Agricultural and Forest Meteorology 168 (2013) 26–35

Contents lists available at SciVerse ScienceDirect

Agricultural and Forest Meteorology

journal homepage: www.elsevier.com/locate/agrformet

Pre-harvest forecasting of county wheat yield and wheat quality using weather information☆

Byoung-Hoon Lee[a,b], Phil Kenkel[b], B. Wade Brorsen[b,*]

[a] Korea Rural Economic Institute, 117-3 Hoegi-ro, Dongdeamun-gu, Seoul 130-710, Republic of Korea
[b] Department of Agricultural Economics, Oklahoma State University, Stillwater, OK 74078, United States

ARTICLE INFO

Article history:
Received 22 November 2011
Received in revised form 9 July 2012
Accepted 16 August 2012

Keywords:
Forecasting
Prediction
Protein
Test weight
Wheat
Yield

ABSTRACT

County wheat yield and wheat quality are forecast using weather information. Regression models are estimated to account for the effect of weather on county wheat yield, protein, and test weight. The explanatory variables include precipitation and temperature for growing periods that correspond to biological wheat development stages. Wheat yield, protein, and test weight are strongly influenced by weather. The forecasting power of the yield and protein models is enhanced by adding a spatial lag effect. Out of sample forecasting tests confirm the models' usefulness in predicting wheat yield and wheat quality.

1. Introduction

Winter wheat in the Southern Plains is a mostly dry land crop with substantial year-to-year variations in yields and quality due to rainfall, temperature, and other weather events. If wheat yield and wheat quality response to weather conditions could be predicted early and accurately, the information could be widely used. The information could be particularly important to farmers optimizing late season agronomic decisions such as fungicide treatments and marketing decisions such as hedging or forward contracting. Smith and Gooding (1999) and Woolfolk et al. (2002) argue that predicting grain quality before wheat harvest would be important information to grain buyers, especially those entering forward contracts. Predictions of wheat yield and quality would also be useful to grain elevators and millers as they make decisions on repositioning inventories to make space for harvest deliveries and to plan segregation strategies. Thus, there has been increasing interest in the use and development of robust crop weather response models. The research reported here was undertaken at the specific request of Oklahoma wheat industry leaders and the resulting forecasts are already being communicated through an extension program.

Two main approaches to predicting crop yield based on weather conditions are simulation models and multiple regression models. A number of comprehensive agricultural simulation models have been used to predict yield and yield variability of wheat. Jones and Kiniry (1986) suggest a model to simulate the effects of genotype and weather conditions on crop yield, Duchon (1986), Claborn (1998), Bannayan et al. (2003), and Tsvetsinskaya et al. (2003) predict yields using weather forecasts and scenarios using the Crop Environment Resource Synthesis (CERES) simulation model. For the Great Plains, Eastering et al. (1998) and Wang et al. (2006) use the Erosion Productivity Impact Calculator (EPIC). Eastering et al. (1998) find that spatial disaggregation of climate data enhances predictions.

The simulation models designed to forecast crop yield use details about crop biology. However, as noted by Walker (1989), simulation requires extensive information such as soil type, plant parameters, and weather data related to the crop development stage, which are often not readily available. Tannura et al. (2008) argue that an important limitation of crop simulation models is that they are likely to ignore the influence of technology development over time. Bechter and Rutner (1978) and Just and Rausser (1981) find single-equation models forecast more accurately than large econometric models, and we expect a similar result for agronomic models in this study.

Because of the complexity and data requirements of simulation models, many previous studies have preferred a regression approach to a large simulation model for forecasting crop yield and quality. Studies using the multiple regression approach include

☆ Partial funding was provided by the Oklahoma Agricultural Experiment Station through Hatch project OKL02721.

* Corresponding author. Tel.: +1 405 744 6836; fax: +1 405 744 8210.
E-mail address: wade.brorsen@okstate.edu (B.W. Brorsen).

0168-1923/$ – see front matter
http://dx.doi.org/10.1016/j.agrformet.2012.08.010

Yang et al. (1992), Dixon et al. (1994), Kandiannan et al. (2002), and Chen and Chang (2005), who use various production functions to capture the effect of climate variables on crop yields. Irwin et al. (2008) and Tannura et al. (2008) modify Thompson's (1969, 1970) corn and soybean regression model and find crop yield strongly relates to weather conditions such as temperature, rainfall, technology, and other weather variables. As Tannura et al. (2008) and other studies have proven, multiple regression models have high explanatory power and can represent relationships between weather conditions and crop yield. Thus, the multiple regression model approach is not only easier to use, it is also likely to be more accurate than the simulation model approach.

Several studies have investigated the influences of weather conditions on wheat quality. The crop maturation period, includes milk development, heading, and ripening stages, which are the critical stages in determining wheat quality (FAO, 2002). Graybosch et al. (1995), Johansson and Svensson (1998), Smith and Gooding (1999), Guttieri et al. (2000), and Johansson et al. (2008) developed quality models that show that the effects of weather and environment strongly influence protein content and wheat test weight. Regnier et al. (2007) investigate the variations in flour and dough functionality traits associated with environmental factors and find the interaction between crop years and production regions is a significant factor for flour and dough qualities since growing conditions and climate conditions differ among the regions and across years.

Previous studies on the effect of weather on wheat quality have several limitations. Most of the previous quality models are not designed to produce preharvest predictions. The studies which developed prediction models do not consider out of sample forecasts but measure in-sample fit. In-sample fit can be inaccurate because most models, including ours, are developed from pretesting over a large number of alternative specifications.

Another major limitation of previous quality prediction studies is studies have either used data from a single location or have not used the extra information provided by spatial data. The increasing availability of spatial climate information makes it important to incorporate this new level of information to improve forecasts. Incorporating spatial data also requires changes in the model estimation procedures such as using a spatial lag model. Oklahoma has two unique resources for examining the relationship between weather and wheat yields and quality. The Oklahoma Mesonet consists of 120 automated stations that cover Oklahoma with one or more stations in each of Oklahoma's 77 counties. Plains Grains, Inc. (PGI) is a private, nonprofit wheat marketing organization based in Stillwater, Oklahoma. PGI evaluates wheat quality, including milling and baking quality, from an extensive network of samples at the county level. These two unique data sets provide the opportunity to examine the ability to predict wheat yield and quality with weather data. These two data sets (mesoscale weather data and elevator scale quality data) are highly disaggregated. Thus, the disaggregated data sets could provide more precise wheat yield and quality predictions than has been possible with the data sets used in past research.

The objective of this study is to develop wheat regression models to account for the impact of weather on wheat yield and quality and to predict (forecast) wheat yield and quality levels accurately. In other words, the primary purpose of the study is to use weather information to predict wheat yield and wheat quality and to select variables and functional forms to estimate parameters and then measure how well the developed models forecast.

2. Conceptual framework

Numerous studies (Dixon et al. (1994) and Kafumann and Snell (1997) and others) have used weather data to predict crop yields. Most studies have used precipitation and average temperature during the growing season as explanatory variables. For example, Yang et al. (1992) use planting season and growing season precipitation and average temperature to predict corn yields while. Hansen (1991) and Tannura et al. (2008) use calendar month precipitation and temperature variables to predict both soybean and corn yields. Even though biological stages of crops do not precisely correspond with calendar months, the use of monthly averages for weather variables as a proxy for crop stage weather is well supported. For example, Dixon et al. (1994) compare weather variables based on biological stages with variables based on fixed calendar months and find the forecasting performance and R^2 of the two models only change slightly.

As illustrated in Slafer and Rawson (1994, p. 398) and Acevedo et al. (2002, Fig. 3.1), wheat is typically classified in four stages of development: germination to emergence (E), from germination to double ridge (GS1), from double ridge to anthesis (GS2), and the grain filling period from anthesis to maturity (GS3) (FAO, 2002). The most crucial stages of wheat yield are from double ridge to anthesis (flowering) and from anthesis to maturity, since kernel number and weight are being determined at that time. Aitken (1974), Miralles and Slafer (1999), and Acevedo et al. (2002) argue that temperature and precipitation during the period from flowering to maturity have the primary influence on wheat yield.

Meanwhile, temperature and precipitation during grain filling are widely known to influence wheat quality characteristics. Graybosch et al. (1995), Johansson and Svensson (1998), Stone and Savin (1999), and Smith and Gooding (1999) find weather has deep impacts on grain quality. For instance, increased temperatures during grain filling tend to increase protein and reduce test weight. Stone and Savin (1999) argue that 70–80% of total protein is accumulated during grain filling.

Winter wheat in Oklahoma is typically planted in early September through the middle of November. Winter wheat harvest begins toward the end of May in southern Oklahoma and continues until about the middle of July (IPM Center, 2005). According to crop weather summary in Oklahoma Department of Agriculture (2000), wheat begins to double ridge and joint in February. In southwestern counties, wheat begins to head by the end of March. In April, anthesis begins, and some wheat in southern Oklahoma begins the grain filling period. Finally, wheat harvest begins approximately May 20th in the southern counties. Therefore, we concentrate on weather variables representing the February through April period for yield predictions and February through May for the quality predictions. We use the longer period for quality prediction both because late season weather has a larger impact on quality relative to yield and because the most obvious use of a pre-harvest quality prediction (establishing segregation strategies) do not require as much lead time as the obvious uses for yield predictions (re-positioning grain inventories and forward pricing grain).

As discussed previously, spatial weather data has the potential to improve crop yield and quality estimates. Eastering et al. (1998) use a fine spatial scale to reduce statistical bias from aggregation and confirm that the difference between observed and estimated yields is greatly reduced when data scale is disaggregated to around 37 mile × 50 mile, which is roughly the size of a county. Anselin (1988) assumes the dependent variable or residual at each location may be correlated with neighboring locations' dependent variables or residuals. For this spatially correlated data or residuals, the dependence is termed as spatial autocorrelation or spatial lag (contiguity) effect. This spatial correlation indicates that dependent variables or residuals are spatially autocorrelated and then violate the general assumption of statistically independent explanatory variables and errors. If the spatial lag effect is not considered, estimates will be biased and inconsistent.

In the context of our data, the weather, yield, and quality observations near the border of neighboring regions offers an opportunity for spatial autocorrelation. For instance, grain produced in one county could be shipped to an adjoining county. This could create spatial autocorrelation in the wheat quality data. The wheat yield data would not be affected because the yield data are based on Agriculture Research Service (ARS) yields, which are in turn based on producer reports of harvested production). The weather data could potentially create spatial autocorrelation in both the yield and quality data. Some cropland will be closer to a weather station in a neighboring county than to weather stations in its own county. Thus, weather measures in a neighboring county should help predict yield and quality and a spatial lag model should increase forecast accuracy.

Anselin et al. (2008) and Anselin and Bera (1998) express the neighbor relation with a spatial weights matrix, with elements w_{ij} that reflect the potential spatial relations between observations. The spatial weights matrix can be expressed as binary contiguity sharing a common border, distance contiguity including nearest neighbor locations, and inverse distance between two observations. For this study, we use the inverse of distance.

Anselin and Bera (1998) suggest two main alternative models of spatial autocorrelation: the spatial lag model, and the spatial error model. The main purpose of the former is to predict spatial patterns, while the latter's purpose is to increase the efficiency of estimates (Bongiovanni and Lowenberg-Deboer, 2001). A spatial lag model is used here since the explanatory variables in neighboring counties are expected to help predict our dependent variables. The general regression function can be expressed as:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \tag{1}$$

where $\boldsymbol{y}$ is a vector of dependent variables, $\boldsymbol{X}$ is the matrix of independent variables, and $\varepsilon \sim N(0,\ \sigma_\varepsilon^2 \boldsymbol{I})$ is a vector of stochastic error terms. The spatial lag model is

$$\boldsymbol{y} = \rho \boldsymbol{Wy} + X\beta + \boldsymbol{\varepsilon} \tag{2}$$

where ρ is the spatial autoregressive coefficient, $\boldsymbol{W}$ is a $N \times N$ spatial weight matrix (Greene, 2008). The spatial lag model is similar to including a lagged dependent variable in a time series model, except that endogeneity is created because the lagged effects go both directions. The weights matrix is standardized so that rows sum to 1. If $\rho > 0$, the dependent variable at each location is positively correlated with other location's dependent variables. The spatial lag model can be estimated with instrumental variables such as two stage least squares (2SLS) and generalized method of moments (GMM) or with maximum likelihood (ML) (Lambert et al., 2004). We use 2SLS here.

2.1. Data

The wheat yield data (from 1994 to 2009) are from 67 counties in Oklahoma and were obtained from National Agricultural Statistics Service (1994–2010). The time period for the yield model corresponds to the entire period during which Mesonet weather data was recorded. The yield is based on harvested acres. Oklahoma has 77 counties, but ten of them are not included due to having little wheat acreage. The cross-sectional time-series data are composed of 1072 observations (16 years times 67 counties). The wheat quality data are from Plains Grains, Inc (PGI)[1] with 2004–2009 used for estimation and 2010–2011 used for out-of-sample testing. The time period for the quality model corresponds to the entire period during which PGI recorded wheat quality information. PGI tests samples from up to 96 grain elevators and the sampling procedures are designed to make the samples representative of the elevator's total harvest receipts. PGI also collects samples from other states, but these states do not have the Mesonet weather data.

PGI collects representative wheat quality samples from both county and terminal elevators. Generally elevators take samples from each truckload arriving at the elevator, and the grain is sampled using a hand grain probe. Each elevator directly uses these samples to measure test weight and moisture content, and then these samples typically accumulate in a barrel. Finally, the elevator's barrel is sampled by PGI's representative using a hand grain probe.

The samples from county and terminal elevators are sent to the USDA, ARS Hard Winter Wheat Quality Laboratory in Manhattan, KS. Twenty-five quality parameters are analyzed in order to provide data that specifically describes the quality of wheat (PGI, 2009). Table 1 shows a list of wheat quality characteristics and basic descriptive statistics. Test weight (kg/hl) reflects the density of the sample and is an indicator of milling yield and the general condition of the wheat since problems that occur during the growing season or at harvest often reduce test weight. Wheat protein content measured on a 12% moisture base correlates with many important processing properties, such as water absorption and gluten strength, and to the texture and appearance of the finished product. "Higher protein dough usually absorbs more water and takes longer to mix" (PGI, 2009, p. 15). The PGI data also included the percentage of damaged kernels, shrunken and broken kernels, foreign material and total defects (the sum of the previous three factors). These quality factors are used in the determination of the numerical wheat grade under U.S. grain grading standards. Data on moisture content and dockage (non-wheat material) were also included.

In the later years of the data set the PGI quality measures were expanded to include measures based on the Single Kernel Characterization System (SKCS), and whole kernel near-infrared (NIR) tests for protein content. Two measures of flour quality (ash and falling number) were also added. This expanded quality data is not reported in Table 1 as it was unavailable for the first portion of the data set and so there were insufficient observations to use in regressions. The wheat quality predictions considered only test weight (lb/bu) and protein (12% mb) which have been recorded consistently and also have higher correlations with weather variables than that of the other quality data. These two quality characteristics also have economic importance since they have an important role in determining price received. Regressions were estimated using other wheat quality variables such as total defects and falling number as the dependent variable. These models were abandoned due to low in-sample predictability. Moisture content was not examined because it was assumed to be strongly affected by the daily weather conditions during harvest and does not reflect intrinsic quality.

The procedure matches elevators' quality data (test weight and protein) with weather data from the closest Mesonet station. This strategy means that one weather station per elevator, rather than a county average, is used to estimate wheat quality models. The quality data are from 96 elevators in 2010 located in 31 separate counties (Fig. 1).

Weather data (from January 1, 1994 to May 31, 2010) are obtained from the Oklahoma Mesonet (Fig. 1). Each of Oklahoma's 77 counties has one or more Mesonet stations. Weather data from the 67 counties with wheat production was used in the yield models while data from 31 stations were used in the quality models. The selected daily data are daily rainfall (cm), daily maximum (minimum) air temperature (°F), daily average air temperature (°F),

[1] Plains Grains Inc. (PGI) is located in Oklahoma and does a wheat quality survey and quality testing of hard red winter wheat to provide end-use quality information to the wheat buyer and producer and to publish the Wheat Quality Report PGI (2009).

Table 1
Descriptive statistics for wheat quality characteristics, 2004–2010.

	N	Mean	Maximum	Minimum	SD
Test weight (kg/hl)	492	79.0	85.4	69.6	2.41
Damaged kernels total (%)	492	0.19	1.8	0	0.25
Foreign material (%)	492	0.30	4.7	0	0.55
Shrunken and broken kernels (%)	492	1.38	5.1	0.2	0.67
Total defects (%)	492	1.86	6.5	0.3	1.02
Protein (12% mb)	492	12.2	16.0	8.8	1.30

Note: Kansas Grain Inspection Service (KGIS), Single Kernel Characterization System (SKCS), near-infrared (NIR).
Mb (moisture base), and SD (standard deviation).

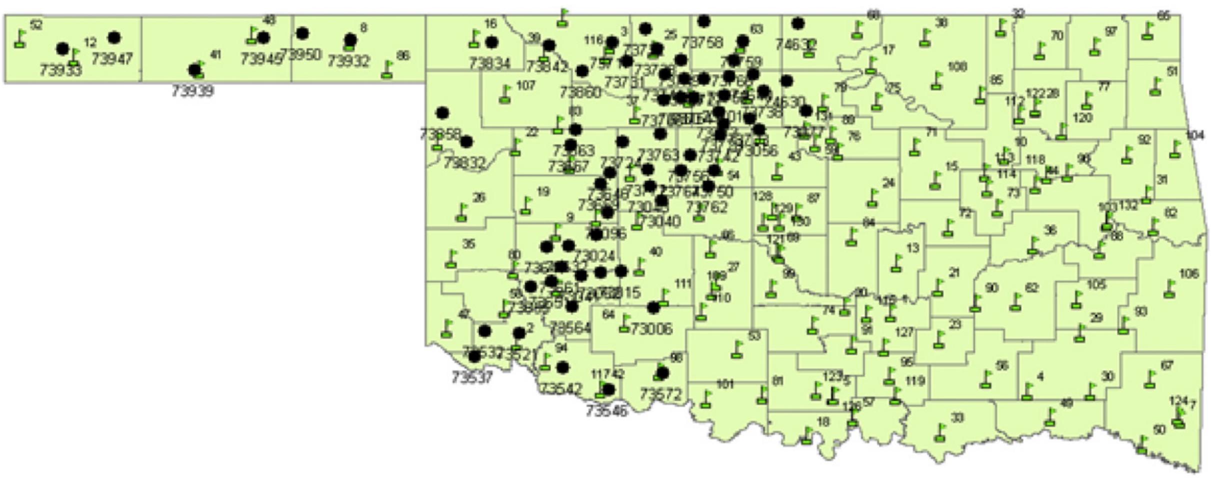

Fig. 1. Location of 2010 sample elevators and Mesonet Stations.
Notes: The study matches elevators' quality data (points) with Mesonet stations' weather data (flags) using the closest weather station.

total solar radiation (MJ m^{-2} d^{-1}), and growth degree days (GDD)[2]. For all Mesonet stations, the daily observations are aggregated to monthly averages. Most counties have only one weather station. For counties with multiple stations, an average of all stations in the county is used for yield models. The quality models use data from the weather station closest to the elevator where the quality sample was taken. Ten weather stations were added to the Mesonet system during the study period so for 3 elevators the closest weather station varies by year. In all three cases, the additional sites were within 3 miles of the original site so it is unlikely that this adjustment had significant impact on the model results.

3. Empirical model specification

To accurately specify the underlying relationships between yield and quality variables and weather variables, the study first examined the relationships between weather variables and yield and quality level using correlation coefficients and graphical displays of single variable nonparametric regression. Nonparametric regression allows exploring data and visualizing structure, and is useful for investigating nonlinear relations between dependent and independent variables[3]. Even though solar radiation and GDD were correlated with yields (Appendix, Table A.1), these variables were not statistically significant in regression models in the presence of other weather variables. Therefore those variables were excluded in the model specification. This result disagrees with Dixon et al. (1994) since the solar radiation variable in their model is essential. Precipitation is quadratically related with yield indicating that increased rainfall increases yield but at a decreasing rate. This suggests that after the crop's water deficit was decreased the response to additional rainfall was diminished. Temperature has a negative linear relation with yield. The optimum temperature for grain fill in wheat is generally considered to be between 20 °C and 22 °C which is lower than the average daily high temperature in April in most of Oklahoma's wheat producing counties. In light of these apparent relationships the yield response model uses linear and quadratic terms of precipitation and a linear term for temperature. The quality response model uses linear terms for precipitation, maximum temperature and minimum temperature since there was no evidence of the weather variables having a nonlinear relation with quality.

Several alternative functional forms are potentially appropriate for modeling linear relationships, including the linear, Cobb-Douglas, translog, square root, spline, and a semi-parametric method, which does not assume a specific functional form. Cobb–Douglas and linear model estimates show statistically significant individual coefficients and also relatively high pseudo R^2 (variance ratio) between in-sample annual predicted yield and annual actual yield during 1994–2009. Therefore, we selected the linear form and Cobb–Douglas form for the yield response model. Meanwhile, the quality response model adopts only a linear form. The models all have individual fixed effects and year random effects. The functional form can be written as[4]

[2] GDD = $[(T\max + T\min)/2] - \mathrm{Tb}$, 32 °F or 39.2 °F is used as the base temperature (Tb) for physiological process in wheat (Cao and Moss, 1989). The GDD vary with growing stage and provide a rough estimate of when a given growth stage occurs at a particular site.

[3] These figures are available in Lee (2011), but are too voluminous to present here.

[4] Linear form, Eq. (3), and the Cobb–Douglas form, Eq. (4), can be represented as matrices and vectors: $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{e}$ and $\ln \boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{e}$, $\boldsymbol{e} \sim \boldsymbol{N}(0, \sigma^2)$ and can be also

$$yield_{it} = \alpha_i + X_{it}\beta + \nu_t + \varepsilon_{it} \text{ (Linear)} \quad (3)$$

$$\ln(yield)_{it} = \alpha_i + \sum_j^n \beta_j \ln X_{jit} + \nu_t + \varepsilon_{it} \text{(Cobb} - \text{Douglas)} \quad (4)$$

and also can be expressed as a spatial lag model:

$$yield_{it} = \rho \boldsymbol{W}_N yield_{it} + \alpha_i + X_{it}\beta + \nu_t + \varepsilon_{it} \quad (3.1)$$

$$\ln(yield)_{it} = \rho \boldsymbol{W}_N \ln(yield)_{it} + \alpha_i + \sum_j^n \beta_j \ln X_{jit} + \nu_t + \varepsilon_{it} \quad (4.1)$$

where $yield_{it}$ is the wheat yield of county i and time t, α_i are individual fixed effects for counties, X_{jit} are the weather variables, and $\boldsymbol{W}_N$ is a $N \times N$ spatial weights matrix for cross-sectional dimension, $\varepsilon_{it} \sim N(0, \sigma_\varepsilon^2)$ is a stochastic error term, $\nu_t \sim N(0, \sigma_\nu^2)$ is a year random effect, and these error terms are assumed to be independent and identically distributed. The yield response model is composed of county fixed effect, year random effect, and three weather variables from February to April such as monthly average rainfall, squared average rainfall, and average temperature that correspond to before and after the anthesis period in Oklahoma, because yield is mostly determined before grain filling.

As discussed, wheat quality depends on the growth periods such as the milk development, heading, and ripening stages. In Oklahoma the wheat growth stages during March to May or June in the northern region contribute to grain filling which relates strongly to wheat quality. Additionally, the quality model reflects the agronomic tradeoff relationship between yield and quality by including yield as an explanatory variable:

$$Quality_{it} = \alpha_i + \delta(yield)_{it} + \sum_j^n \gamma_j X_{jit} + \nu_t + \varepsilon_{it} \quad (5)$$

$$Quality_{it} = \rho W_{N_t} Quality_{it} + \alpha_i + \delta(yield)_{it} + \sum_j^n \gamma_j X_{jit} + \nu_t + \varepsilon_{it} \quad (5.1)$$

where $Quality_{it}$ is either protein content (12% mb: moisture base) or test weight (kg/hl), and $\boldsymbol{W}_{N_t}$ is a $N_t \times N_t$ spatial weights matrix that varies by time t since the number of elevators varies by year. For protein, weather variables are monthly average maximum temperature, and monthly average rainfall from March to May; for test weight, weather variables used are monthly average rainfall for March, April, and May, and maximum and minimum temperatures for April and May.

4. Estimation method and procedure

The most generally used test for spatial autocorrelation is Moran's I test[5] (Griffith, 1987). Proc VARIOGRAM in SAS (SAS Institute Inc., 2004) is used to calculate the Moran's I statistic, Z score, and p-value for testing the hypothesis of no spatial autocorrelation.

Second, the study adopts the maximum likelihood estimation method (Greene, 2008, p. 400) and tests the heteroskedasticity of residuals using a likelihood ratio test. For the models where heteroskedasticity is found, multiplicative heteroskedasticity[6] is assumed (Greene, 2008, p. 170). While generalized method of moments (GMM) estimators are increasingly used and have the advantage of producing standard errors that are robust to nonnormality and heteroskedasticity, GMM is not used here since GMM methods provide less efficient estimates of the parameters of the mean equation than methods that correct for heteroskedasticity.

Yield models with corn and soybeans typically focus considerable effort on estimating time trends (Tannura et al., 2008; Harri et al., 2011). Oklahoma hard red winter wheat has not shown the increases in yields of corn and soybeans (Epplin, 1997). No significant time trend was found for wheat yields in the present study.

If the dependent variable values are correlated with values of nearby locations based on Moran's I statistic, the models include the weighted dependent variable of Eq. (2) and are estimated using instrumental variables. The spatial weights matrix for first ($\boldsymbol{W}$) and second order ($\boldsymbol{W}^2$) are constructed based on the inverse distance between two observations i and j where $w_{ij} = 1/\text{distance}_{ij}$ up to cut off miles and is otherwise 0. *GeoDa* software (Center for Spatially Integrated Social Science, 2004) was used to measure Arc distances among observations for yield with a cut off distance of 49.6 miles. For yield data, distances are measured from the center of the county. For quality data, distances are measured from zip code areas.

In addition, the developed models need to be evaluated for accuracy using out-of-sample forecasting tests rather than only a fitness test using historical data. Since the models are selected by pretesting, in-sample tests will overestimate their accuracy. To test the out-of-sample forecasting power for the developed models, the yield and quality forecasts are evaluated for 2010 and 2011. Also the forecasts are compared against a benchmark forecast of the previous actual six-year average. These tests are truly out of sample since the models and forecasts were developed before the 2010 and 2011 harvests. RMSE, MAE, and Theil's U_1 coefficient[7] as measures of forecasting accuracy for all developed models are used to evaluate the forecasting performance of the models. The first two forecast error statistics (RMSE and MAE) depend on the scale of the dependent variable as relative measures. The Theil coefficient is scale invariant and always lies between zero and one, that is, zero means a perfect fit (Quantitative Micro Software, 2000). Finally, in order to measure how the weather in the period immediately before harvest impacts yield and quality, the study conducts additionally the out-of-sample fit test using the longer period of weather data for the additional month of weather data on the estimates. Figure 2 provides a flow chart of the estimation and forecasting procedure.

5. Empirical results

This study first tests spatial autocorrelation for dependent variables. Table 2 shows a strong spatial lag effect with a Moran's I statistic of 0.0078 for yield and 0.0254 for protein with p-values of 0.0001. For test weight data, however, the p-value is 0.2642, indicating the null hypothesis $H_0: \rho = 0$, no spatial lag effect, could not be rejected. Therefore, the yield response and protein response models are estimated using spatial lags.

Second, the study estimated Eqs. (3)–(5.1) and then the residuals of the estimated models are tested for heteroskedasticity and

rewritten in expected mean form as $\boldsymbol{E}(\boldsymbol{y}) = \boldsymbol{X\beta}$ and $\boldsymbol{E}(\boldsymbol{y}) = \exp(\boldsymbol{X\beta} + \boldsymbol{\sigma}^2/2)$ respectively. Therefore, when we compare predictions (expected values) between the two functional forms, these mean forms are used.

[5] Moran's $\boldsymbol{I}$ statistic is $\boldsymbol{I} = [\boldsymbol{N}/\boldsymbol{S}][\boldsymbol{y}_t' \boldsymbol{W} \boldsymbol{y}_t / \boldsymbol{y}_t' y_t]$where, $\boldsymbol{y}_t$ is a vector of dependent values for each time period $\boldsymbol{t}$, $\boldsymbol{W}$ is a spatial weights matrix, $\boldsymbol{N}$ is observations, and $\boldsymbol{S}$ is the aggregation of all elements in $\boldsymbol{W}$. In general, a Moran's $\boldsymbol{I}$ statistic that is positive and near one indicates large positive autocorrelation while that is negative near one indicates large negative autocorrelation (ESRI 2006).

[6] If residuals are heteroskedastic, multiplicative heteroskedasticity is assumed: $\boldsymbol{\varepsilon}_{it} \sim \boldsymbol{N}(0, \sigma_{it}^2)$ or $\boldsymbol{\sigma}_{it}^2 = \exp(\boldsymbol{\alpha}'' \boldsymbol{Z}_{it})$ where $\boldsymbol{\alpha}$ is a parameter vector and $\boldsymbol{Z}_{it}$ is the matrix of independent variables.

[7] $U_1 = \left(\sqrt{\sum(\hat{y}_i - y_i)^2}\right) / \left(\sqrt{\sum(\hat{y}_i)^2} + \sqrt{\sum(y_i)^2}\right)$ where $\hat{y}_i$ and y_i are the prediction value and the corresponding actual value of county i respectively (Eviews, 2000, p. 337).

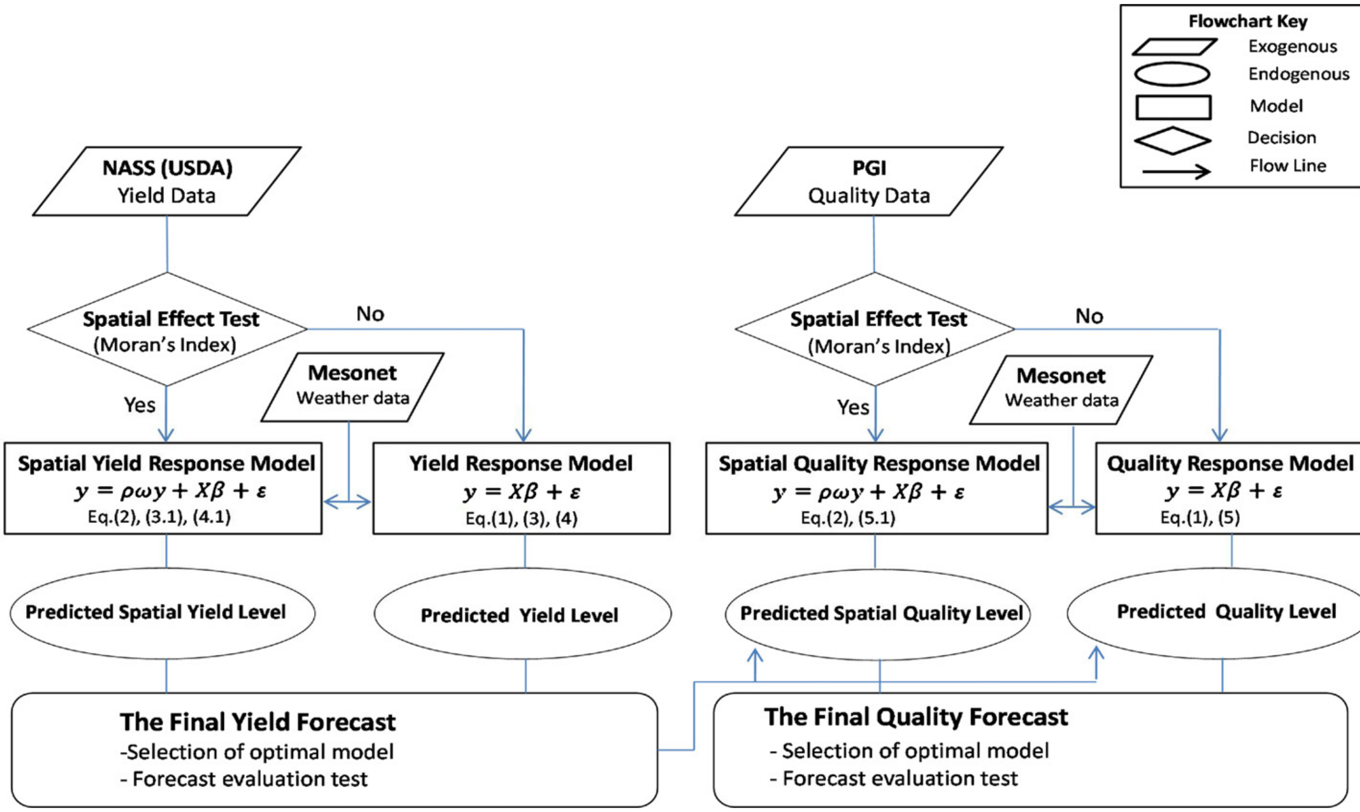

Fig. 2. The prediction procedure for yield and quality models.

Table 2
Tests of no spatial autocorrelation for wheat yield, protein, and test weight.

	Moran's index	Expected index	SD	*z*-score	*p*-value
Yield	0.00784***	−0.00091	0.000512	17.09	<.0001
Protein	0.02540***	−0.00219	0.00106	26.03	<.0001
Test weight	−0.00101	−0.00219	0.00106	1.12	0.2642

*** Significant at 1%, H_0: no spatial autocorrelation.

nonnormality. The test results show that linear yield models' LR statistics of homoskedasticity are smaller than the $\boldsymbol{X}^2$ critical value at the 5% level ($\boldsymbol{X}^2_{3,\,0.05} = 7.82$), that is, the null hypothesis of homoskedasticity is not rejected for linear yield models; while, the Cobb–Douglas yield models' calculated LR statistics are 19.1 for the general model and 16.1 for the spatial model, and thus the null hypothesis of homoskedasticity is rejected at the 5% level. On the other hand, all quality models' LR statistics are greater than the $\boldsymbol{X}^2$ critical value at the 5% level ($\boldsymbol{X}^2_{4,\,0.05} = 9.47$). When the null that residuals are homoskedastic is rejected, we assume multiplicative heteroskedasticity (see Greene, 2008, p. 523). Nonnormality tests show we can reject the null of normality for all models except test weight[8].

6. Comparing yield response models and spatial yield response models

Table 3 shows the estimated yield response models and spatial yield response models. Log likelihood statistics are used to confirm the spatial lag effect found with Moran's I statistic. The null hypothesis of no spatial lag effect ($\rho = 0$) is rejected. The estimated coefficients indicate how weather variables affect wheat yield. The weather variables are all significant at a critical level of 5% for all yield models. Precipitation has a positive relation with yield; while, squared precipitation and temperature are negatively related to yield. This suggests that wheat plants respond to rainfall but the response is diminished as the plant's water deficit is addressed. Higher average daily temperatures result in the crop being above the temperature for optimal grain filling, and are thus associated with lower yields. Finally, spatial yield response model's log likelihood statistic confirms that the accuracy of the yield response models is significantly improved by adding the spatially lagged dependent variable. When mesoscale weather information is used the effects of weather in adjoining areas should be considered.

Out-of-sample forecast error statistics for all yield models are summarized in Table 4. The models for yield perform better than the benchmark six-year average confirmed that weather information can be used by producers and agribusiness to improve yield forecasts.

The calculated statistics show that forecasts from the Cobb-Douglas spatial yield response model are slightly more accurate than forecasts from the linear yield response model. The linear model is slightly less accurate out of sample, as it was in sample.

[8] Our maximum likelihood estimation assumes normality, but possible consequences of making such an assumption are not of major concern here since our objective is forecasting.

Table 3
Yield model and spatial yield (kg/ha) model estimates, 1994–2009.

Variable	Yield response				Spatial yield response			
	Linear		Cobb–Douglas		Linear		Cobb–Douglas	
	Coeff.	*p*-Value	Coeff.	*p*-Value	Coeff.	*p*-Value	Coeff.	p-Value
Intercept	6632.26	<.0001	16.678	<0001	4579.52	0.0010	14.22	0.0003
Precipitation (cm)	15.06	<.0001	0.379	<.0001	7.92	0.0020	0.176	0.0135
Precipitation^2	−0.08	<.0001	−0.049	<.0001	−0.052	0.0003	−0.026	0.0019
Temperature	−108.27	<.0001	−2.605	<.0001	−91.66	<.0001	−2.557	<.0001
Spatial lag					0.742	0.0002	0.843	<.0001
−2 Log Likelihood	6496.2		-696.8		6481.7		−713.6	

Note: A first-order and second-order spatial weight matrices are used as instruments for the spatial lag term as WX, W^2X.

Table 4
Out of sample forecast error statistics for yield models, 2010–2011.

Forecast errors	Average (2004–2009)	Yield response		Spatial yield response	
	wo/weather effects[a]	Linear	Cobb–Douglas	Linear	Cobb–Douglas
RMSE	7.436	6.623	5.775	6.397	5.680
MAE	6.136	5.063	4.108	4.793	4.068
Theil U_1	0.0081	0.00826	0.00761	0.00817	0.00744

Note: These statistics were computed on the basis of the yield observations which cover the years 2010–11.
[a] $yield_{it} = \alpha_i + \nu_t + \varepsilon_{it}$, α_i and ν_t are county fixed effects and random year effects.

7. Protein response model and weight response model

The estimated quality response models of Eqs. (5) and (5.1) for wheat characteristics: protein and test weight are reported in Table 5. Yield and weather variables are all significant at the 5% level. Yield and protein are commonly thought to be inversely related in wheat production with the layman's explanation being that higher yields spread the available nitrogen, which is basis of protein content, over more bushels. Understanding this relationship is helpful in considering the model results. Precipitation and maximum temperature positively affect protein and test weight. Higher temperatures create larger, more denser kernels (increased test weight) while reducing the number of kernels per plant (decreased yield). This lower yield in turn increases the protein percentage. Minimum temperature is negatively related with test weight. As expected, yield was negatively related to protein and positively related to test weight. In terms of elasticities, a 1% rise in yield decreases average protein by 0.25% and increases average test weight by 0.28%. These relationships between weather variables and wheat quality are consistent with the findings of Johansson and Svensson (1998) and Smith and Gooding (1999), who find warm temperatures affect crude protein positively, and precipitation at the end of the season has significant positive correlation with protein concentration.

Even though the spatial lag term is not significant using a Wald test, the more reliable likelihood ratio test does reject the null hypothesis of no spatial lag. The consideration of weather in adjoining areas does not appear to be as essential in predicting wheat quality as it was for wheat yield but the model's result still indicate it is useful.

The forecasting accuracy of the quality models was evaluated over the 2010 and 2011 crop years. Table 6 shows the forecast error statistics with protein and test weight response models. First, the study estimates the general quality models using actual yield. Second, predictions from the spatial quality models are obtained by substituting predicted yields for actual yields. The forecast error values indicate that the accuracy of the quality models can be improved by adding predicted yield from the spatial protein response model rather than that of the no spatial protein response model. Also RMSE values show that the forecasting performance of the models combining spatially predicted yield are improved. Additionally, average values (2004–2009) are used as benchmark forecasts and as expected, all forecast error values for average values are higher than those of the weather models. The models for protein perform better relative to the benchmark six-year average than do the test weight models.

8. Implications of the value of additional month of weather data on estimates

Our baseline procedures assumed that predictions using weather data from late in the growing season would be less

Table 5
Protein model and test weight model estimates, 2004–2009.

Variable	Protein (12%, mb)				Test weight (kg/hl)	
	No-spatial		Spatial		No-spatial	
	Coeff.	*p*-value	Coeff.	*p*-value	Coeff.	*p*-value
Intercept	2.986	0.0587	2.708	0.1417	43.739	0.0058
Yield (kg/ha)	−0.0015	<.0001	−0.0014	<.0001	0.0021	<.0001
Precipitation (cm)	0.0055	0.0006	0.0067	0.0009	0.0261	0.0003
Max. temp.	0.149	<.0001	0.130	<.0001	0.5800	0.0001
Min. temp.					-0.3644	0.0108
Spatial lag			0.138	0.3843		
−2 Log likelihood	1045.5		1019.1		1482.9	

Note: A first and second-order spatial weight matrices are used as instruments for the spatial lag term as WX, W^2X.

Table 6
Out of sample forecast error statistics of wheat quality, 2010–2011.

	Average (2004–2009)	2010–2011			
		No-spatial		Spatial	
	wo/weather effects	W/spatially predicted yield	W/generally predicted yield	W/spatially predicted yield	W/generally predicted yield
Protein					
RMSE	1.080	0.990	1.004	0.933	0.956
MAE	0.870	0.772	0.777	0.753	0.760
Theil U_1	0.0074	0.00685	0.0065	0.006091	0.00625
Test weight					
RMSE	1.495	1.777	1.826		
MAE	1.196	1.379	1.431		
Theil U_1	0.00041	0.000485	0.000499		

Note: These statistics were computed on the basis of the protein and test weight observations which cover the years 2010–2011.

Table 7
Out of sample fit test using the longer period (Feb–May) for yield models.

Forecast errors	Average (2005–2009)	Yield response		Spatial yield response	
	wo/weather effects	Linear	Cobb–Douglas	Linear	Cobb–Douglas
RMSE	7.436	6.951	6.287	6.419	6.035
MAE	6.136	5.403	4.704	4.871	4.564
Theil U_1	0.0081	0.00855	0.00796	0.00813	0.00777

Note: These statistics were computed on the basis of the yield observations which cover the years 2010–2011.

desirable because the forecasts are not available with enough lead time to make logistical and marketing decisions prior to harvest. However, some decision makers might be able to respond to prediction information within a short lead time so it is useful to explore whether prediction models using late season weather information would result in improved forecasts. Expanding the data periods also provide insights into how the weather conditions immediately prior to harvest impacts yield. In the case of our quality predictions, our baseline procedures assumed that decision makers could respond to quality forecasts with a very short lead time. The baseline data time period also reflected the fact that wheat quality is considered to be determined late in the growing season. For both of these reasons, the quality predictions were based on the longer period of weather data. In this case, it is interesting to examine how a quality prediction model with a shorter data series (longer lead time for decisions) would have performed. As with the yield models, the difference in forecast accuracy between the shorter and longer data series models gives insights as to the extent that wheat quality is impacted by the late season weather conditions.

For the yield model, the study also calculated out-of-sample fit tests using the weather data from May. The extended weather data (Feb–May) are used to estimate new regressions and predict yields using the new estimates. The forecast errors reported in Table 7 are larger than those of the original yield models (Feb–April). Adding the additional weather information immediately prior to harvest appears to only add noise and does not improve yield forecasts. A possible explanation is that since the number of heads and the size of the heads are well established by the beginning of May and so late season weather has little effect on yields.

The regressions for quality data included May so we also consider regressions by dropping the May data. Forecast error statistics for the new quality models (March–April) are summarized in Table 8. The calculated statistics show only small differences between the original quality model and the new quality model; however new forecasts are slightly less accurate than those from the original quality response models. The results confirm the assumption that late season weather impacts wheat quality. A decision maker who could respond to a quality prediction under a short lead time would be somewhat better off with a model based on late season weather data. However, these results show that the quality models using the shorter period of weather data still provide useful forecasts.

9. Summary and conclusions

This study estimates wheat regression models to account for the effect of weather on wheat yield, protein, and test weight and to forecast wheat yield and the two wheat quality measures. The explanatory variables include precipitation and temperature for growing periods that correspond to biological wheat development stages. These models include county fixed effects, crop year random effects, and a spatial lag effect. Yield and quality level are strongly influenced by weather variables. For yield, precipitation has a positive, but non-linear relationship with yield, while average monthly temperatures are negatively related to yield.

Table 8
Out of sample fit test using shorter period (March–April) for quality models.

	Average (2004–2009) wo/weather effects	(March–May) Original	(March–April) Shorter
Protein			
RMSE	1.080	0.990	1.266
MAE	0.870	0.772	0.991
Test weight			
RMSE	1.495	1.777	1.893
MAE	1.196	1.379	1.547

Note: These statistics were computed on the basis of the protein and test weight observations which cover the years 2010–2011.

Precipitation and maximum temperature positively affect protein and test weight while minimum temperature is negatively related with test weight. Yield and protein are inversely related while yield and test weight are positively related. In the forecast evaluation, the forecasting ability of both yield and protein models is enhanced by adding the spatial lag effect, i.e. considering weather information in adjoining areas. Out of sample forecasting tests show the developed models are more accurate than using a benchmark six-year average and suggest that Mesoscale weather information can be used to generate improved yield and quality predictions

These predictions could be widely used and could be important to producers optimizing late season agronomic (such as fungicide applications) and marketing decisions. Grain elevators and agribusinesses could also use predictions to position grain inventories prior to harvest and for making decisions on segregation strategies and sales contracts. For example, our models predicted the low protein that occurred in 2010. During the 2010 season some Oklahoma elevators had contracted for rail shipments which specified a minimum protein level. Due to the low protein content of the crop they were unable to fulfill the contracts and were also subject to demurrage payments on the rail cars they were unable to use. If these elevators had been able to forecast a below average protein crop they could have avoided entering into the contracts and planned to store the wheat until a market for low protein wheat developed. The forecasts are being presented to industry groups and the interest generated shows that there is great potential value in weather based forecasts of wheat yield and quality.

While the data used in the model represent a unique set of Mesoscale weather information and wheat quality data, the results are limited by the relatively short time series of quality data available. The models' performance can likely be improved over time when a longer time series of data can be used. As more data becomes available it may also be possible to predict additional quality factors such as shrunken and broken kernels (a grade factor), falling number (a measure of flour quality) or kernel hardness (Brorsen et al. (2012) find kernel hardness is a predictor of dough quality).

It should also be noted that yield and forecast accuracy in a particular season could also be influenced by extreme weather conditions, such as a late season freeze which could have impacts not captured by our linear model using monthly average weather data. Climate change could also obviously impact a prediction model based on historical weather data.

Appendix A.

Table A.1.

Table A.1
Correlation coefficients among yield, protein, weight and weather variables.

	Yield	*p*-Value	Protein	*p*-Value	Weight	*p*-Value
Yield	1.00	<.0001	−0.47	<.0001	0.15	0.0019
Precipitation						
February	0.49	<.0001	−0.52	<.0001	0.06	0.1912
March	0.04	0.4098	−0.20	<.0001	−0.15	0.0014
April	−0.05	0.3255	−0.16	0.0005	0.08	0.1028
May	−0.22	<.0001	−0.09	0.0613	0.10	0.0406
Avg. temperature						
February	−0.43	<.0001	0.06	0.1734	−0.05	0.2891
March	−0.31	<.0001	−0.02	0.6924	−0.12	0.0094
April	−0.42	<.0001	0.34	<.0001	0.20	<.0001
May	−0.02	0.6899	0.00	0.9968	0.06	0.1994
Max. temperature						
February	−0.50	<.0001	0.29	<.0001	0.08	0.0815
March	−0.35	<.0001	0.10	0.0396	−0.07	0.1247
April	−0.38	<.0001	0.45	<.0001	0.27	<.0001
May	0.19	<.0001	0.08	0.0742	0.13	0.0071
Min. temperature						
February	−0.14	0.0029	−0.22	<.0001	−0.13	0.0045
March	−0.25	<.0001	−0.11	0.0174	−0.18	0.0001
April	−0.35	<.0001	0.11	0.0171	0.10	0.0417
May	−0.20	<.0001	−0.16	0.0004	−0.07	0.1368
Growing degree days						
February	−0.55	<.0001	0.07	0.1295	−0.12	0.0136
March	−0.32	<.0001	−0.01	0.8454	−0.13	0.0069
April	−0.39	<.0001	0.32	<.0001	0.22	<.0001
May	0.02	0.7532	−0.06	0.1994	0.06	0.2246
Solar radiation						
February	−0.50	<.0001	0.35	<.0001	0.14	0.0027
March	0.51	<.0001	−0.12	0.0129	0.13	0.0056
April	0.16	0.0013	0.22	<.0001	0.30	<.0001
May	0.40	<.0001	0.06	0.177	0.28	<.0001

References

Aitken, Y., 1974. Flowering Time, Climate and Genotype. Melbourne University Press, Melbourne.

Anselin, L., 1988. Spatial Econometrics: Methods and Models. Kluwer Academic Publishers, Dordrecht (the Netherlands).

Anselin, L., Bera, A.K., 1998. Spatial dependence in linear regression models with an introduction to spatial economics. In: Ullah, A., Giles, D.E.A. (Eds.), Handbook of Applied Economic Statistics. Marcel Dekker, Inc., New York, pp. 237–289.

Anselin, L., Gallo, J., Jayet, H., 2008. Spatial panel econometrics. In: Mayas, L., Sevestre, P. (Eds.), The Econometrics of Panel Data. Springer-Verlag, New York, pp. 625–660.

Acevedo, E.P., Silva, Silva, H., 2002. Wheat growth and physiology. In: Curtis, B.C., Rajaram, S., Macpherson, H.G. (Eds.), Bread Wheat Improvement and Production. FAO, Roma, Italy, pp. 39–70.

Bannayan, M., Crout, N.M.J., Hoogenboom, G., 2003. Application of the CERES-wheat model for within-season prediction of winter wheat yield in the United Kingdom. Agron. J. 95, 114–125.

Bechter, D.M., Rutner, J.L., 1978. Forecasting with statistical models and a case study of retail sales. Econ. Rev. (Kansas City) 63 (3), 3–11.

Bongiovanni, R., Lowenberg-Deboer, J., (2001). Precision agriculture: economics of nitrogen management in corn using site-specific crop response estimates from a spatial regression model. Paper Presented at American Agricultural Economics Association Annual Meeting, 5–8 August.

Brorsen, B.W., Rayas-Duarte, P., Ji, D., 2012. Predicting rheological properties of wheat dough based on wheat characteristics. J. Agric. Sci. 4 (3), 79–84.

Center for Spatially Integrated Social Science, 2004. GeoDa 0.95-I Release Notes. Urbana. IL.

Chen, C.C., Chang, C.C., 2005. The impact of weather on crop yield distribution in Taiwan: some new evidence from panel data models and implications for crop insurance. Agric. Econ. 33, 503–511.

Claborn, W.B., 1998. Testing and calibration of the CERES-wheat model in Oklahoma. M.S. Dissertation, Oklahoma State University, Stillwater, Oklahoma.

Dixon, B.L., Hollinger, S.E., Garcia, P., Tirupattur, V., 1994. Estimating corn yield response models to predict impact of climate change. J. Agric. Resour. Econ. 19 (1), 58–68.

Duchon, C.E., 1986. Corn yield prediction using climatology. J. Clim. Appl. Meteorol. 25, 581–590.

Eastering, W.E., Weiss, A., Hays, C.J., Mearns, L.O., 1998. Spatial scales of climate information for simulating wheat and maize productivity: the case of the US Great Plains. Agric. Forest Meteorol. 90, 51–63.

Epplin, F.M., 1997. Wheat yield response to changes in production practices induced by program provisions. J. Agric. Resour. Econ. 22, 333–344.

FAO, 2002. Bread Wheat Improvement and Production. FAO, Roma, Italy.

Graybosch, R.A., Peterson, C.J., Baenziger, P.S., Shelton, D.R., 1995. Environmental modification of hard red winter wheat flour protein composition. J. Cereal Sci. 22, 45–51.

Greene, W.H., 2008. Econometric Analysis, sixth ed. Prentice Hall, Upper Saddle River, NJ.

Griffith, D.A., 1987. Spatial Autocorrelation. Association of American Geographers, Washington, DC.

Guttieri, M.J., Ahmad, R., Stark, J.C., Souza, E., 2000. End-use quality of six hard red spring wheat cultivars at different irrigation levels. Crop Sci. 40, 631–635.

Hansen, L., 1991. Farmer response to changes in climate: the case of corn production. J. Agric. Econ. Res. 43, 18–25.

Harri, A., Coble, K.H., Ker, A.P., Goodwin, B.K., 2011. Relaxing heteroskedasticity assumptions in area-yield crop insurance rating. Am. J. Agric. Econ. 93 (3), 703–713.

Irwin, S.H., Good, D.L., Tannura, M.A., 2008. Forming expectations about 2008 U.S. corn and soybean yields-application of crop weather models that incorporate planting progress. Marketing and Outlook Research Report 2008-01, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, Illinois.

Integrated Pest Management Center (IPM Center), 2005. Crop profile for wheat in Oklahoma. Retrieved From:http://www.ipmcenters.org/cropprofiles/docs/okwheat.pdf. (1.11.2010).

Johansson, E., Svensson, G., 1998. Variation in bread-making quality: effects of weather parameters on protein concentration and quality in some Swedish wheat cultivars grown during the period 1975–1996. J. Sci. Food Agric. 78, 109–118.

Johansson, E., Prietolinde, M.L., Gissen, C., 2008. Influences of weather, cultivar and fertilizer rate on grain protein polymer accumulation in field-grown winter wheat, and relations to grain water content and falling number. J. Sci. Food Agric. 88, 2011–2018.

Jones, C.A., Kiniry, J.R., 1986. CERES-Maize: A Simulation Model of Maize Growth and Development. Texas A&M Univ., College Station, Texas.

Just, R.E., Rausser, G.C., 1981. Commodity price forecasting with large-scale econometric models and the futures market. Am. J. Agric. Econ. 63, 197–208.

Kafumann, R.K., Snell, S.E., 1997. A biophysical model of corn yield: integrating climatic and social determinants. Am. J. Agric. Econ. 79, 178–190.

Kandiannan, K., Karithikeyan, R., Krishnan, R., Kailasam, C., Balasubramanian, T.N., 2002. A crop-weather model for prediction of rice yield using an empirical–statistical technique. J. Agron. Crop Sci. 188, 59–62.

Lambert, D.M., Lowenberg-Deboer, J., Bongiovanni, R., 2004. Spatial regression models for yield monitor data: a case study from Argentina. Precision Agric. 5, 579–600.

Lee, B.H., 2011. Forecasting wheat yield and quality conditional on weather information and estimating construction costs of agricultural facilities. Ph.D. dissertation, Department of Agricultural Economics, Oklahoma State University.

Miralles, D.J., Slafer, G.A., 1999. Wheat development. In: Satorre, E.H., Slafer, G.A. (Eds.), Wheat: Ecology and Physiology of Yield Determination. Food Products Press, New York, pp. 13–43.

National Agricultural Statistics Service. Crop Production Report. Washington DC: U.S. Department of Agriculture, January 1994–January 2010. Available at <http://usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do?documentID=1047>.

Oklahoma Department of Agriculture, 2000. Oklahoma Agricultural Statistics 1999. Oklahoma City, OK.

Oklahoma Mesonet County Climate Data. Retrieved from <http://climate.mesonet.org/county_climate/Products/QuickFacts/> (1.06.2010).

Oklahoma Mesonet Premium data available only to subscribed members. Retrieved http://www.mesonet.org/premium/statistics.html (01.06.10).

Plains Grains, Inc. (PGI), 2010. Wheat quality data available only to private contracted members. Retrieved 2004–2010.

Plains Grains, Inc. (PGI), 2009. 2009 Hard Red Winter Wheat Quality Report. The Annual Report, PGI, OK. Available at: <http://www.plainsgrains.org/pdfs/PGI_HRW_2009.pdf>.

Quantitative Micro Software, 2000. Eviews 4.0 User's Guide. Quantitative Micro Software, Irvine, CA, USA.

Regnier, S., Holcomb, R., Rayas-Durate, P., 2007. Relating Wheat Quality to End-Product Quality, FAPC-129. Oklahoma State University, Stillwater, Oklahoma.

SAS Institute, 2004. The Variogram Procedure. SAS 9.2 Help and Documentation. Cary. NC.

Slafer, G.A., Rawson, H.M., 1994. Sensitivity of wheat phasic development to major environmental factors: a re-examination of some assumptions made by physiologists and modellers. Austr. J. Plant Physiol. 21, 393–426.

Smith, G.P., Gooding, M.J., 1999. Models of wheat grain considering climate cultivar and nitrogen effects. Agric. Forest Meteorol. 94, 159–170.

Stone, P.J., Savin, R., 1999. Grain quality and its physiological determinants. In: Satorre, E.H., Slafer, G.A. (Eds.), Wheat: Ecology and Physiology of Yield Determination. Food Products Press, New York, pp. 114–122.

Tannura, M.A., Irwin, S.H., Good, D.L., 2008. Weather, technology, and corn and soybean yields in the U.S. Corn Belt. Marketing and Outlook Research Report 2008–2001, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, Illinois.

Thompson, L.M., 1969. Weather and technology in the production of corn in the U.S. Corn Belt. Agron. J. 61 (1969), 453–456.

Thompson, L.M., 1970. Weather and technology in the production of soybeans in the central U.S. Agron. J. 62, 232–236.

Tsvetsinskaya, E.A., Mearns, L.O., Mavromatis, T., Gao, W., McDaniel, L., Downton, M.W., 2003. The effect of spatial scale of climate change scenarios on simulated maize, winter wheat, and rice production in the southeastern United States. Clim. Change 60, 37–71.

Wang, E., Williams, J.R., Little, B.B., 2006. Prediction of weather event associated crop yield losses in Kansas. Paper Presented Southern Agricultural Economics Association Annual Meeting, Orlando, Florida, 5–8 February.

Walker, G.K., 1989. Model for operational forecasting of western Canada wheat yield. Agricultural and Forest Meteorology 44, 339–351.

Woolfolk, C.W., Raun, W.R., Johnson, G.V., Thomason, W.E., Mullen, R.W., Wynn, K.J., Freeman, K.W., 2002. Influence of late-season foliar nitrogen applications on yield and grain nitrogen in winter wheat. Agron. J. 94, 429–434.

Yang, S.R., Koo, W.W., Wilson, W.W., 1992. Heteroskedasticity in crop yield models. J. Agric. Resour. Econ. 17 (1), 103–109.