5-2011

APPLICATION OF GIS AND GEOGRAPHICALLY WEIGHTED REGRESSION TO EVALUATE THE SPATIAL NON-STATIONARITY RELATIONSHIPS BETWEEN PRECIPITATION VS. IRRIGATED AND RAINFED MAIZE AND SOYBEAN YIELDS

Vivek Sharma
University of Nebraska-Lincoln, vivekpathankot@gmail.com

Ayse Kilic
University of Nebraska-Lincoln, akilic@unl.edu

Isa Kabenge
University of Nebraska-Lincoln, kabenge1@hotmail.com

Suat Irmak
University of Nebraska-Lincoln, suat.irmak@unl.edu

Application of GIS and Geographically Weighted Regression to Evaluate the Spatial Non-Stationarity Relationships between Precipitation vs. Irrigated and Rainfed Maize and Soybean Yields

V. Sharma, A. Irmak, I. Kabenge, S. Irmak

Abstract. *Understanding the relationship between the spatial distribution of precipitation and crop yields on large scales (i.e., county, state, regional) while accounting for the spatial non-stationarity can help managers to better evaluate the long-term trends in agricultural productivity to make better assessments in food security, policy decisions, resource assessments, land and water resources enhancement, and management decisions. A relatively new technique, geographically weighted regression (GWR), has the ability to account for spatial non-stationarity with space. While its application is growing in other scientific disciplines (i.e., social sciences), the application of this new technique in agricultural settings has not been practiced. The geographic information system (GIS), along with two different statistical techniques [GWR and conventional ordinary least square regression (OLS)], was utilized to analyze the relationships between various precipitation categories and irrigated and rainfed maize and soybean yields for all 93 counties in Nebraska from 1996 to 2008. Precipitation was spatially interpolated in ArcGIS using a spline interpolation technique with zonal statistics. Both measured and GWR- and OLS-predicted yields were correlated to spatially interpolated annual (January 1 to December 31), seasonal (May 1 to September 30), and monthly (May, June, July, August, and September) precipitation for each county. Statewide average annual precipitation in Nebraska from 1996 to 2008 was 564 mm, with a maximum of 762 mm and minimum of 300 mm. Mean precipitation decreased gradually from May to September during the growing season. County average yields followed the same temporal trends as precipitation. When the OLS regression model was used, there was a general trend of linear correlation between observed yield and long-term average mean annual total precipitation with a varying coefficient of determination (R^2). For rainfed crops, 67% of the variability in mean yield was explained by the mean annual precipitation. About 23% and 17% of the variability in mean yield was explained by mean annual precipitation for irrigated maize and soybean, respectively. However, the performance of the GWR technique in predicting the yields from spatially interpolated precipitation for irrigated and rainfed maize and soybean was significantly better than the performance of the OLS model. For both rainfed maize and soybean, 77% to 80% of the variation in yield was explained by the mean annual precipitation alone. For irrigated crops, 42% of the variation in the yield was explained by the mean annual precipitation. For rainfed crops, there was a strong correlation between seasonal precipitation and yield, with R^2 values of 0.73 and 0.76 for maize and soybean, respectively. The mean annual total precipitation was a better predictor of rainfed maize yield than rainfed soybean yield. On a statewide average, July precipitation as a predictor had the greatest correlation with yields of both maize and soybean. June, July, and August precipitation had greater impact on maize yield than on soybean under rainfed conditions due to more sensitivity of maize to water stress than soybean. For irrigated yields, July precipitation had more impact on soybean yield than on maize. The performance of the GWR technique was superior to the OLS model in analyzing the relationship between yield and precipitation. The superiority of the GWR technique to OLS is mainly due to its ability to account for the impact of spatial non-stationarity on the precipitation vs. yield relationships.*

Keywords. *Geographically weighted regression, GIS, Maize, Ordinary least square regression, Precipitation, Soybean, Spline.*

One of the important initial steps for evaluating the variation in yield on a regional scale is to understand the relationships between yield and various environmental factors. Precipitation is one of the main drivers of crop productivity, and its variability in many agricultural production areas, including Nebraska, is significant. Both irrigated and rainfed maize and soybean lands comprise a significant portion of the total cultivated lands in Nebraska. The land area for these crops in 2002 and 2007 (table 1) shows changes in the last five years, with maize pro-

Submitted for review in October 2010 as manuscript number SW 8860; approved for publication by the Soil & Water Division of ASABE in May 2011.

The mention of trade names or commercial products is for the information of the reader and does not constitute an endorsement or recommendation for use by the University of Nebraska-Lincoln or the authors.

The authors are **Vivek Sharma, ASABE Member,** Graduate Research Assistant, Department of Biological Systems Engineering, **Ayse Irmak,** Assistant Professor, School of Natural Resources Center for Advanced Land Management Information Technologies (CALMIT) and Department of Civil Engineering, **Isa Kabenge, ASABE Member,** Graduate Research Assistant, Department of Biological Systems Engineering, and **Suat Irmak, ASABE Member,** Associate Professor, Department of Biological Systems Engineering, University of Nebraska-Lincoln, Lincoln, Nebraska. **Corresponding author:** Suat Irmak, Department of Biological Systems Engineering, University of Nebraska-Lincoln, 241 L.W. Chase Hall, Lincoln, NE 68583-0726, phone: 402-472-4865; e-mail: sirmak2@unl.edu.

  ISSN 2151-0032

Table 1. Irrigated and rainfed maize and soybean land area in 2007 and 2002 (USDA-NASS, 2010).

Crop Type	Land Area (ha) in 2007			Land Area (ha) in 2002		
	Total	Irrigated	Rainfed	Total	Irrigated	Rainfed
Maize for grain	1,505,466	956,254	549,212	1,202,832	737,871	464,961
Soybean for grain	628,028	257,134	370,894	748,676	314,937	433,739
Maize for silage	30,393	16,582	13,810	67,054	25,490	41,564

duction being dominant, covering over 1.5 million ha. While extensive irrigation practices provide stable and high-yielding conditions on 3.4 million ha of land (USDA-NASS, 2007), rainfed crop production also represents a significant portion of agricultural productivity in Nebraska. Rainfed yields show significant spatial and temporal variability due to variability in both precipitation totals and distributions during the growing season. In 2007, more than 900,000 ha of rainfed maize and soybean crops were harvested, with significant variability in crop productivity between the 93 counties (USDA-NASS, 2007). Annual precipitation in Nebraska also varies significantly spatially and temporally, ranging from 800 mm in the southeastern part of the state to less than 350 mm in the western portion in the panhandle on a long-term average basis. Practical tools and methods are needed for mapping, analyzing, and more importantly, predicting irrigated and rainfed crop yields and understanding the interactions between yield and primary climatic variables (i.e., precipitation) for decision makers to better plan, manage, and allocate natural resources for crop production on large scales.

The relationship between precipitation and yield is critical for understanding present water needs, ensuring the availability of water for future use, and evaluating the sustainability of production systems. Regional variation in yield can be large because of the differences in irrigated and rainfed productivity. Relationships between yield and climate variables on a spatial and temporal scale are extremely important to make better assessments of food security, policy, and land and water resources management decisions on large (watershed, statewide, regional) scales. Variability in precipitation not only impacts productivity but also increases spatial yield variability due to the interactions of precipitation with other field properties, such as soil and terrain characteristics. For example, Kaspar et al. (2003) reported that when precipitation was lower than the average, maize yield showed negative correlations with elevation, slope, and soil curvature, and when precipitation was above average, the yield was positively correlated with those parameters. Kravchenko et al. (2005) found that the coefficient of variation in yield increased in years with low precipitation (45% below normal) and decreased in years with high precipitation (14% above normal).

Various crop simulation models, such as crop-specific CERES-Maize (Jones and Kiniry, 1986) and Hybrid-Maize (Yang et al., 2004), and generic models that include multiple crops, such as DSSAT (Jones et al., 2003), are used to predict crop yield by integrating various characteristic such as weather, physical and/or chemical soil properties, genetic background, management, and other agronomic practices that operate at uniform or non-uniform areas on a field scale (Hansen and Jones, 2000; Irmak et al., 2001, 2002, 2005, 2006). While these models provide extremely powerful and useful information on predicting spatial variability of crop yields, they require a substantial number of input parameters due to variability in soil, topographical conditions, weather, and management practices at a regional scale. Several investigators have demonstrated the strength of coupling crop models with a geographic information system (GIS) for agricultural decision support and planning at various spatial scales (Dent and Thornton, 1988; Curry et al., 1990; Thornton, 1991; Sarangi et al., 2005). Hansen and Jones (2000) demonstrated several approaches to scale-up crop model predictions to large scales. Other researchers (e.g., Lal et al., 1993; Thornton et al., 1995; Carbone et al., 1996; Rosenthal et al., 1998) applied various crop models in regional estimation of crop yield and variability. With the advances in GIS and remote sensing techniques, utilization of these models can be extended to broader spatial scales. This makes yield predictions more practical and useful for regional planning and management, and for application at various spatial scales (Dent and Thornton, 1988; Curry et al., 1990; Moen et al., 1994). For example, application of GIS for predicting and mapping yield can enable state and federal agency personnel to better assess the spatial and temporal information of crop yields and their interactions with weather parameters. This would enable them to better understand the magnitude and variability of crop productivity between different locations (counties) so that priority areas with limited productivity can be identified and resource allocations to improve productivity could be allocated or restructured.

One commonly used method to predict or map crop yields at a regional scale is interpolation of point-based information to larger scales (Hoogenboom and Thornton, 1990; Thornton and Baanante, 1991). Other variables such as soil and weather parameters that have impact on yield need to be interpolated for mapping regional crop productivity. Various interpolation techniques are available to predict and interpolate information or variables within pre-determined boundaries. Local interpolation techniques usually use weather data from local stations and predict the variables of interest at a given point by using data from nearby stations and mathematical functions (Borrough and McDonnell, 1998). For example, geostatistical techniques (Webster, 1985; Oliver and Webster, 1991; McBratney and Pringle, 1997; Oliver, 1999) enable spatial correlation between the data from various observations. These techniques are similar interpolation techniques to those used with minimum spatial variance (Curran et al., 1997; Curran and Atkinson, 1998). Among many other examples, Lal et al. (1993), Gessler et al. (2000), Mueller and Pierce (2003), Li et al. (2004), Sarangi et al. (2005), Flowers et al. (2005), Lauzon et al. (2005), Kravchenko and Robertson (2007), McKinion et al. (2010), and Irmak et al. (2010) have used various interpolation techniques in combination with GIS to spatially interpolate weather, soil physical and chemical properties (including soil moisture, nutrients, pH, and soil carbon), terrain (slope and elevation), plant properties, and other parameters to predict the impact of these parameters on crop yields and yield variability and to analyze regional productivity.

An important catalyst for better integration of GIS and spatial data analysis for improved interpolation has been the development of local spatial statistical techniques. One of the newer non-parametric modeling methods that provides readily interpretable measures of sources of local variation and reduces variable co-linearity is the geographically weighted regression (GWR) technique. GWR is among the new developments of local spatial analytical techniques. GWR (Fotheringham et al., 1998; Brunsdon et al., 2001; Fotheringham et al., 2002) is a local spatial statistical technique that relies on a form of kernel regression within a multiple linear regression framework to develop local relationships between the dependent and independent variables. When modeling the spatial relationship between weather variables and crop and surface characteristics (e.g., precipitation vs. yield; precipitation vs. soil moisture; vegetation indices vs. radiation or temperature, etc.), the spatial non-stationarity of these relationships has to be taken into account. Spatial non-stationarity is indicated when the measurement or estimation of the relationships among variables differs from location to location (Mennis, 2006). It indicates that the relationship between the variables under study varies from one location to another depending on physical factors of the environment, which are spatially autocorrelated (Propastin et al., 2007). The development of local spatial non-stationary relationships between the variables facilitates an exploratory analysis of the stationary assumption of a global multiple linear regression model. The commonly used regression models, including conventional ordinary least squares (OLS), do not account for the spatial non-stationarity of variables. Unlike OLS and other conventional regression models, which provide a single regression equation to describe general relationships among the explanatory and dependent variables, GWR generates spatial data that express the spatial variation in the relationships among variables. Thus, maps generated from these analyses play a key role in describing and interpreting spatial non-stationarity between the variables (Mennis, 2006).

Recently, a limited number of studies have demonstrated the analytical utility of GWR for investigating a variety of scientific areas, including climatology (Brunsdon et al., 2001), ecological inference problem (Calvo and Escolar, 2003), urban poverty (Longley and Tobon, 2004), environmental justice (Mennis and Jordan, 2005), population density vs. home value relationships (Mennis, 2006), and developing relationships between vegetation (normalized difference vegetation index, NDVI) and climate (precipitation) (Propastin et al., 2007). While these studies showed compelling advantages of using the GWR technique in these topic areas, the advantages of using the GWR technique in mapping yield and other variables that can be spatially interpolated has not yet been studied. The objectives of this research are to: (1) interpolate and map spatial distribution of long-term irrigated and rainfed maize and soybean yields for all 93 counties in Nebraska using GWR, OLS, and GIS (spline interpolation technique) and assess the variability of irrigated and rainfed yields with respect to precipitation; (2) develop relationships between monthly (May, June, July, August, or September), seasonal (May 1 to September 30), and annual (January 1 to December 31) precipitation and crop yields to determine both irrigated and rainfed crop yield response to the magnitude of the three precipitation categories; and (3) investigates the potential differences in using GWR over OLS in predicting crop yields on a county scale.

Materials and Methods

Study Area

The study area is the state of Nebraska (latitude 40° N to 43° N, longitude 95° 19′ W to 104° 3′ W) (figs. 1a and 1b), which has 93 counties with population of 1,796,620 (mean population density of about 9 people km^{-2}). The state's ground and surface water resources are regulated by the Nebraska Department of Natural Resources and 23 natural resources districts. Nebraska is one of the leading farming and ranching states in the U.S. The total area of the state is approximately 200,356 km^2, with a mean elevation of 793 m above mean sea level. Nebraska's climate is mainly continental and is divided into two main parts: the eastern and central parts are humid/sub-humid continental climate, and the

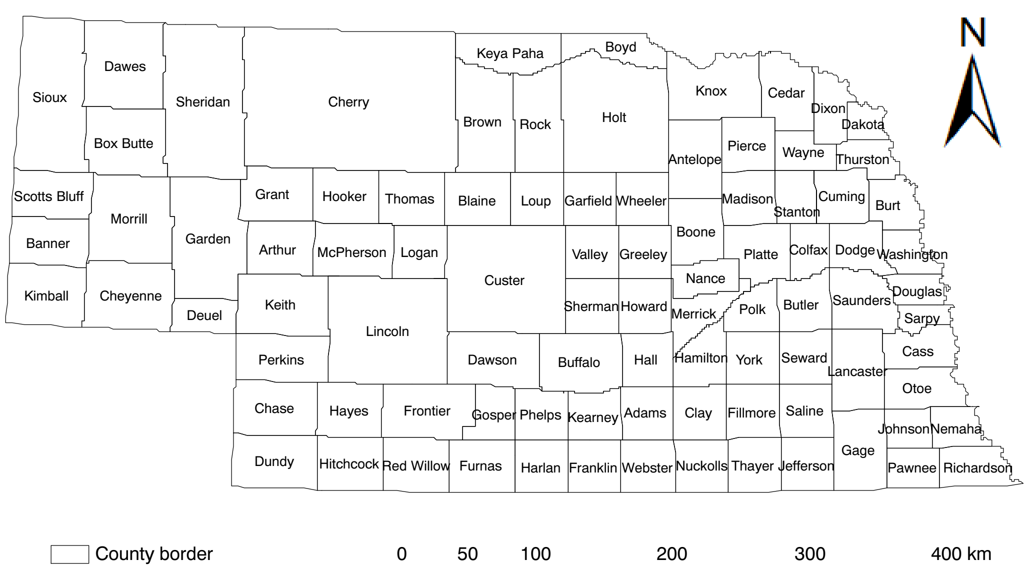

Figure 1a. Nebraska county borders.

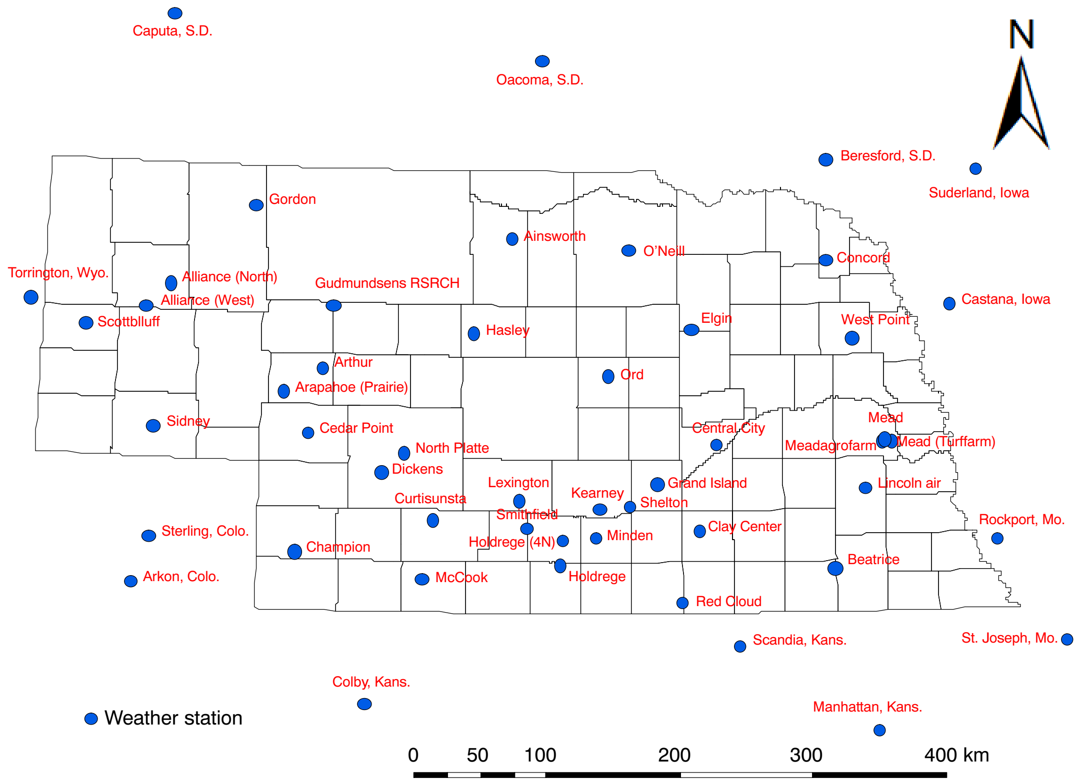

Figure 1b. Locations of the High Plains Regional Climate Center Automated Weather Data Network (HPRCC-AWDN) used to run the interpolation technique (i.e., radial basis function, RBF).

western third has a semiarid/arid climate. The state experiences a wide range of seasonal variation in temperature and precipitation. In general, the weather in Nebraska is influenced by cold dry continental air masses from Canada during winter and warm moist air from the Gulf of Mexico during summer. The highest wind speed usually occurs from January to late May and early June, with daily average wind speed showing significant fluctuation ranging from 2 m s^{-1} to over 8 m s^{-1}. The lowest wind speeds usually occur in the summer months. Summer months are usually hot and humid, averaging 24°C in July, but hot, dry winds often drive summer temperatures above 32°C (Irmak, 2010).

Regional differences in the environmental characteristics, with the combined effect of climatic conditions and soil and topographic characteristics, divides Nebraska into three broad environmental regions: the eastern region is characterized by relatively high amounts of precipitation with superior silt-loam soils rich in organic matter content (i.e., >2.5%) and relatively high agronomic productivity; the central region is generally characterized by rich silt loam soils with high organic matter content with generally flat topography (except sand hills) and moderate precipitation; and the western region has less precipitation and soils with relatively lower potential for agronomic productivity as compared with the eastern and central regions (Searcy and Longwell, 1964). Sand hills area cover a large area of over 50,000 km^2 of sandy soils and dunes with mainly sparse grasslands. The sand hills area covers 21 counties (Cherry, Grant, Hooker, Thomas, Keya Paha, Brown, Rock, Boyd, Holt, Blaine, Loop, Garfield, Wheeler, Arthur, McPherson, Logan, Custer, Valley, Greeley, Sherman, and Howard) in the north and north-central part of the state (figs. 1a and 1b). Water availability is the dominant yield-reducing factor in western Nebraska, where irrigation is necessary for producing average or high yields.

USDA-NASS Crop Yield Data

The USDA National Agricultural Statistical Service (www.nass.usda.gov) is an important source of information for long-term yield data as the service provides yield data predictions almost for every county in the U.S. Yield data for irrigated and rainfed maize and soybean crops were obtained for the 93 counties in Nebraska from 1996 to 2008 from the USDA-NASS website. Yield was defined as the county average yield for either rainfed or irrigated crops as reported. The USDA-NASS does not take into account the average irrigation amount applied per growing season; therefore, irrigation and yield relationships were not included in our study. County yield was averaged across the 13 years (1996 to 2008). Irrigated and rainfed crop yields were used as the dependent variables in the GWR model. The county yield was defined as total harvested yield in kg divided by the total area of harvested yield per year per county, so we expressed yield as kg ha^{-1}. Some of the counties in Nebraska did not report maize or soybean production from 1996 to 2008. In some parts of Nebraska (northwest and northern edge of the state near the sand hills), maize and soybean are not produced. In some counties, there were incomplete or missing yield data. Therefore, the counties with missing data were excluded from our analysis. The number of counties (N) included were 91, 89, 80, and 79 for irrigated maize, rainfed maize, irrigated soybean, and rainfed soybean, respectively.

PRECIPITATION DATA

While other weather variables (i.e., solar radiation, air temperature) have significant impact on crop productivity, precipitation is the main yield-limiting factor in much of Nebraska and was selected as the primary variable that impacts yield and yield variability in this study. Precipitation data were obtained from the High Plains Regional Climate Center (HPRCC; http://hprcc1.unl.edu/cgi-hpcc/home.cgi) automated weather data network (AWDN) for 51 AWDN stations distributed across the state (fig. 1b). To increase the precipitation data density and the robustness of the analyses, some weather stations outside of Nebraska were also used to interpolate precipitation across the boundaries of the Nebraska counties. The stations outside of Nebraska were also part of the HPRCC. Some of the counties did not have weather stations, so a spline interpolation procedure was used to calculate the precipitation for all counties, as described in the next section. Precipitation data were defined in three categories: (1) long-term monthly (May, June, July, August, or September) average precipitation from 1996 to 2008, (2) long-term average seasonal (growing season) total precipitation from 1996 to 2008, and (3) long-term annual total precipitation from 1996 to 2008. Thus, yield vs. precipitation relationship analyses were conducted for all three categories to evaluate both irrigated and rainfed crop yield response to the magnitudes of the three precipitation categories. The growing season was considered from May 1 to September 30, which is typical for maize and soybean production in the region. The major assumption here is that the growing season was assumed to be the same across the state. Other assumptions include similar maize hybrids and soybean varieties across Nebraska. Disease and weed pressure or any other field management issues, such as nitrogen deficiency, that may cause yield reduction were not taken into account in the analyses.

SPATIAL INTERPOLATION OF PRECIPITATION

The spline method was used to predict the spatial distribution of precipitation across Nebraska. This method is a deterministic interpolation method that fits a mathematical function through input data to create a smooth surface. It can generate accurate surfaces from only a few sampled points. These functions allow users to decide between smooth curves or tight straight edges between measured points. Each station is omitted in turn from the estimation of the fitted surface, and the mean square error is calculated. The mean square error calculations are repeated for a range of values of a smoothing parameter, and the value that minimizes the mean square error is used to determine the optimum smoothing. This process is called minimizing the generalized cross-validation (GCV). A regularized spline was selected for the analyses because it creates smoother surface closely constrained with sample data range. The following form of the generalized spline function (Franke, 1982) was used:

$$S(x,y) = T(x,y) + \sum_{j=1}^{N} \lambda_j R(r_j) \tag{1}$$

where T is the constant trend, r_j is the distance from point (x, y) to the jth point, and R is a weighted function of the distance between the interpolated point and the jth data point (j = 1, 2, 3, ..., N, where N is the number of known points, i.e., counties). For the spline function, T and R are defined as:

$$T(x,y) = a_1 + a_2 x + a_3 y \tag{2}$$

$$R(r) = \frac{1}{2\pi}\left\{\frac{r^2}{4}\left[\ln\left(\frac{r}{2\pi}\right) + c - 1\right] + \tau^2\left[K_o\left(\frac{r}{\tau}\right) + c + \ln\left(\frac{r}{2\pi}\right)\right]\right\} \tag{3}$$

where τ is a weight parameter that varies between 0 and 0.5 with higher values representing smoother surfaces, r is the distance between the point and the sample, K_o is a modified Bessel function, and c is Euler's constant (0.577215). Coefficients a_1, a_2, and a_3 in equation 2 are found by the solution of a system of linear equations. The weight parameter was optimized using ArcGIS (ver. 9.3.1). In ArcGIS, the spline is the radial basis function (RBF). The spline interpolation was used to interpolate the precipitation data from January 1986 to December 2009 from 51 weather stations across Nebraska and surrounding areas using ArcGIS.

GWR MODEL

The relationship between the yield and precipitation was modeled by using the OLS regression and GWR. The model derived by the OLS regression was assumed to apply globally to the entire study region, from which measured data have been taken, based on the assumption of spatial stationary in the relationship between the variables under study (Foody, 2003). Spatial stationary assumes that statistical properties of an attribute are independent of a location and that the mean and variance of observed attribute values at different locations across the study region are constant. For example, precipitation might not vary across a small area. However, if there is spatial non-stationarity, then the global prediction of spatial relationships using OLS regression will misrepresent the relationship between precipitation and yield. Therefore, the relationship between these two variables is also examined with the GWR technique. GWR is a local spatial statistical technique used to analyze spatial non-stationarity when the input variable differs from location to location. It provides a local model to predict an independent variable or process (e.g., plant growth, yield) by fitting a regression equation to the available datasets of dependent variables. It enables identification of the yield stability of a region as well as the association of the independent environmental factors to the yield. The main advantage of GWR over OLS regression is its ability to deal with spatial non-stationarity (Propastin et al., 2007). Global regression techniques such as OLS may ignore local information and, therefore, indicate incorrectly that a large part of the variance in yield was unexplained.

GWR is a useful and practical tool for evaluating the spatial heterogeneity of a dependent variable. Spatial heterogeneity can exist when the structure of the process being modeled varies across the study area (county or state). The GWR method constructs separate equations by incorporating the dependent and explanatory variables of features falling within the bandwidth (distance) of each target feature. The shape and size of the bandwidth is dependent on user input for the kernel type, bandwidth method, distance, and number of features. Instead of calibrating a single equation, GWR generates a separate regression equation for each observation (i.e., spatially interpolated precipitation and yield) and thus

Table 2. Statewide average intercept (β_o) and linear coefficient (β_1) for yield vs. precipitation.

	Annual		Seasonal		May		June		July		August		September	
Crop	β_o	β_1	β_o	β_1	β_o	β_1	β_o	β_1	β_o	β_1	β_o	β_1	β_o	β_1
Rainfed maize	-2408	13.26	-3318	23.1	-856.4	69.1	-3096	98.7	2616.6	43.3	487.8	69.7	4957.1	11.3
Rainfed soybean	-828.8	5.05	-828.8	5.05	-265.6	26.5	-718.8	33.4	1138.6	16.5	537.1	23.8	1566.4	12.9
Irrigated maize	9302.9	2.18	8798.6	4.60	9273.2	13.7	9848.7	8.5	8857.8	24.5	9561.2	14.7	10538.9	-1.1
Irrigated soybean	3007.7	0.83	2683.1	2.09	2914.6	6.1	3470.4	0.45	2528.6	13.9	3156.7	4.9	3097.7	6.7

allows parameter values to vary continuously in geographical space. Each equation is calibrated using a different weighting of the observations contained in the dataset. Because the regression equation is calibrated independently for each observation, a separate parameter prediction (z-value) and R^2 value are calculated for each observation. In GWR, a form of kernel regression and multi-linear regression was used to build a model that can be generally expressed as:

$$Y_i = \beta_{o(i)} + \beta_{1(i)}X_{1(i)} + \beta_{2(i)}X_{2(i)}$$

$$+ \beta_{3(i)}X_{3(i)} + \text{K} + \beta_{k(n)}X_{K(n)} + \varepsilon_i \quad (4)$$

where $Y_{(i)}$ is the interpolated yield (kg ha^{-1}) at location i (where i captures the coordinate location), β_o is the intercept, $\beta_{k(i)}$ is the kth local parameter prediction at the ith location (i.e., coefficient for the independent variable), $X_{K(i)}$ is the kth independent variable (precipitation) value (mm) at the ith location, and n represents the last location. In GWR, the weight assigned to each observation is based on a distance decay function centered on observation i. The distance decay function, which may take a variety of forms, is modified by a bandwidth setting at which distance the weight rapidly approaches zero (Mennis, 2006). The coefficients in equation 4 vary with each county and period under consideration (annual, seasonal, monthly), and they are not presented in this article. The statewide average values for β_o and β_1, are shown in table 2.

A spatial kernel is used to provide the geographic weighting in the model. A key coefficient in the kernel is the bandwidth, which controls the size of the kernel. Bandwidths can be considered as smoothing functions of the local parameter predictions (Fotheringham et al., 2002). A fixed bandwidth is based on a defined diameter of a circular search neighborhood. The diameter scalar units are the same as the location variables. There are different choices for the bandwidth, including the Akaike Information Criterion (AIC). The complexity of a GWR model depends not only on the number of variables in the model but also on the bandwidth. This interaction between the bandwidth and the complexity of the model is the reason for our preference for the AIC over the other bandwidths. The geographic weighting occurs once a regression model (Gaussian, Logistic, or Poisson), bandwidth, and kernel type are selected. The local parameter predictions are derived from the regression of data points within the bandwidth of a kernel. The influence of a data point on the local parameter prediction is weighted on the basis of the geographic distance from the regression point. Locations that are close to the regression point of interest are weighted heavier than points located farther away (Fotheringham et al., 2002).

Statistical Analyses

Mean annual, growing season, and monthly precipitation values for each county are essential requirements of this study. The zonal statistic was used to calculate the precipitation values for each county defined by name (string attribute field) of the Nebraska county feature class based on the precipitation value from the precipitation raster dataset. The zonal statistics tool (Spatial Analyst tool of ArcGIS ver. 9.3.1) calculates statistics on the value of a raster within the zone of another dataset. The zonal statistic tool summarizes the value of the precipitation raster within the county and reports the result as the mean, maximum, minimum, and range values. Some studies have used zonal statistics for computing the average elevation, aspect, slope (topographic attributes), and NDVI (Bakhash and Kanwar, 2004; Tiwari and Sharma, 2009). Other studies have used zonal analysis to calculate the soybean yield for different grids (Kulkarni et al., 2008).

Moran's I index of spatial autocorrelation (Fortin and Dale, 2005) in ArcGIS was used to determine whether the pattern of mean yield among counties was randomly distributed, evenly distributed, or clustered. It is possible that low mean yield can occur in a stable system (Berzsenyi et al., 2000; Mead et al., 1986). However, we hypothesized that the mean county yields would be randomly distributed. The geographic distribution of the mean yield of irrigated and rainfed maize and soybean were clustered according to Moran's I test (high z-value and low p-value). The GWR-predicted mean yields were also analyzed using Moran's I test to determine if the predicted yields were significantly different from the observed yields at the 5% significance level.

Results and Discussion

The statewide average β_o and β_1 coefficients from equation 4 for annual, seasonal, and individual months (May, June, July, August, and September) are presented in table 2. The GWR approach produced four β_o and four β_1 coefficients (one for each irrigated and rainfed maize and irrigated and rainfed soybean) for each county as a result of producing a regression equation for each of the yield vs. precipitation relationships. Irrigated maize was produced in 91 of the 93 counties, and rainfed maize was produced in 89 of the 93 counties. Irrigated soybean was produced in 80 counties, and rainfed soybean was grown in 75 counties. Thus, a total of 364 and 356 β_o and β_1 values were calculated for irrigated maize and rainfed maize, respectively, and a total of 320 and 300 β_o and β_1 values were calculated for irrigated and rainfed soybean, respectively. These values were calculated for each precipitation category (annual, seasonal, and individual months) vs. yield relationship. As a result, a total of 9,380 β_o values and 9,380 β_1 values were calculated for each crop and each precipitation vs. yield category. Therefore, only the statewide average β_o and β_1 coefficients are presented in table 2.

Precipitation Amounts and Distribution

The spatial distributions of long-term average annual (January 1 to December 31) precipitation and long-term

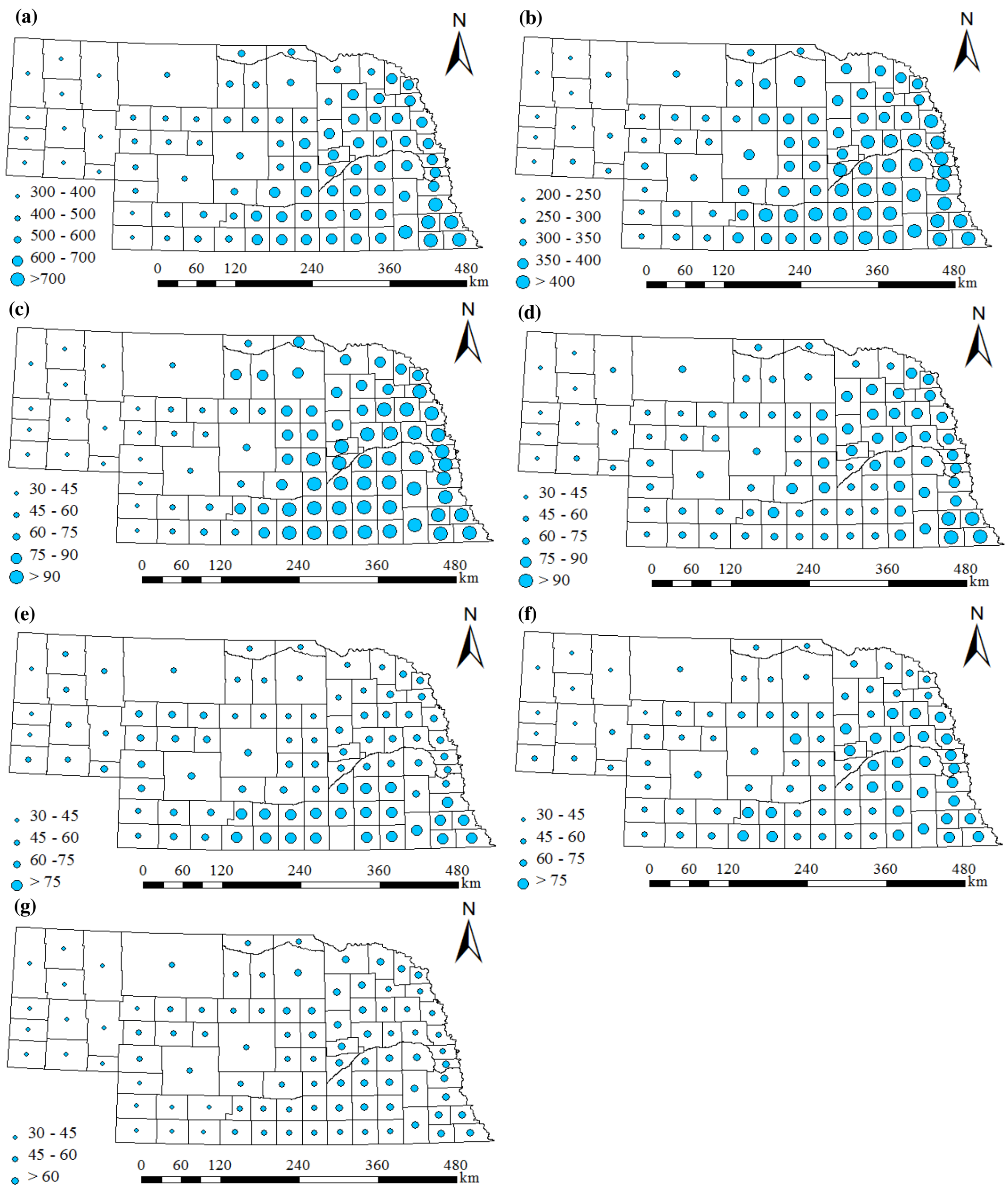

Figure 2. Variation of long-term average (1996-2008) (a) annual, (b) seasonal, (c) May, (d) June, (e) July, (f) August, and (g) September precipitation (mm) across Nebraska.

Table 3. Precipitation (P) statistics for the observation period of 1996-2008 for 93 counties in Nebraska.

Parameter	Mean (mm)	Max (mm)	Min (mm)	County with Max P	County with Min P
Long-term average annual P	564	762	301	Richardson	Scotts Bluff
Long-term average growing season P	364	469	197	Richardson	Scotts Bluff
Long-term average May P	86	111	44	Lancaster	Scotts Bluff
Long-term average June P	83	116	45	Richardson	Scotts Bluff
Long-term average July P	67	88	40	Nuckolls	Scotts Bluff
Long-term average August P	67	92	33	Richardson	Sioux
Long-term average September P	54	67	33	Richardson	Sioux

average growing season (May 1 to September 30) total precipitation for all Nebraska counties are presented in figures 2a and 2b, respectively. Long-term monthly (May, June, July, August, and September) average precipitation for all counties is presented in figures 2c through 2g, respectively. The quantitative values of precipitation are shown as a series of graduated circles arranged from smallest to largest. The precipitation data statistics are presented in table 3. Annual total precipitation showed strong spatial patterns (trends) across the state (fig. 2a). There was a gradual decrease in precipitation totals from the southeast part of the state to the northwest. There was a 400 mm difference in the annual precipitation amounts between the southeast and northwest (see legend in fig. 2a). From the middle of the state to the west, precipitation becomes a limiting factor for crop production. There is also a trend of decreasing precipitation from south to north in the eastern edge of the state. The south-central portion of Nebraska is the most extensively irrigated area in the state, and approximately 75,000 of more than 105,000 irrigation wells are located in the central and south-central parts of the state. Depending on the year, precipitation usually starts becoming a limiting factor for crop production at the western edge of central Nebraska. Based on the annual average precipitation data, there was approximately a 25 mm decrease in precipitation for every 40 km going from east to west. The seasonal total precipitation showed similar patterns as the annual precipitation (fig. 2b vs. fig. 2a).

Average monthly precipitation for the months of May, June, July, August, and September varied substantially across Nebraska. In general, May was the wettest month, with an average of 86 mm and ranging from 110 mm in Lancaster county to 45 mm in Scotts Bluff county (figs. 2c through 2g). Average annual precipitation in the state during the period of 1996 to 2008 was 564 mm, with a maximum of 762 mm in Richardson county and minimum of 300 mm in Scotts Bluff county (table 3, fig. 2a). Mean precipitation decreased gradually from May through September during the growing season. While significant variations exist with location, management practices, and other factors, the typical average seasonal crop water use for maize in south-central Nebraska is about 640 mm (Irmak et al., 2008). With typical and dominant silt-loam soil type in south-central Nebraska, approximately 165 mm of the seasonal crop water use can be supplied from soil moisture available in the 0.90 m (typical maize crop root zone) soil profile from spring precipitation. Thus, if state average data are considered (table 3), about 110 mm [640 - (364 + 165) = 110 mm] is supplied with irrigation. However, soil type and water-holding capacity, crop evapotranspiration amounts, irrigation methods, management practices, and other factors that influence irrigation practices vary significantly across the state. Thus, the crop irrigation requirement also varies substantially not only spatially but also temporally. Since growing season precipitation varies significantly across the state, the crop irrigation requirement also exhibits significant spatial and temporal variability. We are in the process of quantifying long-term average and spatial and temporal variability in irrigation requirement of each county as a follow-up study.

Irrigated and Rainfed Maize and Soybean Yields

The 13-year average rainfed and irrigated maize and soybean yields are presented in figures 3a through 3d. Summary statistics of historical crop yields are provided in table 4. The white-colored counties in figures 3a through 3d indicate counties with no data. There was a general tendency for the rainfed maize and soybean yields to be greater in the northeast part of the state and gradually decrease from the eastern part toward the western part, following the precipitation trends (table 3, figs. 2a and 2b). The highest irrigated maize and soybean yields are both concentrated in the south-central part of the state, where the heaviest irrigation is concentrated.

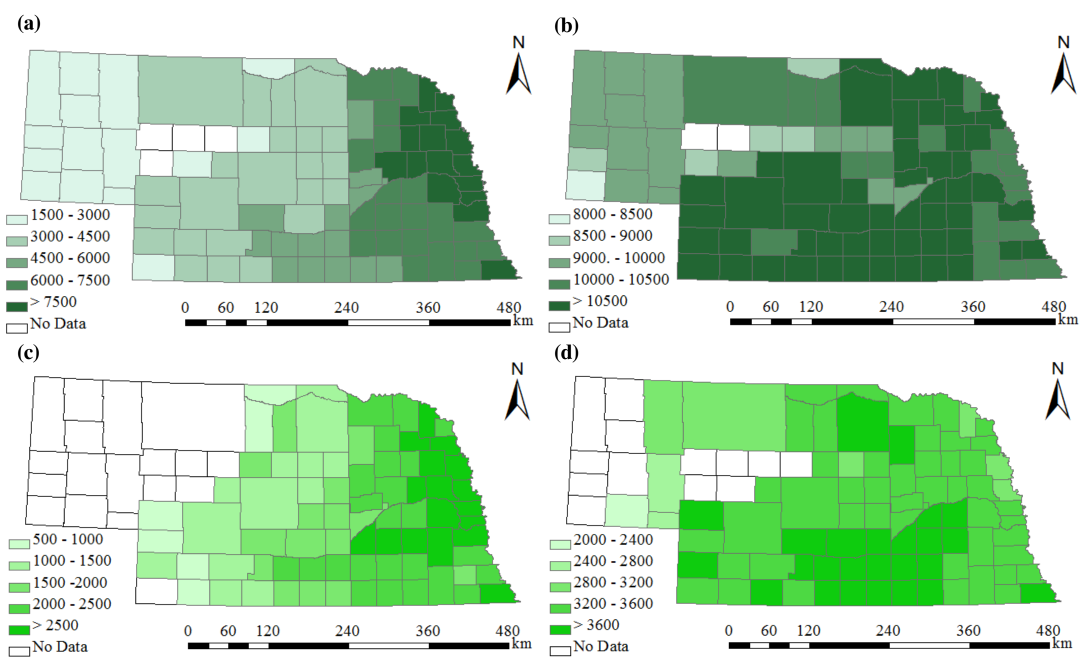

Figure 3. Distribution of long-term (1996-2008) average mean yields (kg ha^{-1}) for (a) rainfed maize, (b) irrigated maize, (c) rainfed soybean, and (d) irrigated soybean across Nebraska counties.

Table 4. Statewide average crop yield (kg ha^{-1}) statistics from 1996 to 2008 in Nebraska (N = number of counties; SD = standard deviation).

Crop	N	Mean	Max	Min	SD	Median	Skewness
Rainfed maize	89	5,410	8,458	1,883	1,727	5,403	-0.069
Rainfed soybean	75	2,215	2,861	639	478	2,354	-0.8514
Irrigated maize	91	10,401	11,661	8,385	692	10,565	-0.9461
Irrigated soybean	80	3,432	3,910	2,098	290	3,435	-1.703

Table 5. Statistics of spatial autocorrelation (Moran's I test) for spatial randomness of the mean yields.

	Mean Rainfed Maize Yield	Mean Irrigated Maize Yield	Mean Rainfed Soybean Yield	Mean Irrigated Soybean Yield
Moran's I (p-value)	0.76	0.39	0.65	0.34
z-value	18.76	8.11	9.68	8.43

Rainfed and irrigated maize and soybean yields varied greatly over the 13-year of study period across Nebraska (long-term yield data not shown). The rainfed maize yield ranged from as high as 8,458 kg ha^{-1} in Dakota county (northeast Nebraska, fig. 1a) to as low as 1,883 kg ha^{-1} in McPherson county (west-central Nebraska), with a mean values of 5,410 kg ha^{-1}. The statewide average irrigated maize yield was 10,401 kg ha^{-1}, with a maximum of 11,661 kg ha^{-1} in Phelps county (south-central Nebraska) and minimum of 8,385 kg ha^{-1} in Kimball county (southwest Nebraska). Interestingly, rainfed soybean yields showed similar spatial patterns as irrigated maize. The statewide average yields for rainfed soybean was 2,215 kg ha^{-1}, with a maximum of 2,861 kg ha^{-1} in Cuming county (northeast Nebraska) and minimum of 639 kg ha^{-1} in Phelps county (south-central Nebraska). The irrigated soybean yield varied from 2,098 kg ha^{-1} in Cheyenne county (western Nebraska) to 3,910 kg ha^{-1} in Hayes county (southwest Nebraska), with a statewide average value of 3,432 kg ha^{-1}. On average, the irrigated soybean yields were about 30% higher than the rainfed soybean yields. Irrigated maize had a much higher (50%) rate of higher yield than rainfed maize, indicating the higher level of susceptibility of maize to water stress.

Yield distribution of all the crops exhibits a negative skewness (table 4), where the number of counties with below-mean yield exceeded the number of counties with above-mean yield. The geographically weighted distribution of long-term average yield for all four crops was significantly clustered with spatial randomness of mean yield (autocorrelation, Moran's I; table 5). Although Nebraska possesses some of the best agricultural soils in the country (in general, deep silt-loam soils with high organic matter content and soil water-holding capacity), the skewness distribution of the mean yields indicates that only several counties produce the highest yields, or there are differences among counties in the amount and use of marginal land. While the actual amount of agronomic crop production on marginal land over the observation period is unknown, the greater amount of marginal land is in the western part of the state. Furthermore, the largest land use class (rainfed grassland/rangeland) covers about 57% of the state (CALMIT, 2005). The main reason for the rainfed maize and soybean yields to be lower in the western part of the state is the lower amount of precipitation as compared with the eastern and central portions. Another reason is that the soils in the western part of Nebraska are a combination of Tassel-Busher and rocky association generally having the characteristics of strongly sloping to steep, excessively drained, weathered sandstone that has lower water-holding capacity and organic matter content as compared with the soils in the west-central, central, and eastern portions. Topography of the eastern and central parts is alluvial lowlands with flat topography and soils with high organic matter content and soil water-holding capacity (i.e., mainly silt loam soils with about 60 mm water per 0.30 m soil depth water-holding capacity). The north-central part of the state, which contains a large portion of the state's total land area, is the sand hills area, which is generally covered by grass/rangelands with very limited farming of agronomical crops.

Precipitation vs. Observed Yield Using OLS Regression

The coefficient of determination (R^2) between observed yield and precipitation (using long-term average mean annual, seasonal, and monthly total precipitation for each county) for irrigated and rainfed maize and soybean using the OLS model are presented in table 6. In the OLS model, two variables that were related to each other were the county average yield and precipitation. The performance of the observed yield vs. long-term average mean annual total precipitation along with the R^2 using OLS shows a good agreement for rainfed maize and soybean. For rainfed crops, 67% of the variability in mean yield is explained by the mean annual precipitation. About 23% and 17% of the variability in mean yield was explained by mean annual precipitation for irrigated maize and soybean, respectively. For rainfed crops, there was a strong correlation between seasonal precipitation and yield, with an R^2 of 0.65 and 0.63 for maize and soybean, respectively. The R^2 was much lower for irrigated crops. As found with the annual total precipitation and yield, the correlation between seasonal precipitation and yield for rainfed crops was stronger than for irrigated crops (table 6). When correlating to the yield, the annual precipitation may have an advantage in that it accounts for the dormant season precipitation that is also carried over and available to the crop as available soil moisture in the beginning of the growing season. In terms of individual-month total precipitation, the R^2 for rainfed maize ranged from 0.57 to 0.67, except for July where the R^2 was 0.34. Similar results were found for rainfed soybean, where R^2 ranged from 0.35 to 0.70, with the exception of August where yield and precipitation had poor correlation (R^2 = 0.20).

Table 6. Relationships between observed yield vs. long-term average mean annual, seasonal, and monthly precipitation by OLS (*N* = number of counties).

Crop	*N*	Coefficient of Determination (R^2) Annual	Seasonal	May	June	July	August	September
Rainfed maize	89	0.67	0.65	0.67	0.60	0.34	0.57	0.57
Rainfed soybean	75	0.67	0.63	0.66	0.36	0.36	0.20	0.41
Irrigated maize	91	0.23	0.27	0.23	0.19	0.30	0.28	0.09
Irrigated soybean	80	0.17	0.20	0.18	0.05	0.22	0.18	0.08

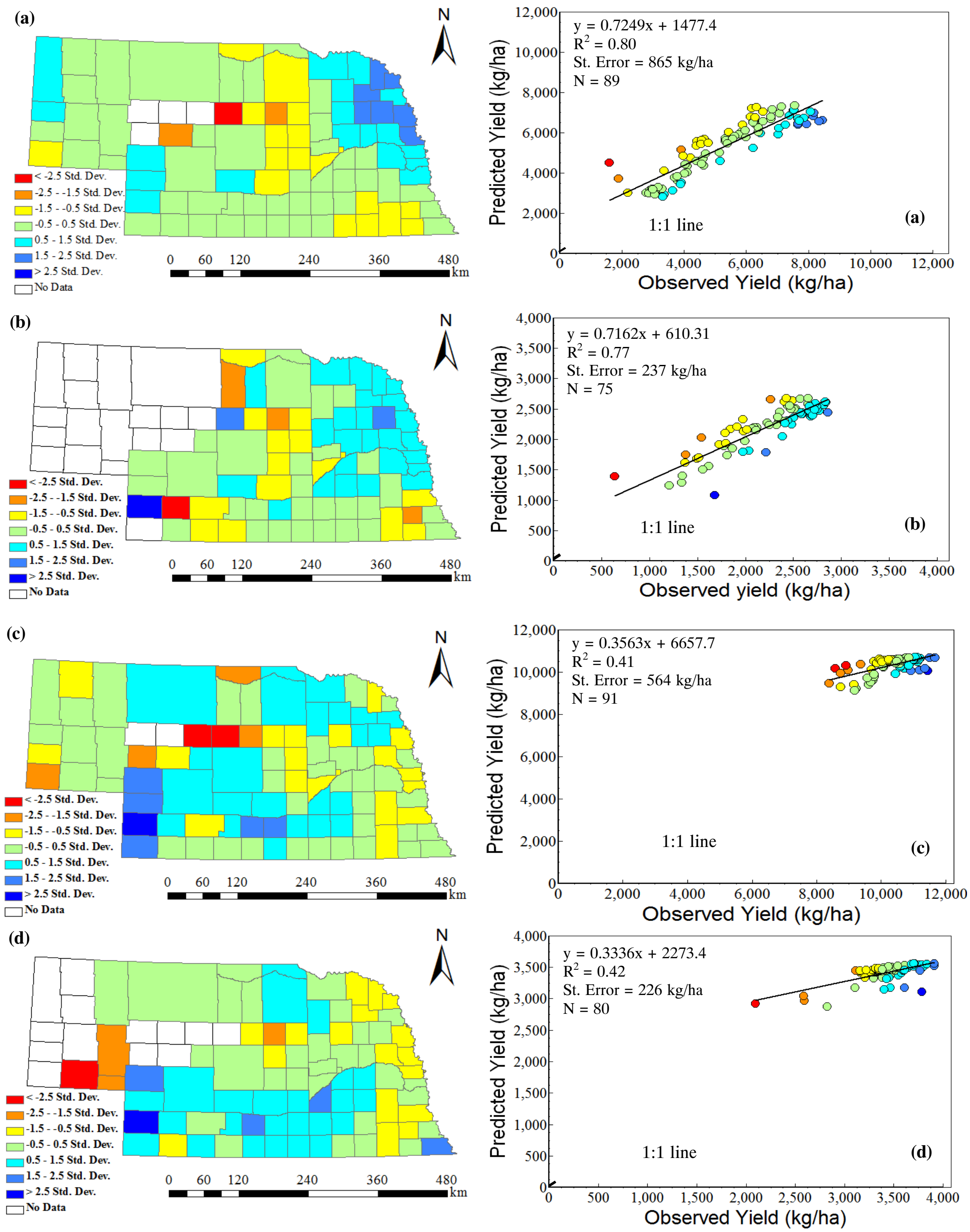

Figure 4. Predicted vs. (a) observed rainfed maize, (b) rainfed soybean, (c) irrigated maize, and (d) irrigated soybean yields across Nebraska using long-term average annual precipitation. Data points in each graph represent the yield of each county. The color of the data points on the graphs (right) is associated with the color of the counties (left). Residual maps on the left side of each graph show counties where yield is under- or overpredicted.

Long-Term Average Annual Precipitation vs. Predicted Yield Using GWR

The following two sections evaluate the relationship between the long-term mean annual, seasonal total (May-September), and monthly precipitation vs. predicted rainfed maize and soybean yields for each county using the GWR model. The relationship between predicted yield and long-term average mean annual total precipitation and the residuals of the regression between predicted and measured yield for irrigated and rainfed maize and soybean are presented in figures 4a through 4d. The results of the GWR model showed very good performance for rainfed maize and soybean. The residuals are spatially autocorrelated, as table 7 shows the Moran's I value for the residuals for all four crops. There is a general trend of strong linear correlation between predicted and observed yield with varying R^2 and SD values. The statistics between measured and predicted yield are presented in table 8. For rainfed maize and soybean, about 80% and 77%

Table 7. Value of spatial autocorrelation (Moran's I test) for spatial randomness of mean yield vs. mean annual precipitation GWR residuals.

	Rainfed Maize Yield Residual	Irrigated Maize Yield Residual	Rainfed Soybean Yield Residual	Irrigated Soybean Yield Residual
Moran's I (p-value)	0.65 (<0.001)	0.22 (<0.001)	0.35 (<0.001)	0.30 (<0.001)

Table 8. Statistics for the state-average observed vs. predicted yields using GWR for the period of 1996-2008. Yield was predicted using mean annual precipitation (*N* = number of counties, SE = standard error, and DF = degrees of freedom).

Crop	*N*	R^2	SE (kg ha^{-1})	DF	F-Ratio	p > F
Rainfed maize	89	0.80	793	87	183.03	<0.0001
Rainfed soybean	75	0.77	237	73	150.45	<0.0001
Irrigated maize	91	0.41	564	89	27.86	<0.0001
Irrigated soybean	80	0.42	266	78	15.44	<0.0002

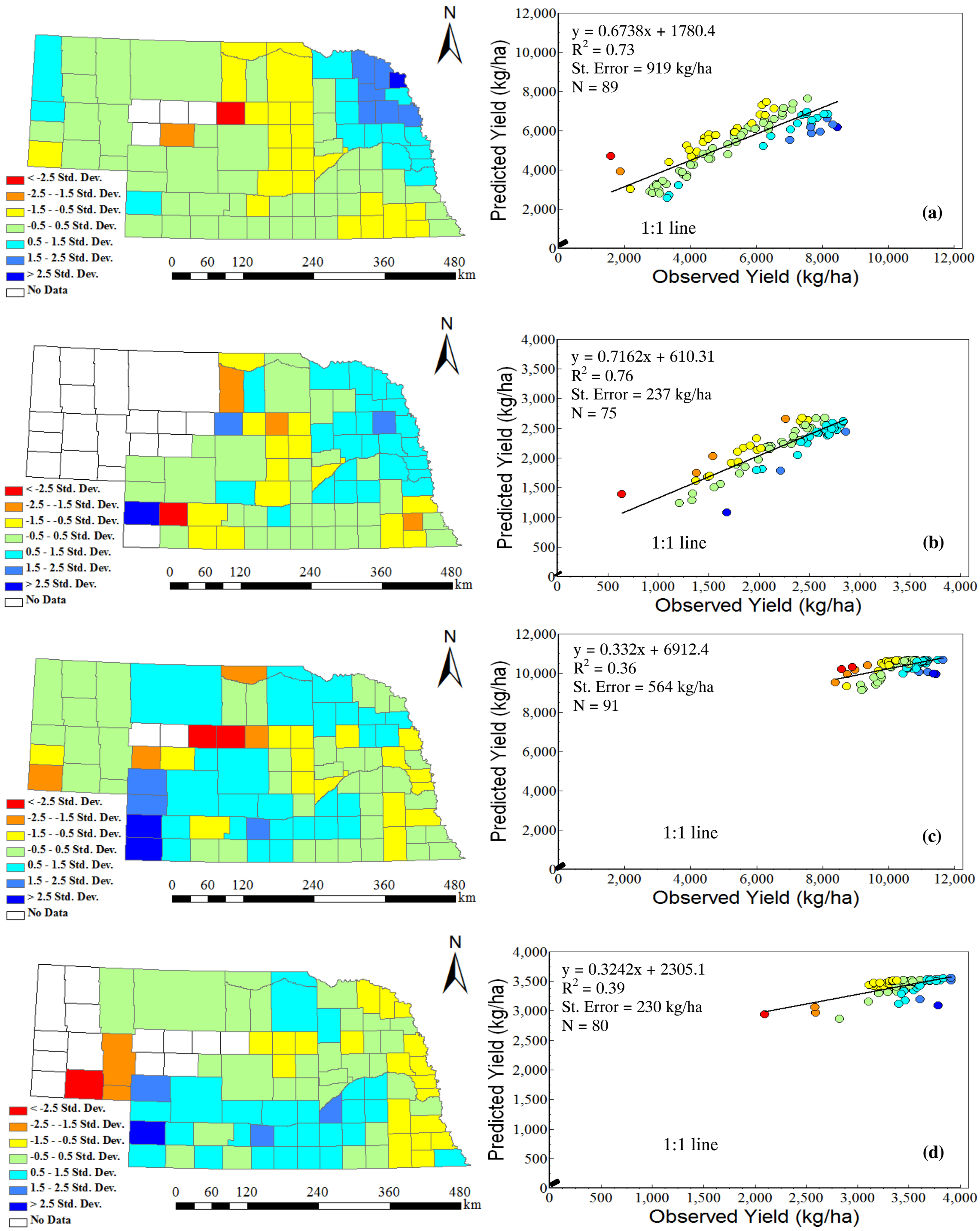

Figure 5. Predicted vs. (a) observed rainfed maize, (b) rainfed soybean, (c) irrigated maize, and (d) irrigated soybean yields across Nebraska using long-term average seasonal precipitation. Data points in each graph represent the yield of each county. The color of the data points on the graphs (right) is associated with the color of the counties (left). Residual maps on the left side of each graph show counties where yield is under- or overpredicted.

of the variation in yield was explained by the mean annual precipitation alone. For irrigated maize and soybean crops, about 41% to 42% of the variation in yield was explained by the mean annual precipitation (table 8).

The results of the standard deviation (SD) of the GWR predictions and the residual maps indicated that the residuals (observed-predicted yield) were within the 2.5 of the SD. For all crops, less than 2% of the counties fell outside the 1.5 SD range. GWR analysis with the observed and predicted yields using mean annual precipitation showed a strong relationship at 5% significance level (table 8). The t-test showed that the intercept and slope of the regression line were significantly different from unity ($p < 0.05$). Further analyses of the regression model showed that for most of the counties of the northeastern part of the state (including Dixon, Dakota, Wayne, Cuming, Thurston, and Burt), yields for rainfed crops were underpredicted (higher value of residual SD) (figs. 4a and 4b) even though precipitation is in the adequate range for crop production in those counties (fig. 2a). This is because precipitation is not the only limiting factor driving crop yield, and other factors such as evapotranspiration, air temperature, solar radiation, soil type and water-holding capacity, organic matter content, pH, irrigation management, crop characteristics, disease and pests pressure, soil management, and other management practices impact the yield, influencing the relationship between precipitation and irrigated and rainfed yields. Similarly for the irrigated crops, yield was underpredicted for the counties in the central part of the state. Since precipitation is less in these counties as compared with the counties in the eastern part, irrigation has more influence on yield predictions. In general, the regression model overpredicted the yield for high-yielding conditions and underpredicted for low-yielding conditions. On a statewide average, the R^2 values were greater for rainfed maize and soybean than for irrigated maize and soybean (i.e., R^2 = 0.80, 0.77, 0.41, and 0.42 for rainfed maize, rainfed soybean, irrigated maize, and irrigated soybean, respectively; table 8).

When compared to the OLS predictions, the GWR technique provided better predictions of yields for both irrigated and rainfed crops. When the R^2 values in table 6 and 8 are considered, the R^2 values for the GWR predictions were 13% and 16% higher for irrigated maize and soybean than the R^2 values for the OLS regression. The GWR predictions were further improved for rainfed maize and soybean, with 44% and 42% greater R^2 for GWR as compared to the OLS regression. The ability of GWR to account for spatial nonstationarity, especially for the rainfed conditions, provided enhanced yield predictions across the Nebraska counties.

Long-Term Average Seasonal Precipitation vs. Predicted Yield Using GWR

The measured and predicted maize and soybean yields for irrigated and rainfed conditions using seasonal total (May 1 to September 30) precipitation are presented in figures 5a through 5d. The values of spatial autocorrelation (Moran's I test) for the spatial randomness of the mean yield vs. mean seasonal total precipitation are presented in table 9, and the statistical analyses are presented in table 10. For rainfed crops, there was a strong correlation between seasonal precipitation and yield, with R^2 of 0.73 and 0.76 for maize and soybean, respectively. The residuals are spatially autocorrelated for all four crops (table 9).

Table 9. Values of the spatial autocorrelation (Moran's I test) for spatial randomness of mean yield vs. mean seasonal precipitation GWR residuals.

	Rainfed Maize Yield Residual	Irrigated Maize Yield Residual	Rainfed Soybean Yield Residual	Irrigated Soybean Yield Residual
Moran's I (p-value)	0.69 (<0.001)	0.25 (<0.001)	0.35 (<0.001)	0.42 (<0.001)

Table 10. Predicted yields for Nebraska counties for the observation period of 1996-2008 using mean seasonal total precipitation (*N* = number of counties, SE = standard error, and DF = degrees of freedom).

Crop	*N*	R^2	SE	DF	F-Ratio	p > F
Rainfed maize	89	0.73	919	87	147.68	<0.0001
Rainfed soybean	75	0.76	237	73	150.44	<0.0001
Irrigated maize	91	0.36	564	89	32.35	<0.0001
Irrigated soybean	80	0.39	230	78	19.98	<0.0001

As found with yield vs. annual total precipitation, the correlation between seasonal precipitation and yield for rainfed crops was stronger than for irrigated crops (table 9). For irrigated crops, 36% of the variability in the mean yield was explained by the mean seasonal precipitation for irrigated maize and 39% for irrigated soybean. The GWR SD residual maps indicate that the residuals were within the 2.5 SD range, which was higher than the SD with the annual precipitation. For all four crops, less than 2% of the counties fell outside the 1.5 SD range. The standard error of estimation for rainfed maize was higher when using seasonal precipitation than when using annual precipitation. The slopes of the regression line between precipitation and yield were significantly different from unity ($p < 0.05$).

Long-Term Average Monthly Precipitation vs. Predicted Yield Using GWR

Correlations between long-term average monthly total precipitation for May, June, July, August, and September vs. long-term county average yield for rainfed maize, rainfed soybean, irrigated maize, and irrigated soybean are presented in figures 6 through 9. The R^2 values between monthly precipitation and yields are summarized in figure 10. The residual maps for yield vs. precipitation for each month (May, June, July, August, and September) for each county for rainfed maize, rainfed soybean, irrigated maize, and irrigated soybean are presented in figures 11, 12, 13, and 14, respectively. The R^2 values ranged from 0.80 to 0.90 for rainfed crops and from 0.65 to 0.75 for irrigated crops. There were differences in the impact of individual monthly total precipitation on crop yield across the state (i.e., the R^2 between individual months' precipitation vs. yield varied from the eastern part to the western part; data not shown). This is because while the precipitation varied from east to west, differences in the impact of precipitation on yield are also due to differences in the planting date across the state. In the eastern part, maize hybrids that have a longer maturity date (114 to 120 days) are planted. The maturity date of the maize hybrids planted in the central region is usually 112 to 113 days, and the maturity date of the hybrids planted in the western portion of the state is shorter (90 to 95 days). Depending on the location in the state, planting date, climate conditions, and hybrid, the potential for kernel development for maize normally begins in June, tasseling usually begins in mid-July, silking/pollination

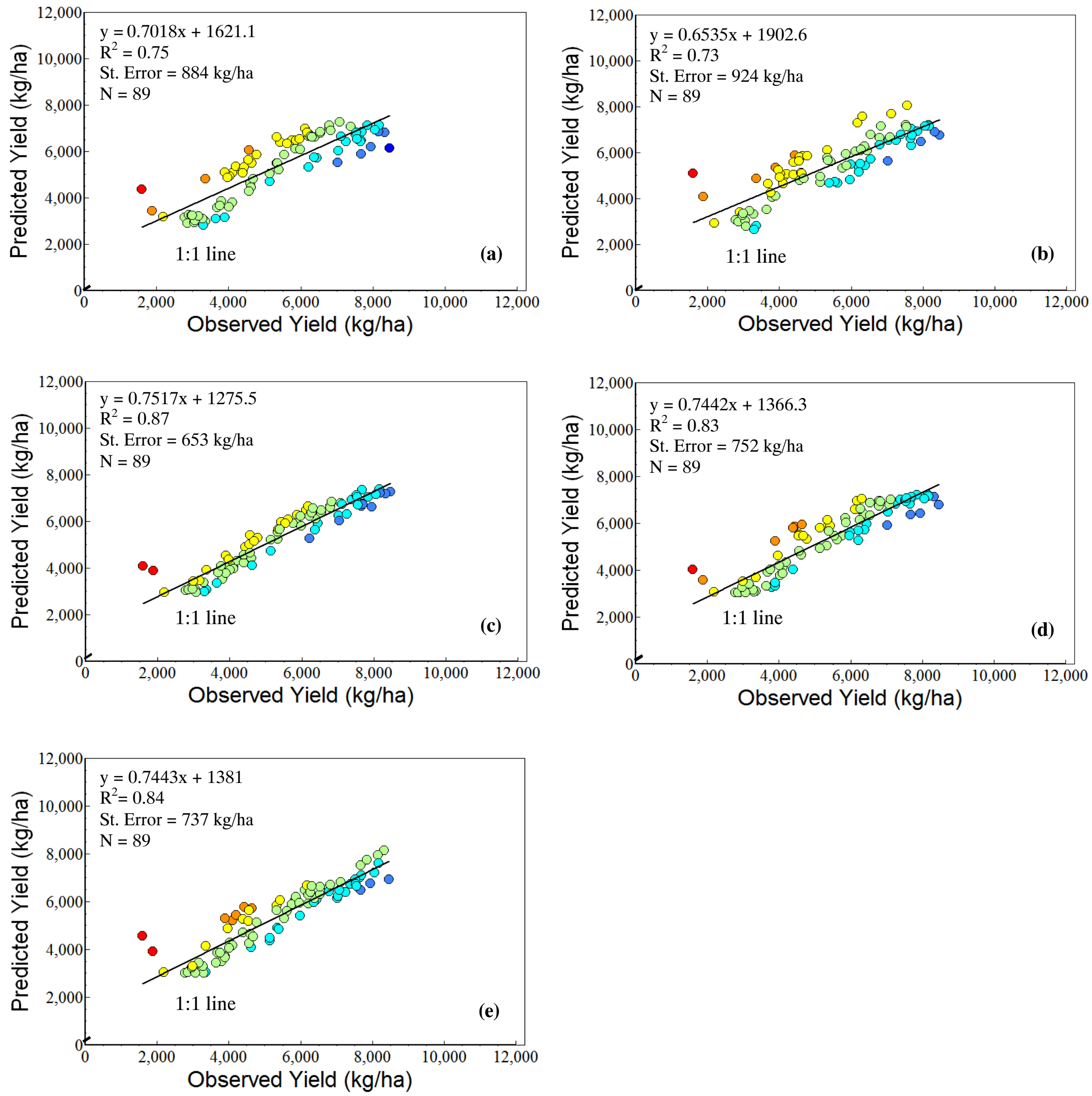

Figure 6. Predicted vs. observed rainfed maize yields across Nebraska using geographically weighted regression from mean monthly precipitation for (a) May, (b) June, (c) July, (d) August, and (e) September.

occurs in late July, and grain fill occurs during early to mid-August across the study region. The most critical growth stage for maize is usually between tasseling and silking. During this stage, plant water stress can delay silking relative to pollen shedding and can reduce seed set. Usually, the water stress during the vegetative growth period is not as critical as the tasseling-silking stage, and the stress during the grain fill can be intermediate in terms of its impact on yield (Musick and Dusek, 1980), although maintaining healthy plants in all growth stages is important for achieving high yields.

Another variable that can cause spatial non-stationarity of the relationship between the yield and precipitation is the significant variations in tillage managements that are practiced by maize and soybean farmers in Nebraska. Based on the survey conducted by the USDA Natural Resources Conservation Service (USDA-NRCS, 2009), the tillage practices not only change with locations but also show great variability for the same location and same crop between farmers. For example, based on the assessment by the USDA-NRCS, a larger percentage of maize was planted on no-till in the eastern part of Nebraska than in the central and western parts. Counties like Madison, Douglas, Johnson, Sarpy, Gage, and Jefferson in eastern Nebraska had $\geq$77% of the planted maize as no-till. Only one county (Banner) in western Nebraska had more than 70% of the maize land area planted as no-till. For soybean, eastern Nebraska had a higher percentage of land planted on no-till than the central part. Soybean is not grown extensively in western and north-central Nebraska (USDA-NRCS, 2009). Thus, there is a gradual decrease in both maize and soybean no-till planting land area from eastern to western Nebraska. Interestingly, for maize, the percentage of no-till maize planting followed an opposite trend with precipitation. Disk-till is another commonly used tillage practice and is usually concentrated in central and west-central Nebraska. These different tillage practices also impact the spatial non-stationarity relationships between precipitation and yield, as tillage practice influence the available soil water and precipitation relationship (i.e., depending on several factors, disk-

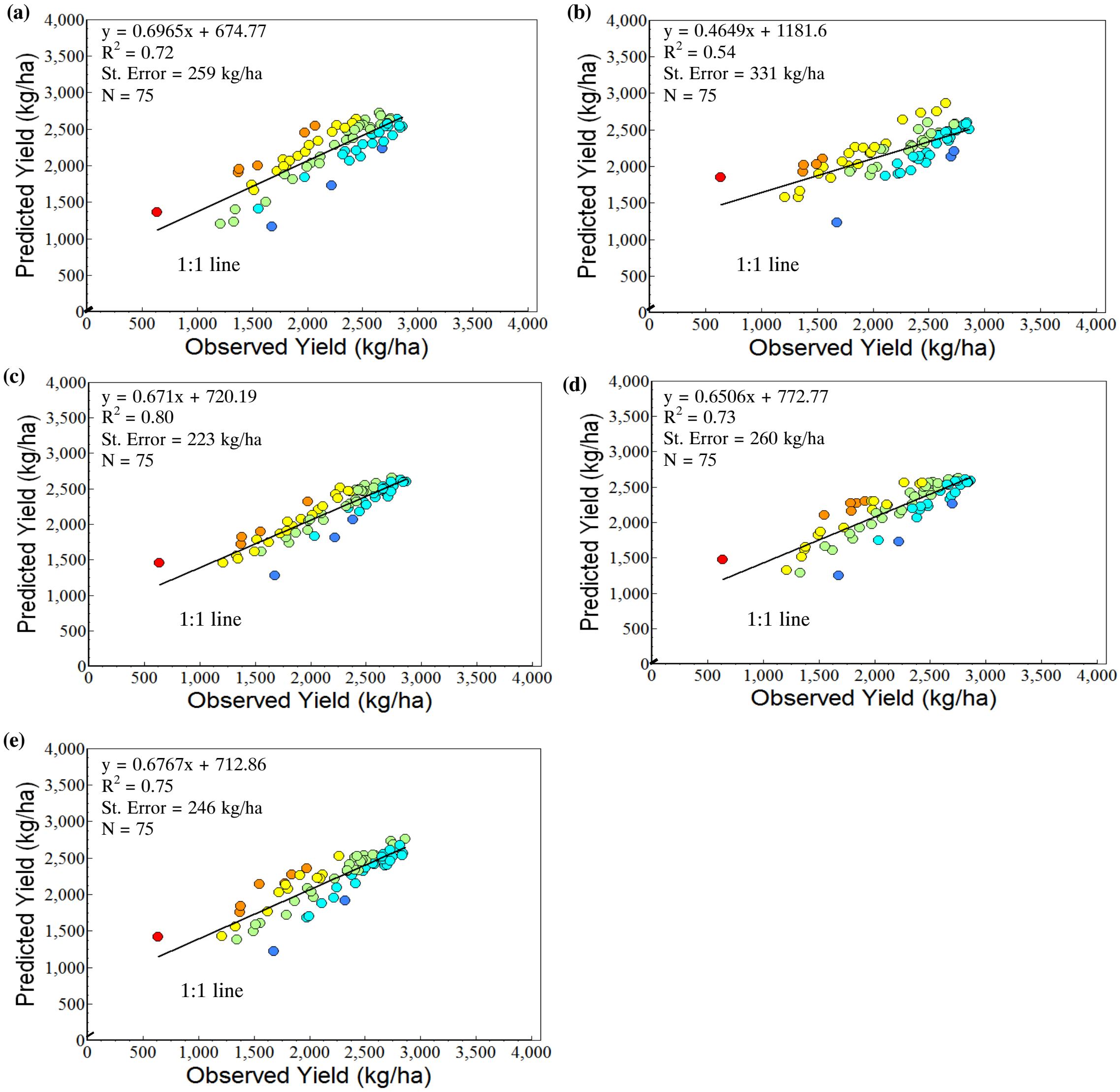

Figure 7. Predicted vs. observed rainfed soybean yields across Nebraska using geographically weighted regression from mean monthly precipitation for (a) May, (b) June, (c) July, (d) August, and (e) September.

Table 11. Predicted vs. observed yields for Nebraska counties for the period of 1996-2008 using mean monthly precipitation (May 1 to September 30). The intercept and slope were obtained from regression: *Y* = *a* + *bx*, where *Y* = predicted yield, and *x* = precipitation.

		Coefficient of Determination (R^2)					Standard Error (kg ha^{-1})				
Crop	*N*	May	June	July	Aug.	Sept.	May	June	July	Aug.	Sept.
Rainfed maize	89	0.75	0.73	0.87	0.83	0.84	884	924	653	752	737
Rainfed soybean	75	0.72	0.54	0.80	0.73	0.75	259	331	223	260	246
Irrigated maize	91	0.35	0.40	0.56	0.47	0.50	567	548	485	526	510
Irrigated soybean	80	0.36	0.50	0.67	0.52	0.52	235	210	176	213	212

till fields, in general, may have greater soil evaporation than no-till fields).

The July precipitation was the most critical for the high crop yield for both crops under irrigated and rainfed conditions. This is most likely related to the sensitivity of maize to water stress during the critical growth stages (tasseling and silking), which usually occurs in July, depending on the location in the state, as planting date varies from east to west. July precipitation had more impact on maize yield than on soybean under rainfed conditions. In all months and in both irrigated and rainfed treatments, yield was proportional to precipitation. For irrigated yields, July precipitation had more impact on soybean yield than on maize. The R^2 values with standard error values are presented in table 11. For rainfed maize, the initial precipitation in May is more important for setting potential for the number and size of kernels, whereas the precipitation in July is most important for yield potential. Counties with more July precipitation usually had higher yields. Our findings are in agreement with those of other researchers who also reported that less precipitation during the months of July, August, and September in the Great Plains region can reduce yield (Robin and Domingo, 1953; Denmead and Shaw, 1960; Musick and Dusek, 1980; Schlenker and Roberts, 2006). In a field study with maize in

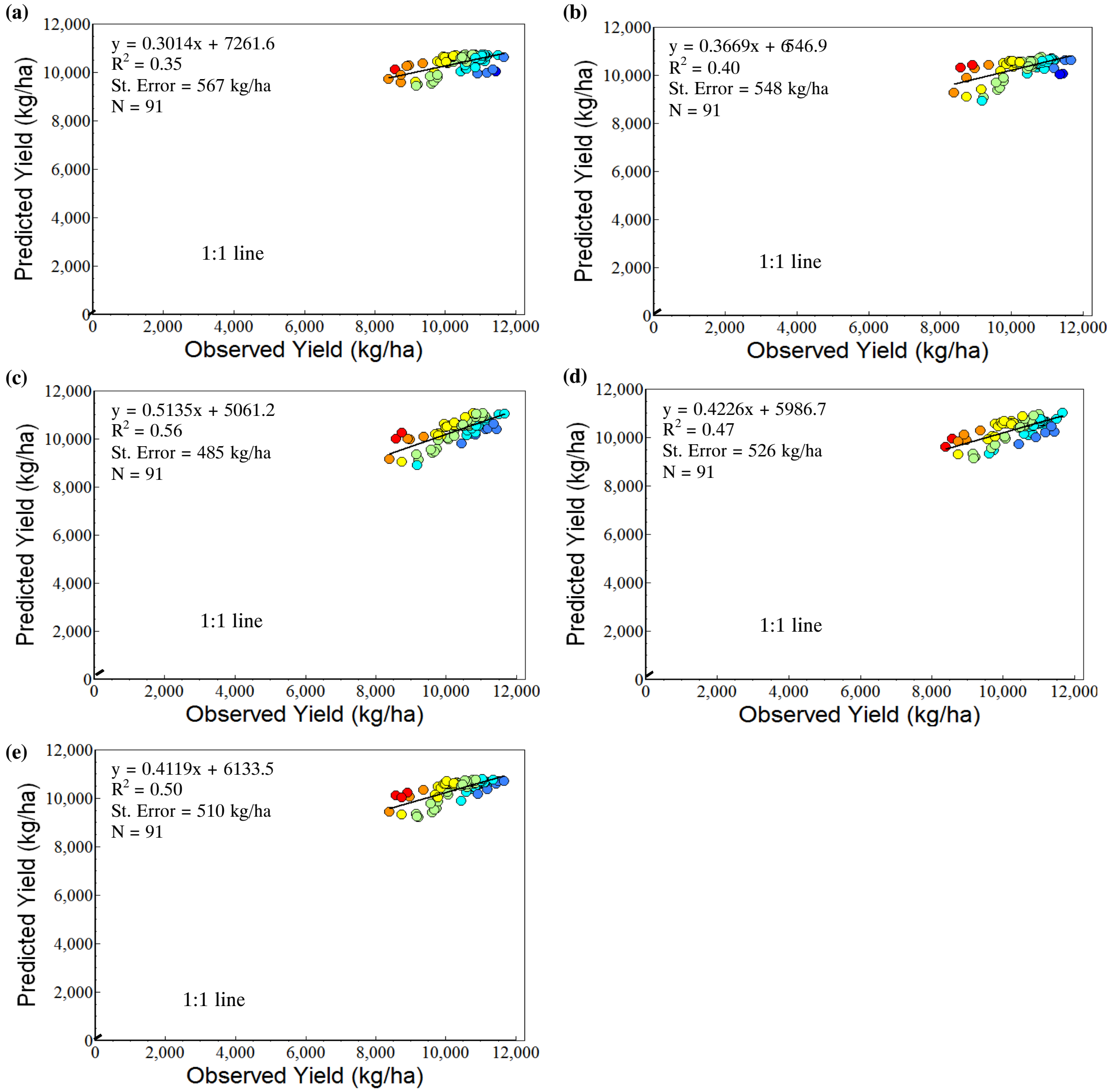

Figure 8. Predicted vs. observed irrigated maize yields across Nebraska using geographically weighted regression from mean monthly precipitation for (a) May, (b) June, (c) July, (d) August, and (e) September.

west-central Nebraska, Payero et al. (2009) found that irrigations applied in July had the highest positive correlation with yield. This high correlation decreased considerably for irrigations applied in August and became negative for irrigations applied in September. The best positive correlation between the soil water deficit factor and yield occurred during weeks 12 through 14 from crop emergence, during the “milk” and “dough” growth stages. Yield was poorly correlated to stress during weeks 15 and 16, and the correlation became negative after week 17. They reported that if water is limiting, then applying a larger proportion of the allocation in July is a good strategy, which supports our findings that for both irrigated and rainfed maize and soybean, July precipitation had the strongest correlation with yield (table 11, fig. 10).

The GWR SD residual maps indicate that the residuals were usually within the 1.5 SD range (figs. 11 through 14). The GWR analysis showed a significant relationship between observed and predicted yield (table 11). The t-test showed that the predicted values for intercept and slope of the regression line were significantly different from unity ($p < 0.05$). Similarly, for rainfed soybean, May precipitation is important for the initial potential pod development, and July precipitation is most important for maximum yield.

The R^2 values were lower for irrigated maize and soybean than for rainfed conditions and ranged from 0.30 to 0.70 and from 0.56 and 0.67 for irrigated maize and soybean, respectively, for July. Counties with higher July precipitation had higher yields. The GWR SD residual maps indicate that the residuals for irrigated crops were usually within the 2.5 SD range (figs. 11 through 14). When all the months are considered, none of the counties fell outside the 2.5 SD range for July. The relationship between observed and predicted yield was significant ($p < 0.05$) (table 12). In most cases, the GWR model performed well in predicting county-average yields as a function of precipitation across Nebraska, except for underpredicting for the northeast part of the state. In general, the model overpredicted yields for high-yielding conditions and underpredicted for low-yielding conditions.

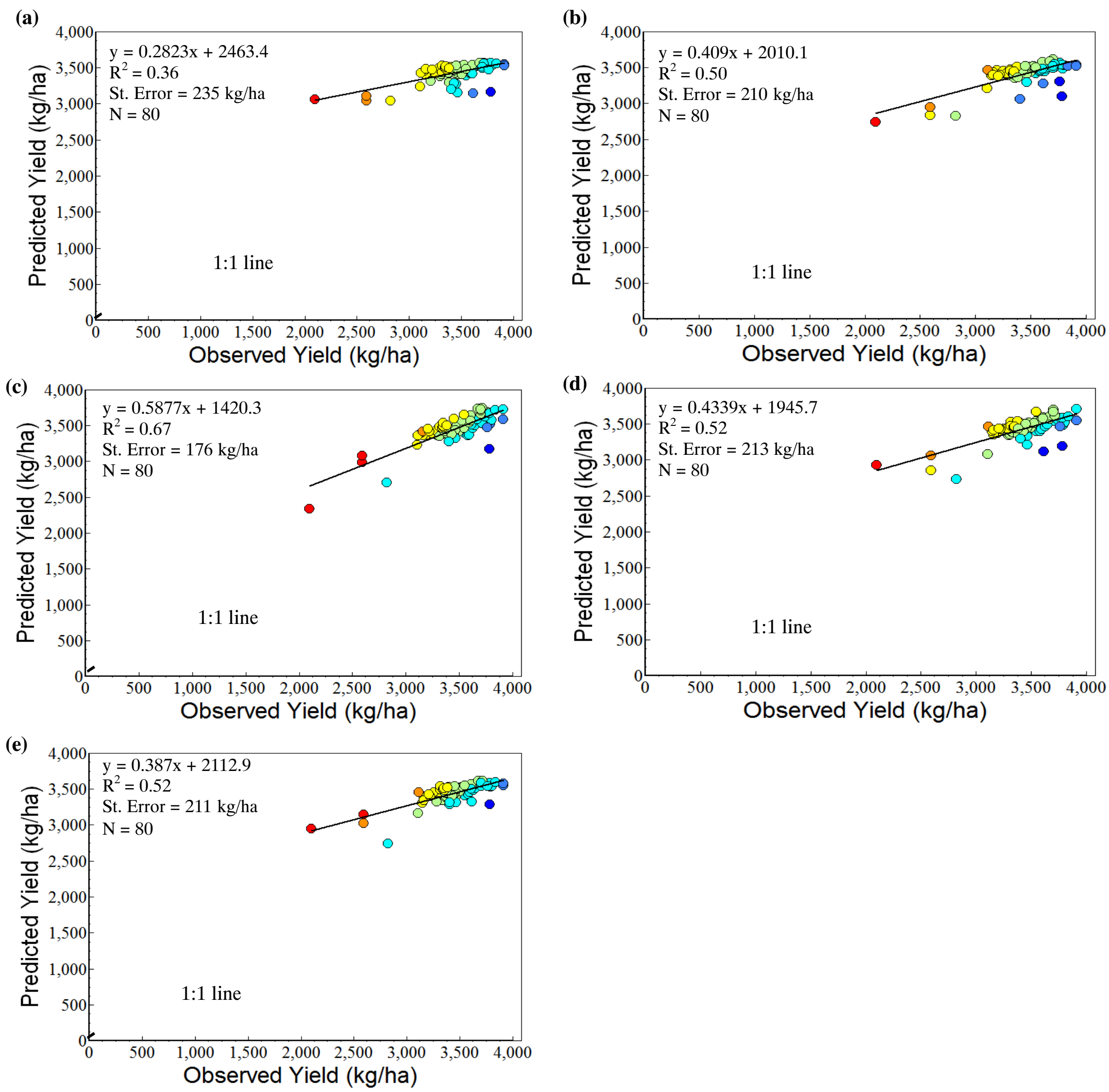

Figure 9. Predicted vs. observed irrigated soybean yields across Nebraska using geographically weighted regression from mean monthly precipitation for (a) May, (b) June, (c) July, (d) August, and (e) September.

Table 12. F-ratio statistics of mean yields for observation period of 1996-2008 using mean monthly precipitation (May 1 to September 30).

Crop	*N*	DF	F-Ratio					p > F
			May	June	July	August	September	
Rainfed maize	89	87	169.45	115.32	46.36	111.57	53.51	<0.0001
Rainfed soybean	75	73	145.55	41.59	18.34	51.8	30.17	<0.0001
Irrigated maize	91	89	26.75	21.19	39.87	35.27	9.8	<0.0001
Irrigated soybean	80	78	17.92	4.9	22.8	17.22	6.88	<0.0001

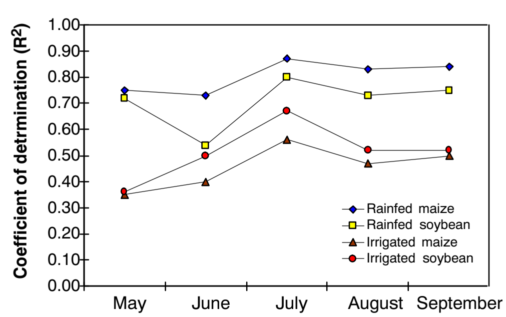

Figure 10. Coefficient of determination (R^2) between observed and predicted irrigated and rainfed maize and soybean yields using geographically weighted regression from mean monthly precipitation (May-September).

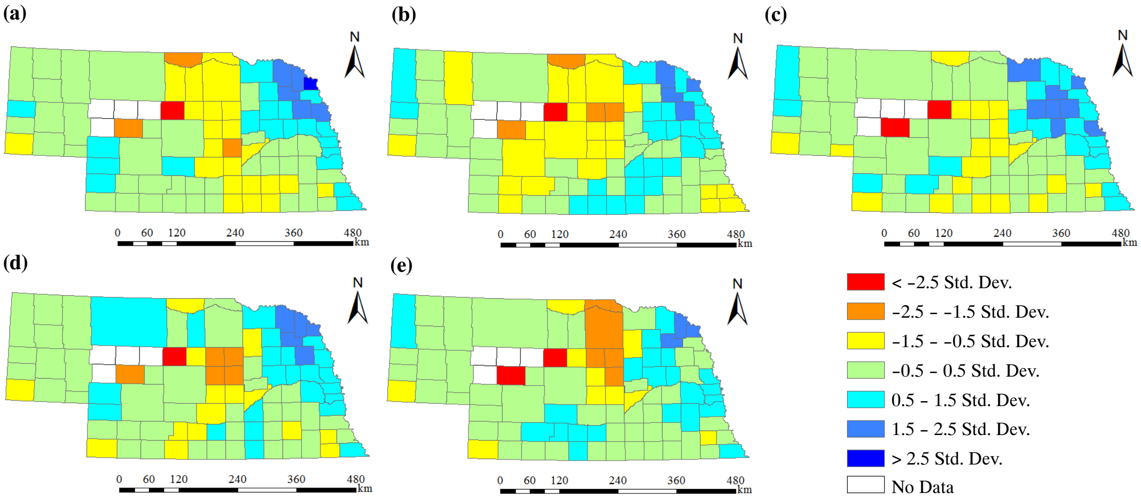

Figure 11. Residual maps for (a) May, (b) June, (c) July, (d) August, and (e) September for rainfed maize showing counties where yield is under- or overpredicted based on the data presented in figure 6.

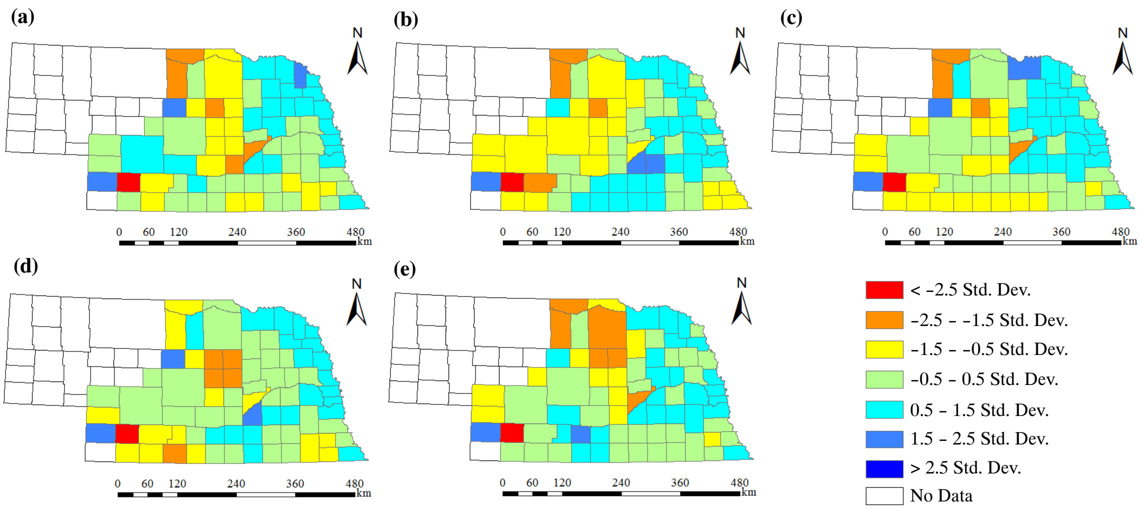

Figure 12. Residual maps for (a) May, (b) June, (c) July, (d) August, and (e) September for rainfed soybean showing counties where yield is under- or overpredicted based on the data presented in figure 7.

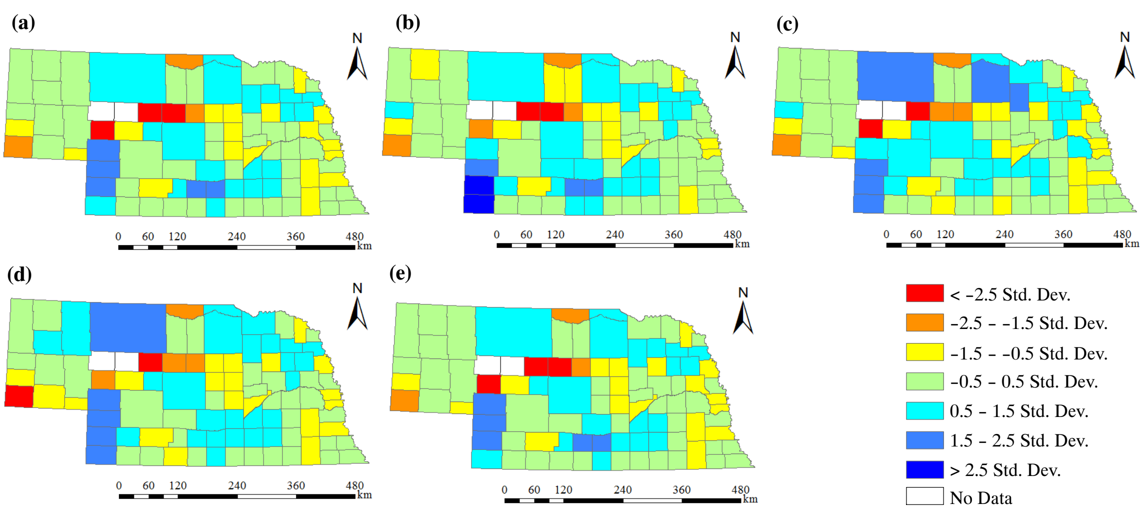

Figure 13. Residual maps for (a) May, (b) June, (c) July, (d) August, and (e) September for irrigated maize showing counties where yield is under- or overpredicted based on the data presented in figure 8.

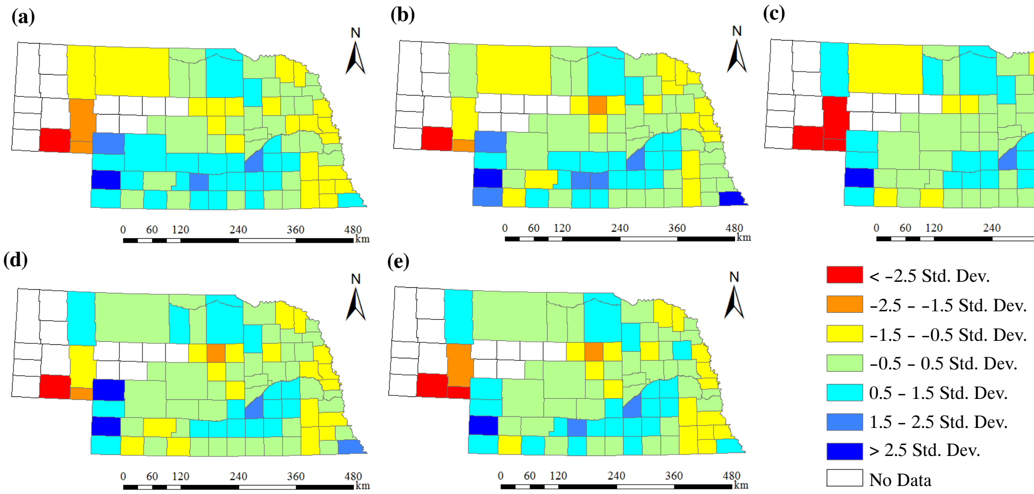

Figure 14. Residual maps for (a) May, (b) June, (c) July, (d) August, and (e) September for irrigated soybean showing counties where yield is under- or overpredicted based on the data presented in figure 9.

Conclusions

Variations in irrigated and rainfed maize and soybean yield at the county level were analyzed for the state of Nebraska for 1996 to 2008. Geographically weighted regression (GWR) and ordinary least square (OLS) models were used to develop relationships between yield and precipitation. The GWR techniques accounts for the spatial non-stationarity of the relationships between precipitation and crop yields. The yields on a county level and statewide average basis were correlated to both stationary (point measurements and spatially interpolated) annual (January to December), seasonal (May-September), and monthly (May, June, July, August, and September) precipitation for all 93 counties. The spline method was used to interpolate the spatial distribution of precipitation across the state. Precipitation was regressed with irrigated and rainfed maize and soybean to describe the maize and soybean yield as a function of precipitation. When GWR model-predicted yields were considered, the performance of GWR in estimating yield for both maize and soybean under irrigated and rainfed conditions was significantly better than the performance of OLS. When the OLS regression model was used, there was a general trend of linear correlation between observed yield and long-term average mean annual total precipitation, with a varying coefficient of determination (R^2). For rainfed crops, 67% of the variability in mean yield was explained by the mean annual precipitation. About 23% and 17% of the variability in mean yield was explained by mean annual precipitation for irrigated maize and soybean, respectively.

The performance of the GWR technique in predicting yields was significantly better than the performance of OLS. When the GWR technique was used, for both rainfed maize and soybean, 77% to 80% of variation in yield was explained by the mean annual precipitation alone. For irrigated crops, 42% of the variation in the yield was explained by the mean annual precipitation. For rainfed crops, there was a strong correlation between seasonal precipitation and yield, with R^2 of 0.73 and 0.76 for maize and soybean, respectively. The mean annual total precipitation was a better predictor of rainfed maize yield than rainfed soybean yield. Correlation between precipitation and yield were lower for irrigated maize and soybean, and irrigated yields had lower standard error of prediction than rainfed yields. On a statewide average, July precipitation had the highest correlation with yield for both maize and soybean. Our study showed the superiority of the GWR technique over the conventional OLS regression approach in analyzing the relationship between yield and precipitation and predicting yields for irrigated and rainfed maize and soybean. The analyses of this study can be useful to predict maize and soybean yields as a function of spatially interpolated monthly precipitation ahead of the harvest season. The superiority of the GWR technique is mainly due to accounting for the impact of the spatial non-stationarity of the precipitation vs. yield relationships. The stationary attributes of the OLS model present a challenge in predicting yields on large scales. In addition to the precipitation, further application and evaluations of the relatively new GWR technique in similar agricultural research topics can improve the yield predictions by accounting for the spatial non-stationarity of other climatic variables (i.e., air temperature, solar radiation, etc.) and management practices.

References

Bakhash, A., and R. S. Kanwar. 2004. Using discriminant analysis and GIS to delineate subsurface drainage patterns. *Trans. ASAE* 47(3): 689-699.

Berzsenyi, Z., B. Gyorffy, and D. Q. Lap. 2000. Effect of crop rotation and fertilization on maize and wheat yields and yield stability in long-term experiment. *European J. Agron*. 13(2-3): 225-244.

Borrough. P. A., and R. A. McDonnell. 1998. *Principles of Geographical Information System: Spatial Information Systems and Geostatistics.* New York, N.Y.: Oxford University Press.

Brunsdon, C., J. McClatchey, and D. Unwin. 2001. Spatial variations in the average rainfall-altitude relationships in Great Britain: An approach using geographically weighted regression. *Intl. J. Climatol*. 21(4): 455-466.

CALMIT. 2005. Center for Advanced Land Management Information Technologies. Lincoln, Neb.: University of Nebraska-Lincoln, School of Natural Resources. Available at: www.calmit.unl.edu.

Calvo, C., and M. Escolar. 2003. The local voter: A geographically weighted approach to ecological inference. *American J. Political Sci*. 47(1): 189-204.
Carbone, G. J., S. Narumalani, and M. King. 1996. Application of remote sensing and GIS technologies with physiological crop models. *Photogram. Eng. and Remote Sensing* 62(2): 171-179.
Curran, P. J., and P. M. Atkinson. 1998. Geostatistics and remote sensing. *Prog. Physical Geography* 22(1): 61-78.
Curran, P. J., G. M. Foody, and P. R. van Gardingen. 1997. Scaling-up. In *Scaling-up from Cell to Landscape*, 1-5. P. R. van Gardingen, G. M. Foody, and P. J. Curran, eds. Cambridge, U.K.: Cambridge University Press.
Curry, R. B., R. M. Peart, J. W. Jones, K. J. Boote, and L. H. Allen, Jr. 1990. Simulation as a tool for analyzing crop response to climate change. *Trans. ASAE* 33(3): 981-990.
Denmead, O. T., and R. H. Shaw. 1960. The effects of soil moisture stress at different stages of growth on the development and yield of corn. *Agron. J*. 52(5): 272-274.
Dent, J. B., and P. K. Thornton. 1988. The role of biological simulation models in farming systems research. *Agric. Admin. and Extension* 29(2): 111-122.
Flowers, M., R. Weisz, and J. G. White. 2005. Yield-based management zones and grid sampling strategies: Describing soil test and nutrient variability. *Agron. J*. 97(3): 968-982.
Foody, G. M. 2003. Geographical weighting as a further refinement to regression modeling: An example focused on the NDVI-rainfall relationships. *Remote Sensing Environ.* 88(3): 283-293.
Fotheringham, A. S., C. Brundson, and M. E. Charlton. 1998. Geographically weighted regression: A natural evolution of the expansion method for spatial data analysis. *Environ. and Planning* 30(11): 1905-1927.
Fotheringham, A. S., C. Brunsdon, and M. E. Charlton. 2002. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships.* West Sussex, U.K.: John Wiley and Sons.
Fortin. M. J., and M. Dale. 2005. *Spatial Analysis: A Guide for Ecologists.* Cambridge, U.K.: Cambridge University Press.
Franke, R. 1982. Smooth interpolation of scattered data by local thin plate splines. *Comp. and Math. with Appl.* 8(4): 273-281.
Gessler, P. E., O. A. Chadwick, F. Chamran, L. Althouse, and K. Holmes. 2000. Modeling soil-landscape and ecosystem properties using terrain attributes. *SSSA J*. 64(6): 2046-2056.
Hansen, J. W., and J. W. Jones. 2000. Scaling-up crop models for climate variability applications. *Agric. Systems* 65(1): 43-72.
Hoogenboom, G., and P. K. Thornton. 1990. A GIS for agrotechnology transfer application in Guatemala. In *Application of GIS, Simulation Models, and Knowledge-Based System for Landuse Management*, 61-70. Blacksburg, Va.: Virginia Tech, Department of Agricultural Engineering.
Irmak, A., J. W. Jones, W. D. Batchelor, and J. O. Paz. 2001. Estimating spatially variable soil properties for crop model use in precision farming. *Trans. ASAE* 42(6): 1867-1877.
Irmak, A., J. W. Jones, W. D. Batchelor, and J. O. Paz. 2002. Linking multi-variables for diagnosing causes of spatial yield variability in soybean. *Trans. ASAE* 45(3): 839-849.
Irmak, A., J. W. Jones, and S. S. Jagtap. 2005. Evaluation of the CROPGRO-Soybean model for assessing climate impacts on regional soybean yields. *Trans. ASAE* 48(6): 2343-2353.
Irmak, A., J. W. Jones, W. D. Batchelor, S. Irmak, J. O. Paz and K. J. Boote. 2006. Analysis of spatial yield variability using a combined crop model-empirical approach. *Trans. ASABE* 49(3): 811-818.
Irmak, A., P. Ranade, D. Marx, S. Irmak, K. G. Hubbard, G. Meyer, and D. L. Martin. 2010. Spatial interpolation of climate variables in Nebraska. Trans. ASABE *53(6): 1759-1771*.
Irmak, S. 2010. Nebraska Water and Energy Flux Measurement, Modeling, and Research Network (NEBFLUX). *Trans. ASABE* 53(4): 1097-1115.
Irmak, S., E. Istanbulluoglu, and A. Irmak. 2008. An evaluation of evapotranspiration model complexity versus performance in comparison with Bowen ratio energy balance measurements. *Trans. ASABE* 51(4): 1295-1310.
Jones, C. A., and J. R. Kiniry, eds. 1986. *CERES-Maize: A Simulation Model of Maize Growth and Development.* College Station, Tex.: Texas A&M University Press.
Jones, J. W., G. Hoogenboom, C. H. Porter, K. J. Boote, W. D. Batchelor, L. A. Hunt, P. W. Wilkens, U. Singh, A. J. Gijsman, and J. T. Ritchie. 2003. The DSSAT cropping system model. *European J. Agron.* 18(3-4): 235-265.
Kaspar, T. C., T. S. Colvin, D. B. Jaynes, D. L. Karlen, D. E. James, and D. W. Meek. 2003. Relationship between six years of corn yields and terrain attributes. *Precision Agric*. 4(1): 87-101.
Kravchenko, A. N., and G. P. Robertson. 2007. Can topographical and yield data substantially improve total soil carbon mapping by regression kriging? *Agron. J*. 99(1): 12-17.
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(2): 514-523.
Kulkarni, S. S., S. G. Bajwa, J. C. Rupe, and T. L. Kirkpatrick. 2008. Spatial correlation of crop response to soybean cyst nematode (*Heterodera glycines*). *Trans. ASABE* 51(4): 1451-1459.
Lal, H., G. Hoogenboom, J.-P. Calixte, J. W. Jones, and F. H. Beinroth. 1993. Using crop simulation models and GIS for regional productivity analyses. *Trans. ASAE* 36(1): 175-184.
Lauzon, J. D., I. P. O'Halloran, D. J. Fallow, A. P. von Bertoldi, and D. Aspinall. 2005. Spatial variability of soil test phosphorous, potassium, and pH of Ontario soils. *Agron. J*. 97(2): 524-532.
Li, B., R. W. McClendon, and G. Hoogenboom. 2004. Spatial interpolation of weather variables for single locations using artificial neural network. *Trans. ASAE* 47(2): 629-637.
Longley, P. A., and C. Tobon. 2004. Spatial dependence and heterogeneity in patterns of hardship: An intra-urban analysis. *Annals Assoc. American Geographers* 94(3): 503-519.
McBratney, A. B., and M. J. Pringle. 1997. Spatial variability in soil- implications for precision agriculture. In *Proc. 1st European Conf. on Precision Agric.*, vol. 1: 3-31. J. V. Stafford, ed. Oxford, U.K.: BIOS Scientific Publishers.
McKinion, J. M., J. L. Willers, and J. N. Jenkins. 2010. Spatial analyses to evaluate multi-crop yield stability for a field. *Computers and Electronics in Agric*. 70(1): 187-198.
Mead, R., J. Riley, K. Dear, and S. P. Singh. 1986. Stability comparison of intercropping and monocropping systems. *Biometrics* 42(2): 253-266.
Mennis, J. 2006. Mapping the results of geographically weighted regression. *Cartographic J. British Cartographic Soc*. 43(2): 171-179.
Mennis, J., and L. Jordan. 2005. The distribution of environmental equity: Exploring spatial nonstationarity in multivariate models of air toxic releases. *Annals Assoc. American Geographers* 95(2): 249-268.
Moen, T. N., H. M. Kaiser, and S. J. Riha. 1994. Regional yield estimation using a crop simulation model: Concepts, methods, and validation. *Agric. Systems*. 46(1): 79-92.
Mueller, T. G., and F. J. Pierce. 2003. Soil carbon maps: Enhancing spatial estimates of soil carbon with terrain attributes. *SSSA J*. 67(1): 258-267.
Musick, J. T., and D. A. Dusek. 1980. Irrigated corn yield response to water. *Trans. ASAE* 23(1): 92-98.
Oliver, M. A. 1999. Exploring soil spatial variability geostatistically. In *Proc. 2nd European Conf. on Precision Agric.*, part 1: 3-17. J. V. Stafford, ed. Sheffield, U.K.: Sheffield Academic Press.
Oliver, M. A., and R. Webster. 1991. How geostatistics can help you. *Soil Use and Mgmt.* 7(4): 206-217.

Payero, J. O., D. D. Tarkalson, S. Irmak, D. Davison, and J. Petersen. 2009. Effect of timing of a deficient-irrigation allocation on corn evapotranspiration, yield, water use efficiency, and dry mass. *Agric. Water Mgmt.* 96(10): 1387-1397.

Propastin, P., M. Kappas, and S. Erasmi. 2007. Application of geographically weighted regression to investigate the impact of scale on prediction uncertainty by modeling relationships between vegetation and climate. *Intl. J. Spatial Data Infrastructures Res*. 3: 73-94.

Robin, J. S., and C. E. Domingo. 1953. Some effect of severe soil moisture deficits at specific growth stages in corn. *Agron. J*. 45(12): 618-612.

Rosenthal, W. D., G. L. Hammer, and D. Butler. 1998. Predicting regional grain sorghum production in Australia using spatial data and crop simulation modeling. *Agric. Forest Meteorol*. 91(3-4): 263-274.

Sarangi, A., C. A. Cox, and C. A. Madramootoo. 2005. Geostatistical methods for prediction of spatial variability of rainfall in a mountainous region. *Trans. ASAE* 48(3): 943-954.

Schlenker, W., and M. J. Roberts. 2006. Nonlinear effects of weather on corn yields. *Rev. Agric. Econ*. 28(3): 391-398.

Searcy, N. D., and A. R. Longwell. 1964. *Nebraska Atlas.* Kearney: Nebraska Atlas Publishing.

Thornton, P. K. 1991. Application of crop simulation models in agricultural research and development in the tropics and subtropics. Paper series P-15. Muscle Shoals, Ala.: International Fertilizer Development Center.

Thornton, P. K., and C. A. Baanante. 1991. Use of simulation models for evaluating sustainable production systems. In *Proc. Symp. on Sustainability of Pasture Crop Rotation in the Southern Cone of South America*, 22. Montevideo, Uruguay: INIA Publications.

Thornton, P. K., A. R. Saka, U. Singh, J. D. T. Kumwenda, J. E. Brink, and J. B. Dent. 1995. Application of a maize crop simulation model in the central region of Malawi. *Experimental Agric*. 31(2): 213-226.

Tiwari. K. N., and S. K Sharma. 2009. Development of user-friendly software for prediction of monthly runoff and sediment yield of upper Damodar Valley catchment, India. ASABE Paper No. 096497. St. Joseph, Mich.: ASABE.

USDA-NASS. 2007. Census of agriculture: Farm and ranch irrigation survey. Washington, D.C.: USDA National Agricultural Statistics Service.

USDA-NASS. 2010. Yield data: Nebraska. Washington, D.C.: USDA National Agricultural Statistics Service. Available at: http://quickstats.nass.usda.gov.

USDA-NRCS. 2009. Nebraska no-till. Washington, D.C.: USDA Natural Resources Conservation Service. Available at: www.ne.nrcs.usda.gov/technical/notill.html.

Webster, R. 1985. Quantitative spatial analysis of soil in the field. In *Advances in Soil Sci*. 3: 1-70. New York, N.Y.: Springer.

Yang, H. S., A. Dobermann, J. L. Lindquist, D. T. Walters, T. J. Arkebauer, and K. G. Cassman. 2004. Hybrid-Maize: A maize simulation model that combines two modeling approaches. *Field Crops Res.* 87(2-3): 131-154.