soil, and economic variables: Assessing the risk of confounding. Agric. Ecosyst. Environ. 110:195–209.

Bannayan, M., N.M.J. Crout, and G. Hoogenboom. 2007. Application of the CERES-Wheat model for within-season prediction of winter wheat in the United Kingdom. Agron. J. 95:114–125.

Batchelor, W.D., B. Basso, and J.O. Paz. 2002. Examples of strategies to analyze spatial and temporal yield variability using crop models. Eur. J. Agron. 18:141–158.

Bates, J.W. 1995. Numerical analysis of bryophyte-environment relationships in lowland English flora. Fragm. Flor. Geobot. 40:471–490.

Begon, M., J.L. Harper, and C.R. Townsend. 1990. Ecology: Individuals, populations, and communities. Blackwell, Malden, MA.

Berzsenyi, Z., B. Györffy, and D.-Q. Lap. 2000. Effect of crop rotation and fertilisation on maize and wheat yields and yield stability in a long-term experiment. Eur. J. Agron. 13:225–244.

Calvino, P.A., and V.O. Sadras. 1999. Interannual variation in soybean yield: Interaction among rainfall, soil depth and crop management. Field Crops Res. 63:237–246.

Classen, M.M., and R.H. Shaw. 1970. Water deficit effects on corn: II. Grain component. Agron. J. 62:652–655.

Day, K.J., M.J. Hutchings, and E.A. John. 2003. The effects of spatial pattern of nutrient supply on yield, structure and mortality in plant populations. J. Ecol. 91:541–553.

Devillers, J., J.C. Doré, M. Tisseur, S. Cluzeau, and G. Maurin. 2004. Modelling the flight activity of Apis mellifera at the hive entrance. Comput. Electron. Agric. 42:87–109.

DeWit, A.J.W., H.L. Boogaard, and C.A. van Diepen. 2005. Spatial resolution of precipitation and radiation: The effect on regional crop yield forecasts. Agric. For. Meteorol. 135:156–168.

Dobermann, A., J.L. Ping, V.I. Adamchuk, G.C. Simbahan, and R.B. Ferguson. 2003. Classification of crop yield variability in irrigated production fields. Agron. J. 95:1105–1120.

Duvick, D.N., and K.G. Cassman. 1999. Post-green revolution trends in yield potential of temperate maize in the north-central United States. Crop Sci. 39:1622–1630.

ESRI. 2005. ArcGIS 9.1. Environmental Systems Research Institute. Redlands, CA.

Evans, L.T., and R.A. Fischer. 1999. Yield potential: its definition, measurement, and significance. Crop Sci 39:1544–1551.

Fortin, M.-J., and M. Dale. 2005. Spatial analysis: A guide for ecologists. Cambridge Univ. Press, Cambridge, UK.

Frank, I.E., and J.H. Friedman. 1993. A statistical view of some chemometrics regression tools. Technometrics 35:109–135.

Geladi, P., and B.R. Kowalski. 1986. Partial least-squares regression: A tutorial. Anal. Chim. Acta 185:1–17.

Gish, T.J., C.L. Walthall, C.S.T. Daughtry, and K.-J.S. King. 2005. Using soil moisture and spatial yield patterns to identify subsurface flow pathways. J. Environ. Qual. 34:274–286.

Guisan, A., and N. Zimmerman. 2000. Predictive habitat distribution models in ecology. Ecol. Model. 135:147–186.

Hall, M.H., J.A. Jennings, and G.E. Shewmaker. 2004. Alfalfa management guide. Available at http://www.plantmanagementnetwork.org/pub/fg/management/2004/alfalfa/. Forage Grazinglands

Helland, I. 1988. On the structure of partial least squares regression. Commun. Statist. Simulation Comput. 17:581–607.

Holland, J.B., A. Bjørnstad, K.J. Frey, M. Gullord, and D.M. Wesenberg. 2002. Recurrent selection for broad adaptation affects stability of oat. Euphytica 126:265–274.

Hutchings, M.J., and E.A. John. 2004. The effects of environmental heterogeneity on root growth and root/shoot partitioning. Ann. Bot. (Lond.) 94:1–8.

Iowa State University. 2004. Iowa soil properties and interpretations database manual. Version 7.0. Iowa State University, Iowa Agriculture and Home Economics Exp. Stn., Univ. Ext. Serv., Ames.

Iowa State University, Department of Agronomy. 2005. Iowa Environmental Mesonet: IEM networks. Available at http://mesonet.agron.iastate.edu/info/network.phtml?network=IA_COOP (verified 23 Apr. 2008). Iowa State Univ., Ames.

Jagtap, S.S., and J.W. Jones. 2002. Adaptation and evaluation of the CROPGRO-soybean model to predict regional yield and production. Agric. Ecosyst. Environ. 93:73–85.

Janik, L.J., and J.O. Skjemstad. 1995. Characterization and analysis of soils using mid-infrared partial least squares: II. Correlations with some laboratory data. Aust. J. Soil Res. 33:637–650.

Jaynes, D.B., T.C. Kaspar, T.S. Colvin, and D.E. James. 2003. Cluster analysis of spatiotemporal corn yield patterns in an Iowa field. Agron. J. 95:574–586.

Johnson, G.P., R.R. Holmes, Jr., and L.A. Waite. 2004. The great flood of 1993 on the Upper Mississippi River—10 years later. USGS Fact Sheet 2004-3024. Available at http://il.water.usgs.gov/pubs/fs2004-3024.pdf (verified: 15 January 2008). USGS, Reston, VA.

Kaspar, T.C., T.S. Colvin, D.B. Jaynes, D.L. Karlen, D.E. James, D.W. Meek, D. Pulido, and H. Butler. 2003. Relationship between six years of corn yields and terrain attributes. Precis. Agric. 4:87–101.

Kaspar, T.C., D.J. Pulido, T.E. Fenton, T.S. Colvin, D.L. Karlen, D.B. Jaynes, and D.W. Meek. 2004. Relationship of corn and soybean yield to soil and terrain properties. Agron. J. 96:700–709.

Kravchenko, A.N., and D.G. Bullock. 2000. Correlation of corn and soybean grain yield with topography and soil properties. Agron. J. 92:75–83.

Kravchenko, A.N., G.P. Robertson, K.D. Thelen, and R.R. Harwood. 2005. Management, topographical, and weather effects on spatial variability of crop grain yields. Agron. J. 97:514–523.

Launay, M., and M. Guérif. 2003. Ability for a model to predict crop production variability at the regional scale: An evaluation for sugar beet. Agronomie 23:135–146.

Legendre, P., M.R.T. Dale, M.-J. Fortin, J. Gurevitch, M. Hohn, and D. Myers. 2002. The consequences of spatial structure for the design and analysis of ecological field surveys. Ecography 25:601–615.

Lobell, D.B., and J.I. Ortiz-Monasterio. 2006. Regional importance of crop yield constraints: Linking simulation models and geostatistics to interpret spatial patterns. Ecol. Model. 196:173–182.

Long, S.P. 1999. Environmental responses. p. 215–249. *In* R.F. Sage and R.K. Monson (eds). C4 plant biology. Academic Press, San Diego, CA.

Mead, R., J. Riley, K. Dear, and S.P. Singh. 1986. Stability comparison of intercropping and monocropping systems. Biometrics 42:253–266.

Miller, J., J. Franklin, and R. Aspinall. 2007. Incorporating spatial dependence in predictive vegetation models. Ecol. Model. 202:225–242.

NASS. 2005. Agricultural statistics database. Available at http://www.nass.usda.gov/Data_and_Statistics/Quick_Stats/index.asp (verified 23 Apr. 2008). USDA–NASS, Washington, DC.

NeSmith, D.S., and J.T. Ritchie. 1992. Short- and long-term responses of corn to a preanthesis soil water deficit. Agron. J. 84:107–113.

Nguyen, H.T., and B.-W. Lee. 2006. Assessment of rice leaf growth and nitrogen status by hyperspectral canopy reflectance and partial least squares regression. Eur. J. Agron. 24:349–356.

NRCS. 1999. Soil taxonomy, a basic system of soil classification for making and interpreting soil surveys. 2nd ed. USDA–NRCS, Washington, DC.

Ogunremi, L.T., R. Lal, and O. Babalola. 1981. Effects of water table depth and calcium peroxide application on cowpea (*Vigna unguiculata*) and soybean (*Glycine max*). Plant Soil 63:275–281.

Pausas, J.G., C. Jordi, F. Albert, and X. Font. 2003. Coarse-scale plant species richness in relation to environmental heterogeneity. J. Veg. Sci. 14:661–668.

Peters, J.B., K.A. Kelling, and P.E. Speth. 2000. Alfalfa response to K at various soil pH levels. Available at http://www.soils.wisc.edu/extension/FAPM/proceedings/5A.peters.pdf (verified 23 Apr. 2008). p. 397–404. *In* Proc. of the Wisconsin Fertilizer, Aglime, and Pest Management Conf., Madison, WI. Univ. of Wisconsin, Madison.

Popp, M., M. Rudstrom, and P. Manning. 2005. Spatial yield risk across region, crop and aggregation method. Can. J. Agric. Econ. 53:103–115.

Prior, J.C. 1991. Landforms of Iowa. Univ. of Iowa Press, Iowa City.

Riha, S.J., D.S. Wilks, and P. Simoens. 1996. Impact of temperature and precipitation variability on crop model predictions. Clim. Change 32:293–311.

Sanogo, S., and X.B. Yang. 2001. Relation of sand content, pH, and potassium and phosphorus nutrition to development of sudden death syndrome in soybean. Can. J. Plant Pathol. 23:174–180.

SAS Institute. 2005. Statistical Analysis Software. SAS Inst., Cary, NC.

Schepers, A.R., J.F. Shanahan, M.A. Liebig, J.S. Schepers, S.H. Johnson, and A. Luchiari, Jr. 2004. Appropriateness of management zones for characterizing spatial variability of soil properties and irrigated corn yields across years. Agron. J. 96:195–203.

Secchi, S., and B. Babcock. 2007. Impact of high corn prices on conservation reserve program acreage. http://www.card.iastate.edu/iowa_ag_review/spring_07/IAR.pdf. Iowa Ag Rev. 13:4–5, 7.

Si, B.C., and R.E. Farrell. 2004. Scale-dependent relationships between wheat yield and topographic indices: A wavelet approach. Soil Sci. Soc. Am. J. 68:577–587.

Stöckle, C.O., M. Donatelli, and R. Nelson. 2003. CropSyst, a cropping systems simulation model. Eur. J. Agron. 18:289–307.

Stoskopf, N.C. 1985. Cereal Grain Crops. Reston Publishing, Reston, VA.

Strahler, A.N., and A.H. Strahler. 1984. Elements of physical geography. John Wiley & Sons, New York.

Timlin, D.J., Ya. Pachepsky, V.A. Snyder, and R.B. Bryant. 1998. Spatial and temporal variability of corn grain yield on a hillslope. Soil Sci. Soc. Am. J. 62:764–773.

Tobias, R.D. 2007. An introduction to partial least squares regression. Available at http://support.sas.com/rnd/app/papers/pls.pdf (verified 23 Apr. 2008). SAS Inst., Cary, NC.

Tokatlidis, I.S., and S.D. Koutroubas. 2004. A review of maize hybrids' dependence on high plant populations and its implication for crop yield stability. Field Crops Res. 88:103–114.

Tollenaar, M., and E.A. Lee. 2002. Yield potential, yield stability, and stress tolerance in maize. Field Crops Res. 75:161–169.

Troeh, F.R., and L.M. Thompson. 1993. p. 140–141. *In* Soils and soil fertility. Oxford Univ. Press, New York.

USDA–ERS. 2007. State fact sheets: Iowa. Available at http://www.ers.usda.gov/StateFacts/IA.htm (verified 23 Apr. 2008). USDA, Economic Research Service, Washington, DC.

USGS. 1999. National elevation dataset. Available at ftp://ftp.igsb.uiowa.edu/GIS_Library/IA_State/Elevation (verified 23 Apr. 2008). USGS, Natural Resources Geographic Information Systems Library, Geographic Information Systems Support and Research Facility, Iowa State Univ., Ames.

USGS. 2006. National elevation dataset. Available at http://ned.usgs.gov/Ned/faq.asp#vertical (verified 23 Apr. 2008). USGS, Reston, VA.

Vågen, T.-G., K.D. Sheperd, and M.G. Walsh. 2006. Sensing landscape level change in soil fertility following deforestation and conversion in the highlands of Madagascar using Vis-NIR spectroscopy. Geoderma 133:281–294.

van der Voet, H. 1994. Comparing the predictive accuracy of models using a simple randomization test. Chemometr. Intell. Lab. Syst. 25:313–323.

Waltman, W.J., S.W. Waltman, D.A. Mortenson, and D.A. Miller. 2004. A new geography of agriculture: Understanding patterns, trends, and processes at multiple spatial and temporal scales. p. 262–276. *In* S. Stamatiadis, J.M. Lynch, and J.S. Schepers (ed.) Remote sensing and the environment. Peripheral Editions, Larissa, Greece.

Wassenaar, T., P. Lagacherie, J.-P. Legros, and M.D.A. Rounsevell. 1999. Modelling wheat yield responses to soil and climate variability at the regional scale. Clim. Res. 11:209–220.

Widrlechner, M.P. 1999. NC United States moisture balance map. Available at http://www.ars.usda.gov/Main/docs.htm?docid=9815&page=4 (verified 23 Apr. 2008). USDA–ARS, Washington, DC.

Wigley, T.M.L., and T. Qipu. 1983. Crop-climate modeling using spatial patterns of yield and climate: Part 1. Background and an example from Australia. Am. Meteorol. Soc. 22:1831–1841.

Williams, C.L., W.W. Hargrove, M. Liebman, and D.E. James. 2008. Agro-ecoregionalization of Iowa using multivariate geographical clustering. Agric. Ecosyst. Environ. 123:161–174.

Wilson, M. 2007. Statistical learning: Principle components and partial least squares regression. Available at http://mc.stat.ufl.edu/STA%204702/Spring%2007/MVPrinComp.pdf (verified 17 September 2007). Dep. of Statistics, Univ. of Florida, Gainesville.

Wold, S., M. Sjöström, and L. Eriksson. 2001. PLS-regression: A basic tool of chemometrics. Chemometr. Intell. Lab. Syst. 58:109–130.

Wood, S., L. You, and X. Zhang. 2004. Spatial patterns of crop yields in Latin America and the Caribbean. Cuad. Econ. 41:361–381.