# Using Bayesian Kriging for Spatial Smoothing of Trends in the Means and Variances of Crop Yield Densities

**Bart Niyibizi[1], B. Wade Brorsen[2], and Eunchun Park[3]**

[1]PhD Candidate, Department of Agricultural Economics, Oklahoma State University
bart.niyibizi@okstate.edu
[2]Regents Professor and A.J. & Susan Jacques Chair, Department of Agricultural Economics, Oklahoma State University
wade.brorsen@okstate.edu
[3]Assistant Professor, Department of Agricultural Economics, Mississippi State University
eunchun.park@msstate.edu

***Selected Paper prepared for presentation at the 2018 Agricultural & Applied Economics Association Annual Meeting, Washington, D.C., August 5-August 7***

Funding: This research received funding from the Oklahoma Agricultural Experiment Station and the National Institute for Agriculture (NIFA) Hatch project OKL02939 as well as the A.J. and Susan Jacques Chair.

**Abstract**

Crop yield forecasts are useful for several purposes such as rating crop insurance and government budget predictions, and allocation of barges and railcars. We use Bayesian Kriging for spatial smoothing of yield density parameters, including time trends. There is a paucity of useful historical yield data for counties, but properly using other counties' information in the estimation of a county's yield density alleviates the efficiency problem of not having enough observations. Using a Gaussian spatial process, our yield density parameters are spatially correlated. We spatially smooth multiple parameters, including time trends. Our model can handle unbalanced panel data. Using corn yield data from Illinois and Iowa, we find that the yield mean has increased faster in northern counties, but that the yield variance has increased faster in southern counties. Forecast accuracy is similar with our model and with Bayesian Model Averaging and a mixture of normal distribution, but our approach is the only one that provides the spatial smoothing desired by the Risk Management Agency.

Key words: Spatial econometrics, Bayesian Kriging, Bayesian hierarchical modelling, Bayesian spatial smoothing, yield density, corn yield, density forecasting.

## 1. Introduction

Crop yield forecasts are useful for several purposes. Farmers can use yield forecasts in making decisions such as entering commodity contracts, buying crop insurance, or determining land values. The Risk Management Agency (RMA) of the USDA calculates area yield crop insurance premiums based on forecasted county yields. With the emergence of the need for risk management, density forecasts have become attractive rather than point forecasts. There is uncertainty in forecasting, and relative to point forecasts, density forecasts provide more information. In crop insurance, accurate rating of an insurance program is crucial. For the RMA, inaccurate rates could lead to substantial losses that would result from insurance indemnities considerably exceeding collected premiums. As Harri et al. (2011) remark, if an insurance policy is not actuarially fair, producers with high knowledge of yield risk may arbitrage the program or select against it.

Yield density forecasts are generally made before the beginning of a crop's growing season and cannot use yield-determining information observed during the season such as weather or production inputs. In addition, the yield forecast for a county is traditionally based largely on a density estimated from that county's past yield observations. This poses an efficiency problem as the time series of county yields are not long enough. In addition, there is a concern that structural changes in crop genetics, weather patterns, and farming practices may imply that current yield observations are from very different distributions than observations from several decades ago, making yield time series even shorter.

To surmount this problem, spatial data can be useful in estimating crop yield densities. Due to similarities in climate, soil type, and shared information, the assumption that there are

similarities in densities of neighboring counties is plausible. In spatial econometrics, the traditional spatial error model of Anselin (1988) treats spatial correlation as being entirely in the error term and restricts all locations to have the same model parameters.

However, there have been efforts to incorporate spatial data at the parameter level. For example, Harri et al. (2011) estimate parameters of yield densities at the county level but the parameters have restrictions determined at the district level. Ker, Tolhurst, and Liu (2016) estimate the yield density of a county by first estimating a posterior density using the county's own data and then finding a weighted average of posterior densities from all counties, based on the Bayesian Information Criterion (BIC).

Ker, Tolhurst, and Liu (2016) assume that there are yield density similarities, but they do not make assumptions about the source of the similarities. The time trend used in these models is likely the most important parameter in making density forecasts. Yet, their time trend is smoothed in the same way as the other distribution parameters. Park et al. (2016, 2018) and Harri et al. (2011) used two-step methods where the trend was estimated separately from the other parameters.

The objective of this study is to develop Bayesian Kriging for spatial smoothing of yield density parameters. We assume that the similarity in yields comes from the correlation between parameters that define yield densities. This technique allows the smoothing of parameters and gives the ability to estimate yield densities even for counties with missing observations. Previous applications of Bayesian Kriging in the agricultural economics literature (e.g., Ozaki et al. 2008) perform spatial smoothing of the intercept, while Park et al. (2016, 2018) use it for spatial smoothing of two distributional parameters. In this paper, we contribute to the spatial

econometrics literature by extending Bayesian Kriging to the general heteroscedastic error regression problem. Our work is the second paper, after Reich (2012), to extend Bayesian Kriging to the slope parameters in the linear regression problem rather than just the intercept as in Ozaki et al. While we assume normality, our general approach can be extended to nonnormal distributions.

Forecasts of expected yield densities are based on past yield observations. But there are counties that lack yield records for certain years. The Ker, Tolhurst, and Liu (2016) approach is restricted to balanced data. Using interpolation techniques, our Bayesian Kriging approach makes it possible to include counties without a full yield history in the estimation and predictions.

## 2. Methods

Suppose there are *N* counties for which yield densities have to be estimated in year *t* (*t*=1,…,*T*). In this study, each county's yield is assumed to have a normal density. It is common practice in yield forecasts to assume time trends in the yield densities. Examples of what time trends capture include advances in production technology and improvements in seed genetics. Unlike past research, we use a simple time trend rather than a spline because we are limiting the time period. We assume that there are time trends in both the mean and variance. The yield $Y_{it}$ in year *t* for county *i* can thus be expressed as:

$$Y_{it} = \alpha_i + \beta_i t + \varepsilon_{it} \quad (1)$$

where $\alpha_i$ and $\beta_i$ are the parameters that determine the mean of the yield, and $\varepsilon_{it}$ is a random error term with mean zero and variance $\sigma_{it}^2$. As noted above, we assume heteroskedasticity through a time trend in the variance equation. The yield variance $\sigma_{it}^2$ is:

$$\sigma_{it}^2 = \gamma_i + \delta_i t + \nu_{it}, \tag{2}$$

where $\gamma_i$ is the intercept, $\delta_i$ is the slope or time trend, and $\nu_{it}$is a random error term with mean zero. This assumption is different from assumptions made in previous research such as Harri et al. (2011), where the variance is assumed to be proportional to a power of predicted yields.

Note that there are four parameters that define the mean and variance equations for county *i*'s yield densities. For convenience of notation we will sometimes collectively refer to the parameters as $\boldsymbol{\theta_i}$, that is, $\boldsymbol{\theta_i} = (\alpha_i, \beta_i, \gamma_i, \delta_i)$. Equation (1) is then written as:

$$Y_{it} \sim p_1(Y_{it}|\boldsymbol{\theta_i}). \tag{3}$$

Equation (3) can further be written in matrix form as:

$$\boldsymbol{Y} \sim p_1(\boldsymbol{Y}|\boldsymbol{\theta}), \tag{4}$$

where $\boldsymbol{Y}$ is the $N \times T$ matrix of yields from all counties and all years and $\boldsymbol{\theta}$ is the $N \times 4$ matrix of yield density parameters for all counties. In the remainder of the paper whenever we mention parameters without more details we are referring to $\boldsymbol{\theta}$.

The Bayesian approach to inference assumes that model parameters are random variables. This is in contrast to the frequentist approach that assumes that the observed data are random but that the model parameters are fixed. The parameters $\boldsymbol{\theta}$ follow a density

$$\boldsymbol{\theta} \sim p_2(\boldsymbol{\theta}|\boldsymbol{\lambda}). \tag{5}$$

The prior distributions of parameters represent prior beliefs about the parameters before observing the data, although in some cases data can be used in forming priors. The parameters of the prior distributions are called hyperparameters ($\boldsymbol{\lambda}$ in this case). The hyperparameters comprise the parameters that determine the parameters $\boldsymbol{\theta}$, including the Kriging parameters (sill and range) that define the spatial similarities among parameters of counties' yield densities based on distance. The prior distribution for our Kriging approach is a Gaussian spatial process. Bayesian inference is based on the posterior distribution. The posterior is proportional to the product of the likelihood $p_1(\boldsymbol{Y}|\boldsymbol{\theta})$ and the prior $p_2(\boldsymbol{\theta}|\boldsymbol{\lambda})$.

Our Kriging approach uses Bayesian hierarchical modelling, where hyperparameters have prior distributions of their own, called hyperpriors:

(6) $$\boldsymbol{\lambda} \sim p_3(\boldsymbol{\lambda}).$$

In our Bayesian hierarchical modelling, the determination of yield densities is based on the posterior distribution of the parameters. The joint posterior distribution of the parameters is proportional to the product of: (i) the likelihood, (ii) the prior, and (iii) the hyperprior,

(7) $$p(\boldsymbol{\theta}, \boldsymbol{\lambda}|\boldsymbol{Y}) \propto p_1(\boldsymbol{Y}|\boldsymbol{\theta}, \boldsymbol{\lambda}) p_2(\boldsymbol{\theta}|\boldsymbol{\lambda}) p_3(\boldsymbol{\lambda}),$$

and is expressed mathematically as:

(8) $$p(\boldsymbol{\theta}, \boldsymbol{\lambda}|\boldsymbol{Y}) = \frac{p_1(\boldsymbol{Y}|\boldsymbol{\theta},\boldsymbol{\lambda}) p_2(\boldsymbol{\theta}|\boldsymbol{\lambda}) p_3(\boldsymbol{\lambda})}{\iint p_1(\boldsymbol{Y}|\boldsymbol{\theta},\boldsymbol{\lambda}) p_2(\boldsymbol{\theta}|\boldsymbol{\lambda}) p_3(\boldsymbol{\lambda})}\ .$$

Note that the likelihood notations $p_1(\boldsymbol{Y}|\boldsymbol{\theta}, \boldsymbol{\lambda})$ in equations (7) and (8) and $p_1(\boldsymbol{Y}|\boldsymbol{\theta})$ in equation (4) are equivalent, as the likelihood of $\boldsymbol{Y}$ depends on $\boldsymbol{\lambda}$ only through $\boldsymbol{\theta}$.

As mentioned above, we assume that yields are normally distributed. The likelihood function for our data is

$$p_1(\boldsymbol{Y}|\boldsymbol{\theta},\boldsymbol{\lambda}) = \prod_{i=1}^{N}\prod_{t=1}^{T}\frac{1}{\gamma_i+\delta_i t}\exp\left\{-\frac{[Y_{it}-(\alpha_i+\beta_i t)]^2}{2(\gamma_i+\delta_i t)}\right\}. \tag{9}$$

Perhaps the most important aspect of our Bayesian hierarchical modelling is the Gaussian spatial process. It is at this stage that we model the spatial processes of each parameter. The spatial processes of each parameter are assumed to be independent. For brevity, here we only present the spatial process of the intercept of the mean equation.

$$\begin{aligned} \boldsymbol{\alpha} &= MVGP(\boldsymbol{\mu}, \Sigma_{\boldsymbol{\alpha}}), \\ \alpha_i &= \mu_i + u_i, \\ u_i &\sim MVN(0, \Lambda), \\ \Sigma_{\boldsymbol{\alpha}} &= \psi\left(d_{ij}, \rho_{\boldsymbol{\alpha}}, \Phi_{\boldsymbol{\alpha}}\right), \end{aligned} \tag{10}$$

where $\boldsymbol{\alpha} = \alpha_1, \ldots, \alpha_N$ is the vector of intercept parameters for all counties, and is assumed to follow a multivariate Gaussian process (MVGP), $\boldsymbol{\mu}$ is the deterministic part of the Gaussian process, $\Sigma_{\boldsymbol{\alpha}}$ is the covariance matrix of the Gaussian process, $d_{ij}$ is the distance between counties *i* and *j* measured from longitude and latitude coordinates, $\rho_{\boldsymbol{\alpha}}$ is the sill parameter, and $\Phi_{\boldsymbol{\alpha}}$ is the range parameter. Note that the spatial correlation is captured in the stochastic parameters of the MVGP. There is no spatial correlation in errors like in a spatial error model.

The covariance between two counties is a function of the distance separating them and the Kriging parameters (sill and range). There are different possible specifications for this function; we use the exponential:

$$\psi(d_{ij}, \rho_{\alpha}, \Phi_{\alpha}) = \rho_{\alpha} e^{-\frac{d_{ij}}{\Phi_{\alpha}}}. \quad (11)$$

As noted before, the hyperprior is the prior for the vector of hyperparameters. We assume that each hyperprior is independent from the others.

## 3. Data

Historical corn yields from 1984 to 2015 are from the National Agricultural Statistics Service (NASS). We use county corn yields from the states of Illinois and Iowa. We use all counties, including those with an incomplete yield history. When we compare the predictive performance of our approach to that of BMA, we use all counties for our approach but only use counties with a full yield history for BMA. Illinois and Iowa are the two leading producers of corn in the country. Due to the computational needs of the algorithm used in this study, we use data from only two states. The distance between counties is measured using the longitude and latitude coordinates of the county centroids.

## 4. Procedures

Our approach estimates the mean and variance of yield densities together. Yield densities have usually been estimated in two stages. The first stage is to regress a county's historical yield against a trend. The second stage—the determination of the variance of the yield densities— uses the estimated residuals from the first stage. There are different functional forms used by researchers for the mean yield. The trend is usually deterministic. The common approaches are a

simple linear trend and a trend modeled through a spline with knots (e.g. Harri et al. 2011). In our model we use a simple linear trend. One slope suffices because we use a limited time period.

Similarly, in the crop insurance literature, different assumptions have been made about the structure of the variance of yield. Some researchers have assumed a constant coefficient of variation, with changes in the yield standard deviation proportional to the yield mean (e.g. Ker, Tolhurst, and Liu 2016; Deng, Barnett, and Vedenov 2007; Ker and Coble 2003; Miranda and Glauber 1997; Skees, Black and Barnett 1997). Other researchers have assumed homoskedasticity (Coble, Heifner, and Zuniga 2000; Mahul 1999; Miranda 1991). Harri et al. (2011) assume the relationship between yield variance and predicted yield is $\sigma_{it}^2 = \omega^2[\mathrm{E}(\hat{Y}_{it})]^{\kappa}$, so they empirically estimate the functional form of the heteroskedasticity. Our model assumes a linear trend in the variance equation. This form of heteroskedasticity is different from other forms assumed in past research.

There have been different specifications for yield densities. We use the normal density in this study. In addition to the normal, other specifications that have been used are the gamma (Gallagher 1987), the logistic (e.g., Atwood, Shaik, and Watts 2003), and the mixture of normals (e.g. Ker, Tolhurst, and Liu 2016), among others.

Ker, Tolhurst, and Liu (2016) use Bayesian Model Averaging (BMA). BMA is a weighted average of a set of models. They first estimate each county's conditional yield density using its own historical yield data. To obtain the BMA estimate for county *i*, they compute the Bayesian Information Criterion (BIC) of each estimated model using county *i*'s yield data. County *i*'s BMA estimate is the weighted average of all models, where the weights are derived from the BIC of each model given the data in county *i*. Like the BMA approach, our approach

improves estimation efficiency by using data from all counties in the estimation of each county's yield density. Unlike the BMA approach, the Bayesian approach assumes that the similarity in yield density results from proximity in space.

The computation of integrals such as the one in equation (8) generally does not have closed-form solutions. The marginal posterior distributions of the parameters $\boldsymbol{\theta}$, we use Markov Chain Monte Carlo (MCMC) simulations. Specifically, we use the Metropolis Hastings algorithm within a Gibbs sampler. In the Metropolis-Hastings algorithm, random parameter values are drawn from a candidate density and then accepted or rejected; the accepted values are included in the posterior density.

We impose hyperpriors as follows. For the hyperparameters corresponding to prior means of the parameters $\boldsymbol{\theta}$, we impose normal hyperprior distributions with parameters obtained from maximum likelihood estimates of the model that aggregates all counties for the estimation of the parameters of yield densities. Note that there is no spatial process in this maximum likelihood estimation. For the variance hyperparameters, we impose inverse gamma hyperpriors. For the sill and range, we impose hyperpriors based on an empirical variogram.

The posterior distributions obtained from the MCMC are given as samples of values from the posterior distributions. The samples are generated by a Markov Chain whose stationary distribution is the same as the posterior. We assume that the marginal distributions of the simulated values are close to the target distributions of the parameters.

The MCMC contains 15,000 iterations, with a burn-in period of 5,000. Discarding (i.e., "burning") the early iterations diminishes the influence of the starting values. To perform a

diagnosis of the convergence of the MCM simulations, we use the Gelman-Rubin test. The procedure for this test followed the outline in Gelman et al. (2014). Two sequences are simulated, each with 10,000 iterations after burn-in. Then each of the sequences is split into two sequences of length 5,000, and we end up with $m$=4 sequences of $n$=5000 iterations each. Convergence is judged based on the following scale reduction factor:

$$\hat{R} = \sqrt{\frac{n-1}{n} + \frac{1}{n}\frac{B}{W}} \tag{11}$$

where $B$ and $W$ are the between-and within-sequence variances. The MCM is said to converge if the value of $\hat{R}$ is close to 1.

We compare the predictive performance of our approach to that of the BMA model (Ker, Tolhurst, and Liu 2016), using two prediction measures: the root mean squared error (RMSE) and Diebold and Gunther's (1998) probability integral transform test. We first estimate posterior densities using 1984-2010 data and predict 2011 yield means and yield densities. We then estimate posterior densities using 1984-2011 data to predict 2012 yields means and yield densities. The same process is repeated until we use 1984-2014 data to predict 2015 yields. All these predictions are compared to actual observed yields, and the RMSE is computed.

The probability integral transform (PIT) is used to evaluate density forecasts. The PIT is the cumulative density function corresponding to the predicted yield density, evaluated at the observed yield. Deiebold and Gunther (1998) show that under assumptions of a nonzero Jacobian with continuous partial derivatives, the sequence of PITs is Uniform (0, 1). This test is a goodness-of-fit test.

## 5. Results

Parameters of corn yield densities are estimated using Bayesian Kriging. The MCMC provides four parameters for each of the 201 counties in Illinois and Iowa. Because there are many counties in this study, we present the posterior means using maps (Figures 1-6). All results presented in these maps are from the estimation that uses 1984-2014 data. Figures 1-4 contain results from the Bayesian Kriging approach, while figures 5 and 6 have results from the BMA model. The clustering of values on the maps follows the Jenks natural breaks classification method.

Figure 1 is a map of the intercept parameter of the mean of yield density. The posterior mean of the intercept is highest in central Illinois and in northern Iowa. The map in Figure 2 shows the posterior mean of the trend parameter of the mean of yield density. The map shows that yields have grown faster in northern counties for both states. An explanation is perhaps that due to the cooler climate in the north, improvements in production technology have resulted in larger yield increases in the north. Tannura et al. (2008) argue that increased yields are a result of earlier planting and this earlier planting may be more important in northern areas.

Figure 3 shows the posterior mean of the intercept of yield density, and figure 4 shows the posterior mean of the trend of the variance of yield density. As seen in these figures, the variance is higher and has increased faster in southern counties than in northern counties. This result is interesting since the variance did not increase faster in the counties where the mean increased faster.

Figure 5 shows the posterior mean of the intercept of the mean of yield density, computed using BMA. Figure 6 is a map of the posterior mean of the trend of the mean of yield density. In contrasting figures 5 and 6 with figures 1 and 2, respectively, two key differences are observed. First, the blanks in figures 5 and 6 reflect the inability of BMA to estimate yield densities for counties without a complete yield history. Second, the transition from clusters with higher values to clusters with lower values is smoother in figures 1 and 2.

To test the convergence of the MCMC chains, we use the Gelman-Rubin diagnostic. In this test, values close to 1 are desired for the scale reduction factor. The test was mostly satisfied. The highest $\hat{R}$ value obtained was 1.29. Sixty percent of the parameters that were sampled in the MCMC had an $\hat{R}$ value less than 1.2. Figure 7 shows plots of the 10,000 iterations of the 4 parameters of one county. The plots show good mixing of the MCM chains for this county. However, due to the high number of parameters to sample, not all chains had good convergence, and significantly increasing the number of iterations would improve convergence.

We now compare the predictive performance of our approach to that of BMA. Although we make this comparison, we note that the BMA model only uses counties with a full yield history, while our model uses all 201 counties. Because Bayesian Kriging is an interpolation technique, it has the advantage of handling missing values. A major similar aspect between the BMA approach and our approach is that data from all counties included in the model are used in the estimation of each county's yield density.

The RMSE for our Bayesian Kriging model is 31.03 bushels/acre, while the RMSE of BMA is 31.33 bushels/acre. These values are similar, but our model is the only one that has predictions for all counties and provides the spatial smoothing desired by the Risk Management

Agency. Both RMSE values are high, because both models' predictions of 2012 corn yields were poor. Corn yields were atypically low in 2012.We compute separate RMSE's to evaluate how the models performed in each individual year. Our model outperforms BMA in 2009, 2010, 2011, and 2013, while BMA outperforms our model in 2012, 2014, and 2015. These results are summarized in table 1.

Histograms of the probability integral transforms obtained by forecasting yield densities using Bayesian Kriging and BMA are presented in figures 8 and 9, respectively. For accurate density forecasts, the PIT is distributed Uniform (0, 1). Using a formal PIT test, the null hypothesis that the actual data comes from the density estimated by the Bayesian Kriging approach is rejected, while the null hypothesis that the data comes from the distribution estimated using BMA is not rejected. This result is explained by the fact that our specified distribution is normal, while BMA uses a mixture of normals. A better comparison of the two methods of estimation will be included in the next step of this research; in that comparison, BMA will be respecified with the assumption of normality of yields or we will redo the Bayesian Kriging with a mixture of normals.

## 6. Conclusion

Forecasts of yield densities are important for crop insurance rating. The RMA needs accurate forecasts in order to make actuarially fair premiums for insurance policies. Producers and agricultural businesses also need accurate yield forecasts for land valuation and for their own planning. We introduce a new approach for the estimation of yield densities. We use Bayesian Kriging for spatial smoothing of all parameters of counties' yield densities. The intercept and trend parameter are spatially smoothed in both the mean equation and the variance equation. Our

Bayesian Kriging approach assumes that the stochastic parameters are spatially correlated. The spatial autocorrelation in the data is captured by the spatial correlation in parameters rather than spatial correlation in error. A county's estimated yield density is derived not only from its own historical yield observations, but also from other counties.

This paper extends Bayesian Kriging to the general heteroskedastic error regression model. The approach can handle nonnormality, but that increases computational time considerably. This paper has two main contributions to the crop insurance literature. The first is the ability to spatially smooth multiple parameters of yield densities, including trend parameters. The second is the ability to use unbalanced data in the estimation. There are counties whose yield is not recorded in all years; previous approaches that only use balanced data cannot be used directly in forecasting yield densities for such counties. A drawback of our approach is that it is computationally expensive. However, with increases in computers' performance, this problem is expected to lessen. Also, we are using the R software package, which is notorious for being computationally burdensome. In addition, given the aforementioned advantages, the computational time should not be a deterrent.

References

Anselin, L., 1988. The Scope of Spatial Econometrics. In *Spatial Econometrics: Methods and Models* (pp. 7-15). Dordrecht: Springer.

Coble, K.H., Heifner, R.G. and Zuniga, M., 2000. Implications of Crop Yield and Revenue Insurance for Producer Hedging. *Journal of Agricultural and Resource Economics* 25(2):432-452.

Deng, X., Barnett, B.J. and Vedenov, D.V., 2007. Is There a Viable Market for Area-Based Crop Insurance? *American Journal of Agricultural Economics 89*(2):508-519.

Diebold, F.X., Gunther, T.A. and Tay, A.S., 1998. Evaluating Density Forecasts with Applications to Financial Risk Management. *International Economic Review*, vol. 39, no. 4. In *Symposium on Forecasting and Empirical Methods in Macroeconomics and Finance* (p. 863).

Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A. and Rubin, D.B., 2014. *Bayesian Data Analysis* (Vol. 2). Boca Raton, FL: CRC press.

Harri, A., Coble, K.H., Ker, A.P. and Goodwin, B.J., 2011. Relaxing Heteroscedasticity Assumptions in Area-Yield Crop Insurance Rating. *American Journal of Agricultural Economics 93*(3):707-717.

Ker, A.P. and Coble, K., 2003. Modeling Conditional Yield Densities. *American Journal of Agricultural Economics 85*(2):291-304.

Ker, A.P., Tolhurst, T.N. and Liu, Y., 2016. Bayesian Estimation of Possibly Similar Yield Densities: Implications for Rating Crop Insurance Contracts. *American Journal of Agricultural Economics 98*(2):360-382.

Mahul, O., 1999. Optimum Area Yield Crop Insurance. *American Journal of Agricultural Economics 81*(1):75-82.

Miranda, M.J., 1991. Area-Yield Crop Insurance Reconsidered. *American Journal of Agricultural Economics 73*(2):233-242.

Miranda, M.J. and Glauber, J.W., 1997. Systemic Risk, Reinsurance, and the Failure of Crop Insurance Markets. *American Journal of Agricultural Economics 79*(1):206-215.

Ozaki, V.A., Ghosh, S.K., Goodwin, B.K. and Shirota, R., 2008. Spatio-Temporal Modeling of Agricultural Yield Data with an Application to Pricing Crop Insurance Contracts. *American Journal of Agricultural Economics 90*(4):951-961.

Park, E., Brorsen, B.W. and Harri, A., 2016. Using Bayesian Spatial Smoothing and Extreme Value Theory to Develop Area-Yield Crop Insurance Rating. Selected paper, 2016 Annual Meeting, Boston, MA, USA.

Park, E., Brorsen, B.W. and Harri, A., 2018. Using Bayesian Kriging for Spatial Smoothing in Crop Insurance Rating. *American Journal of Agricultural Economics* forthcoming.

Reich, B.J., 2012. Spatiotemporal Quantile Regression for Detecting Distributional Changes in Environmental Processes. *Journal of the Royal Statistical Society: Series C (Applied Statistics) 61(4): 535-553.*

Skees, J.R., Black, J.R. and Barnett, B.J., 1997. Designing and Rating an Area Yield Crop Insurance Contract. *American Journal of Agricultural Economics 79*(2):430-438.

Tannura, M.A., Irwin, S.H., and Good, D.L. 2008. Weather, Technology, and Corn and Soybean Yields in the U.S. Corn Belt. Unpublished working paper. University of Illinois at Urbana-Champaign. Available at SSRN: https://ssrn.com/abstract=1147803.

**Table 1. Root Mean Squared Error (bu/acre) Comparison between BMA and Bayesian Kriging Models**

| | BMA | Bayesian Kriging |
|---|---|---|
| 2009 | 13.71 | 13.14 |
| 2010 | 28.50 | 23.92 |
| 2011 | 18.29 | 16.80 |
| 2012 | 63.33 | 65.32 |
| 2013 | 21.65 | 19.12 |
| 2014 | 28.53 | 29.49 |
| 2015 | 15.53 | 16.71 |
| 2009-2015 | 31.33 | 31.03 |

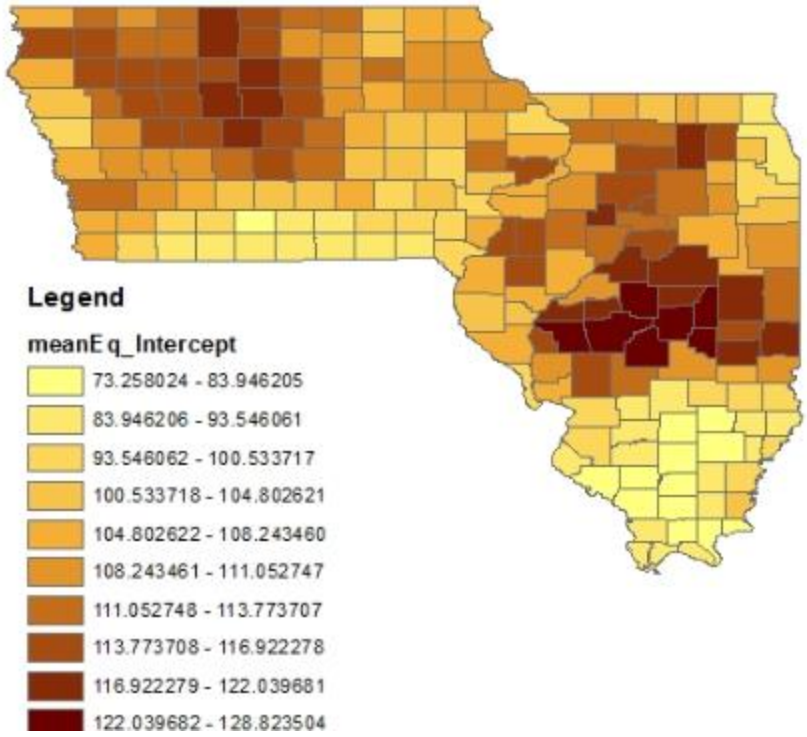


**Figure 1. Intercept of Mean Equation for County Yield Density (Bayesian Kriging)**

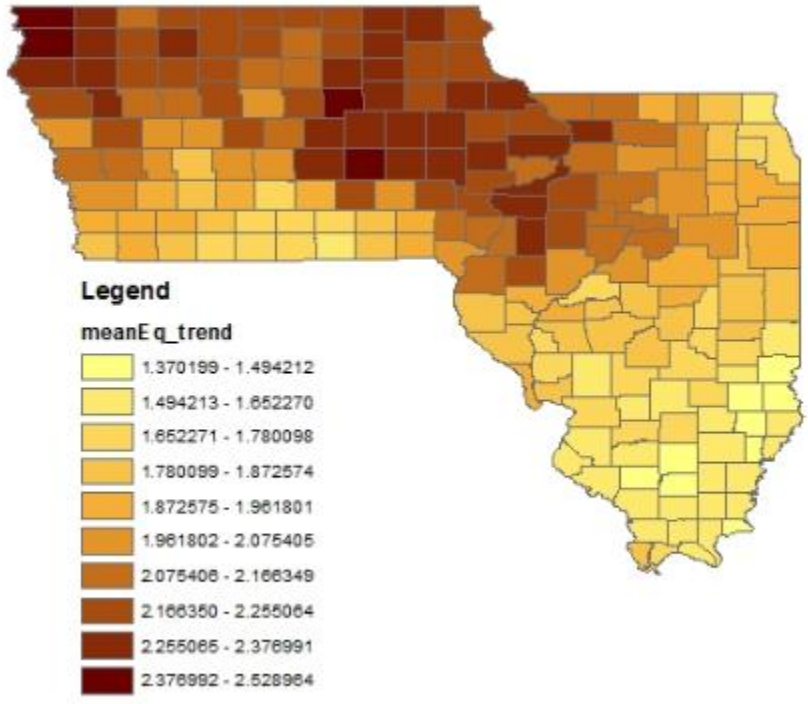


**Figure 2. Trend of Mean Equation for County Yield Density (Bayesian Kriging)**

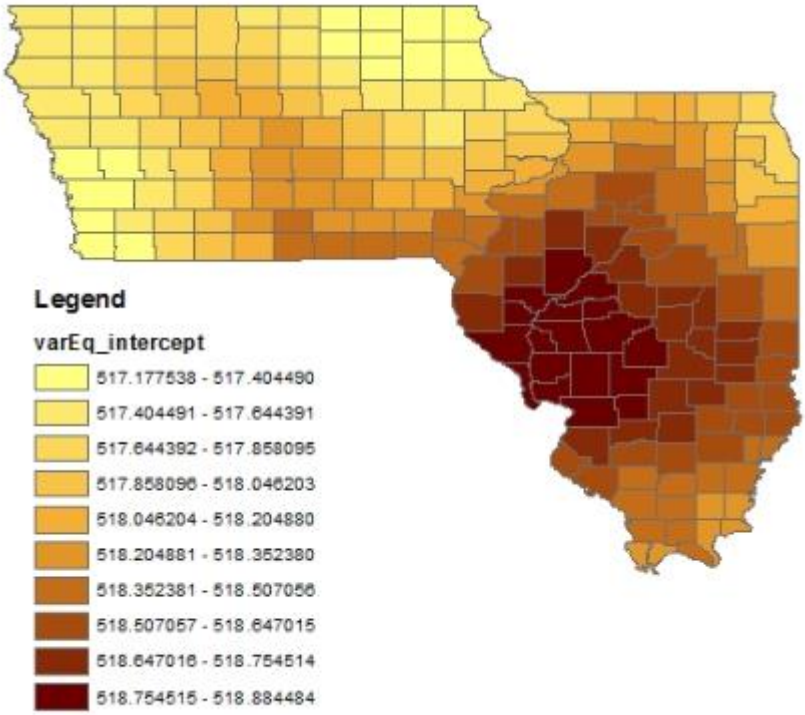


**Figure 3. Intercept of Variance Equation for County Yield Density (Bayesian Kriging)**

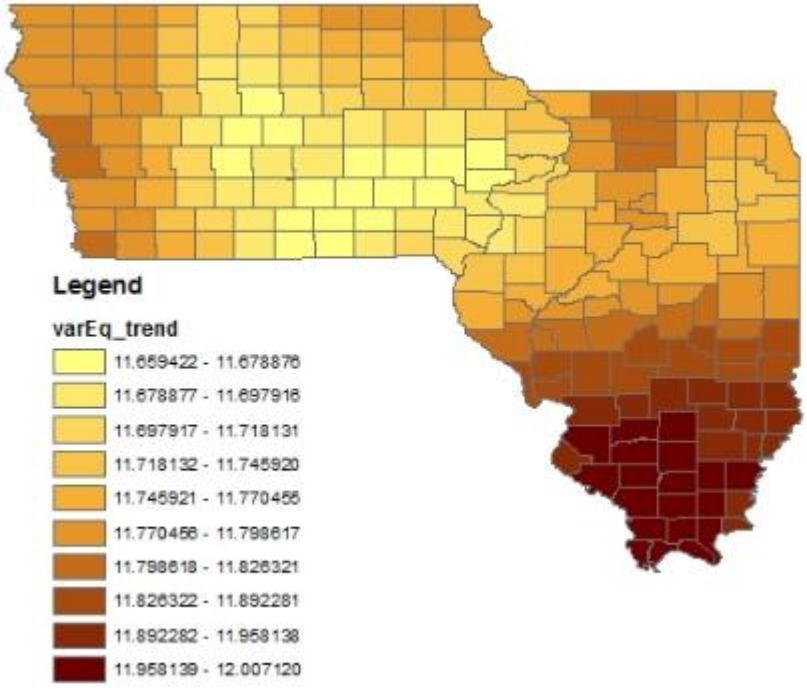


**Figure 4. Trend of Variance Equation for County Yield Density**

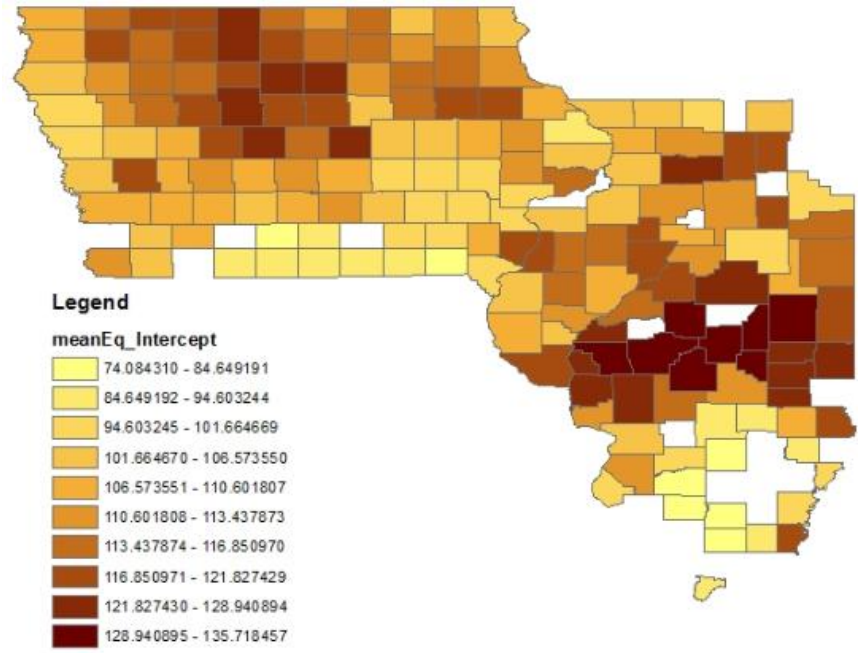


**Figure 5. Intercept of Mean Equation for County Yield Density (BMA)**

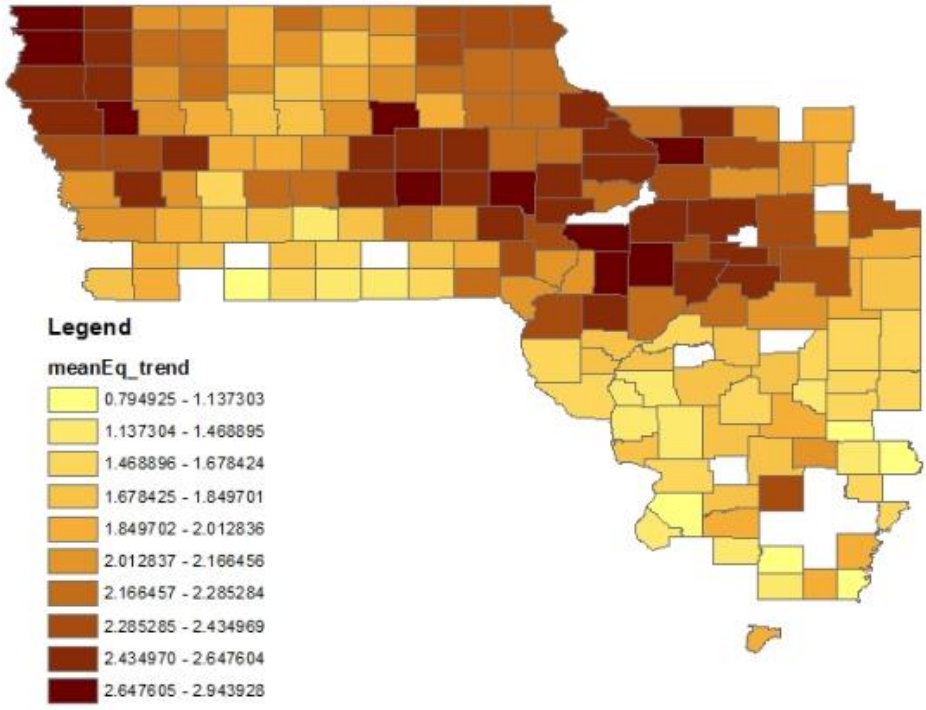


**Figure 6. Trend of Mean Equation for County Yield Density (BMA)**

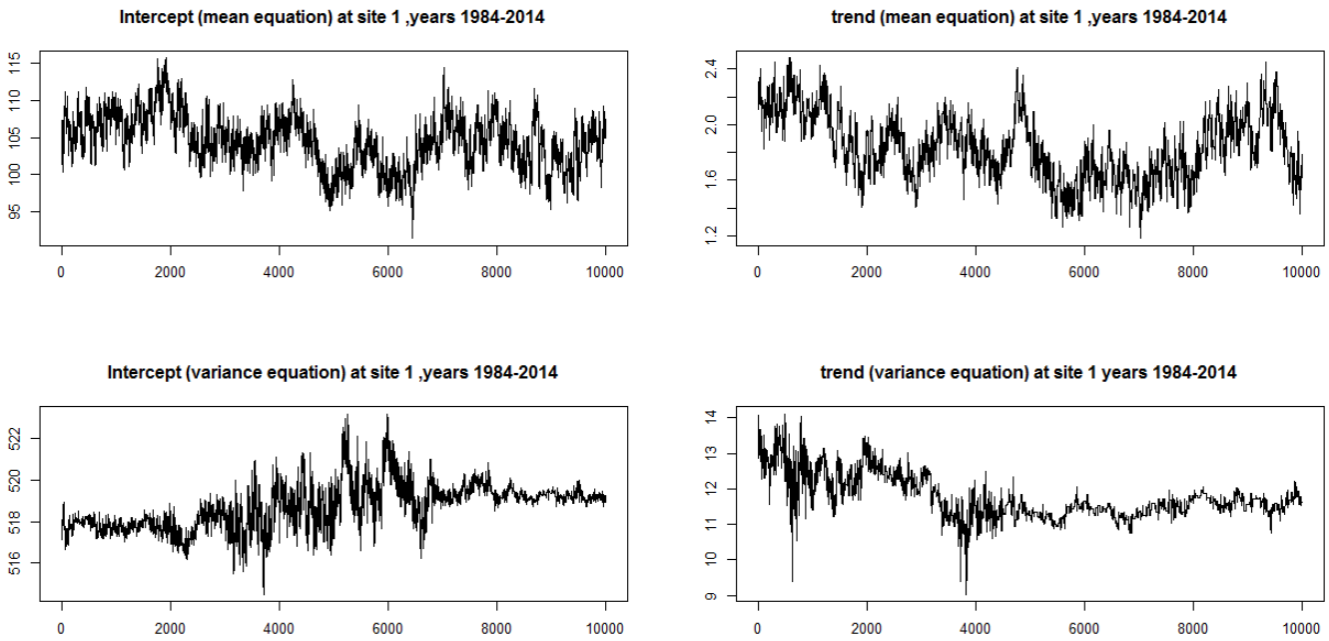


**Figure 7. Iterations of parameters for Adams County, Illinois**

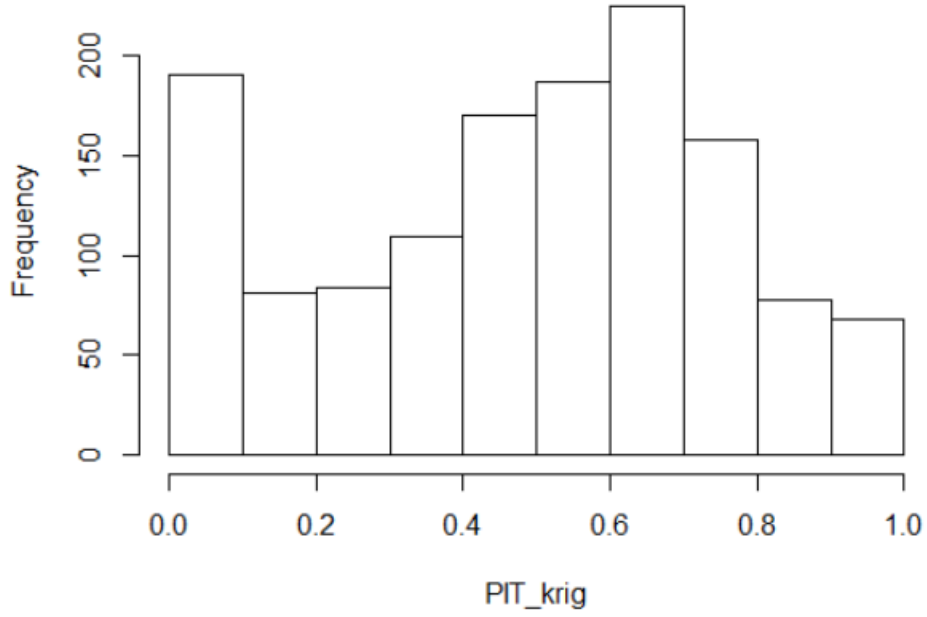


**Figure 8. Histogram of the Probability Integral Transform using Bayesian Kriging**

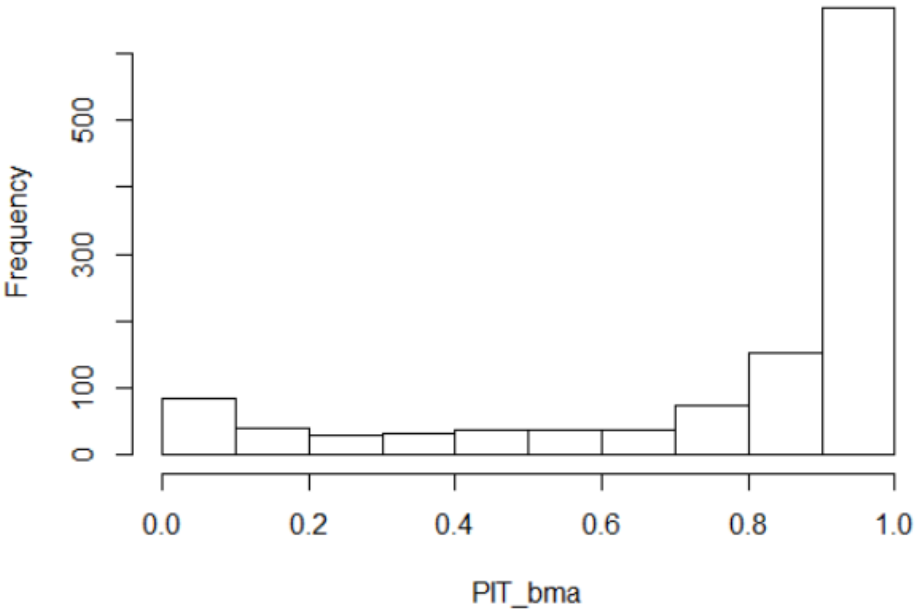


**Figure 9. Histogram of the Probability Integral Transform using BMA**