Using a Spatially Explicit Analysis Model to Evaluate Spatial Variation of Corn Yield

K.C. Stone[1] and E.J. Sadler[2]

[1]USDA-ARS Coastal Plains Soil, Water, and Plant Research Center, Florence, SC, and [2]USDA-ARS Cropping Systems and Water Quality Research Unit, Columbia, MO 6521. **Corresponding author:** Kenneth C. Stone; e-mail: ken.stone@ars.usda.gov.

Abstract: *Spatial irrigation of agricultural crops using site-specific variable-rate irrigation (VRI) systems is beginning to have wide-spread acceptance. However, optimizing the management of these VRI systems to conserve natural resources and increase profitability requires an understanding of the spatial crop responses. In this research, we utilize a recently developed spatially explicit analysis model to analyze spatial corn yield data. The specific objectives of this research are 1) to calculate a suite of estimates needed for the types of analyses mentioned above and to provide credible intervals around these estimates and 2) to examine whether the conclusions from this rigorous re-analysis are different from the prior analysis and if the results force any modifications to the conclusions obtained with the prior analyses. The model simultaneously accounted for spatial correlation as well as relationships within the treatments and has the ability to contribute information to nearby neighbors. The model-based yield estimates were in excellent agreement with the observed spatial corn yields and were able to more accurately estimate the high and low yields. After calculating estimates of yield, we then calculated estimates of other response variables such as rainfed yield, maximum yield, and irrigation at maximum yield. These estimated response variables were then compared with previous results from a classical statistical analysis. Our conclusions supported the original analysis in identifying significant spatial differences in crop responses across and within soil map units. The major improvement in the 2014 re-analysis is that the model explicitly considered the spatial dependence in calculation of the estimated yields and other variables and, thus, should provide improved estimates of their impact in system design and management.*

Keywords. *Varying coefficient model; Response curves; Semiparametric regression; Site-specific agriculture.*

INTRODUCTION

With the convergence of geopositioning (GPS), geospatial (GIS), and increased computing technologies in the 1980's and 1990's, much attention has been placed on precision, or site-specific, agriculture as a way to improve economics and environmental benefits of agriculture. Many crop inputs, including fertilizer, plant population, and pest control, have been examined for potential to reduce costs or increase yield, or perhaps shift the location of use from where the input is not used efficiently to other places in the field where it is. When one considers irrigation as an input to be managed concurrently with all other inputs, it becomes a logical extension to consider if spatially-varying irrigation could be a water-conserving approach. This was demonstrated to be the case based on research in South Carolina (Sadler et al., 2005), where optimal variable-rate irrigation had potential to save from 10 to 15% in three fairly dry years (1999-2001). It was also been shown in other locations and clearly, water conserving aspects of variable rate irrigation merit further examination.

One data source that has potential to inform this question was obtained in Florence, South Carolina, at a variable-rate center pivot research facility (Camp and Sadler 1998; Camp et al. 1998, Sadler et al, 2002). A 3-year corn yield experiment was conducted to examine the simultaneous effects of irrigation and N fertilization. The experiment was intended to be analyzed by both analysis of variance (ANOVA) and geostatistical analyses. The constraints on the experimental design were discussed in detail in Sadler et al. (2002). In brief, there were four irrigation treatments: 0%, 50%, 100%, and 150% of irrigation to replace evapotranspiration (ET) applied simultaneously in all irrigated plots when the mean soil water tension in the 100% plots in 4 (1999) or 6 (2000-2001) soil map units was less than or equal to -30 kPa at the 0.3-m depth. There were two N treatments: 134 kg/ha, which is the extension recommendation for rainfed corn in SC, and 225 kg/ha, for irrigated corn. Nitrogen fertilizer was a blended dry granular pre-plant application common to all plots within a year, followed by treatments applied as urea-ammonium nitrate with sulfur (UAN 24S) injected in a nominal 13-, 11-, and 16-mm irrigation in all plots during the three years. Where sufficient area existed within a map unit, one or more randomized complete blocks (RCB) of these eight treatments were imposed. Where insufficient area existed, a randomized incomplete block (RICB) was used. In total, there were 39 RCBs and 19 RICBs, for a total of 396 plots. Harvest was done with a plot combine from 2 rows 6.1 m in length near the center of the plots. Additional details can be obtained in Sadler et al. (2002).

In the initial ANOVA analysis, the irrigation main effect was, as expected, highly significant in both linear and quadratic contrasts over the irrigation ranges (0 to 150% of full irrigation). Somewhat surprisingly, however, between the two N fertilizer rates of 134 and 225 kg/ha, the main effect of N was significant in only one of three years. The test of whether the soil variation was a significant contributor to variation in yield was significant, despite the limitations of the experimental design. Interaction effects were not consistent, either in 2-way or 3-way interactions. In short, the soil variation was important, the expected irrigation effect was obtained, and the N effect was somewhat surprising. Sadler et al. (2002) further evaluated the quadratic irrigation production functions for map unit means in that experiment, and solved for maximum yield and the irrigation to obtain it. These analyses were all done with soil map units as a class variable, and did not explicitly account for spatial variation within blocks, within soil map units, or among soil map units. The only indication of spatial variation in the production functions was that block-level functions for the most common soil map unit were provided. Further, the map-unit-mean analytical expressions for yield as a function of irrigation were obtained separately for N treatments in only 2001, the year in which the N treatment was significant. These characteristics of the analyses limited interpretation of spatial variation in the production functions themselves.

A first attempt to explicitly account for spatial variation in these data was described by Sadler et al (2002b). The goal of this analysis was not to test statistical significance, but to allow generation of maps of derived characteristics, including rainfed yield, maximum yield, and irrigation water use efficiency, plus the maps of irrigation to achieve those. All derived results were obtained from averages over N treatment, on the basis that the N treatment was not significant in two of three years. This analysis involved two steps – a separate interpolation of yields from each individual treatment to estimate the yield that would have been expected for all four irrigation treatments at each of the 396 plot centers, and a quadratic regression of the four yields on seasonal irrigation rate in mm.

Sadler et al. (2003) described the marginal N response for given irrigation levels, and explained the 2-step process for all eight treatments, producing analytical expressions of irrigation production functions for both N treatments at each of the 396 plot centers. The maps of derived surfaces qualitatively showed distinct spatial patterns in the field. The primary conclusion was that spatial patterns in marginal N response were not stable across years, and further that it was surprisingly variable, ranging from negative to positive in credible areas (i.e., multiple data points in each) of the field. These results demonstrated the utility of having mathematical expressions for irrigation production functions at many areas within a field and were, at that time, to our knowledge, the only such results in existence. Spatial patterns of maximum yield, rainfed yield, irrigation water use efficiency, and with prices, results of marginal economic benefits of irrigation and N can be provided. However, the procedure suffers a number of limitations, including dependence on only spatial variation for step one and on only irrigation for step two (N is accounted for by performing regressions separately for the two N treatments). There is also no good means to provide estimates of uncertainty (variation) around the yield estimates themselves, nor of significance in the estimated yield.

Spatially Explicit Analysis

Despite the benefits of the analyses performed, these data required a spatially explicit analysis. The spatially explicit analysis was achieved using a Bayesian mixed model formulation of bivariate penalized smoothing splines (Holan et al., 2008). This model simultaneously accounted for spatial correlation as well as relationships within the treatments – in other words, considered X, Y, Irrigation, and N, all with the ability to contribute information to nearby neighbors. The outcomes of this analysis were required to include yield estimates, analytical expressions for irrigation and N production functions, and estimates of the uncertainty in yield estimates, coefficients of the analytical expressions, and in the derived variables (e.g., maximum, rainfed, and economic yield), the irrigation required to provide them, irrigation water use efficiency or water productivity. Additionally, the model provided credible intervals around the estimated variablesThis paper employs the method of Holan et al. (2008) to re-analyze the experimental data using a spatially explicit analysis. The specific objectives of this research are to 1) to estimate the suite of variables needed for the types of analyses mentioned above and to provide credible intervals around those estimated variables, and 2) to examine whether the results of this rigorous re-analysis differed from the prior analysis and whether the results force any modifications to the conclusions obtained with the prior analyses.

METHODS

The observed spatial corn yield data collected from Sadler et al. (2002) (2002 analysis) was used to estimate a suite of variables using the recently developed spatially explicit analysis model (2014 analysis). The 2014 analysis, described above inputs the spatial coordinates, imposed irrigation and nitrogen (N) treatments, and observed yield data to estimate the spatial yields. Additionally, the 2014 analysis provides credible intervals (posterior distributions) for the individual estimated yields and uses a spatially-treatment-varying coefficient model to fit the observed yield data.

For a general comparison of the performance of the 2002 and 2014 approaches, the results of both sets of yield estimates were compared to the observed yields using linear regression (SAS, Cary, NC). For a comparison of derived analysis variables (maximum, rainfed, and economic yield, the irrigation required to provide them, irrigation water use efficiency or water productivity), standard summary statistics (means, standard deviations, minimum, and maximum values) were calculated (calculations were carried out using SAS), and the differences between the estimated variables were calculated (i.e., 2002 versus 2014). The point estimates from both analyses were also compared using linear regression (i.e., perfect agreement would result in an intercept=0, slope=1, and R^2=1.0). Contour maps of the calculated variables were generated using the Surfer software (Golden Software, Inc., Golden, CO) using default interpolation parameters (point Kriging, slope=1, aniso=1.0).

RESULTS

Estimated Yield: The observed yields were fit using the spatially explicit analysis (Holan et al., 2008). The 1999-2001 estimated corn yields were calculated using the 2014 analysis (figure1). The yield estimates were then plotted against the observed yields for comparison. In 1999, the 2014 analysis estimated the yields very well in terms of R^2. The estimated yield slope was 0.82 Mg/mm (R^2=0.83, rmse=0.82, Figure 2). Additionally, the upper and lower credible intervals for the 2014 estimated yields are plotted along with the regression (Figure 2). The 1999 growing season was generally considered a drought year; it required the greatest irrigation depths and had the widest variation in corn yields. The 2002 yield estimates were also plotted for comparison (Figure 2; slope=0.72 Mg/ha, R^2=0.84, rmse=0.68). The 2014 yield estimate had a slope closer to 1.0, indicating it did a better job of estimating the high and low yields than the 2002 yield estimates.

In 2000, the estimated yield slope was 0.78 Mg/ha and had an R^2 of 0.80. The 2000 slope was lower than in 1999 and may be attributed to the lower irrigation amounts applied. Again, the 2002 estimated slope was lower than that found for the 2014 analysis. The 2001 season had the lowest irrigation applications of the 3-yr study, indicating a better weather year than either of the first two, and correspondingly lower observed variation in yield. The 2014 estimated yield slope was 0.64 Mg/ha (R^2 = 0.66, rmse=0.67) while the 2002 analysis slope was much lower, 0.39 Mg/ha (R^2=0.77, rmse=0.32). The overall estimated yield fit over the three years decreased from 1999-2001 due to the decreasing irrigation depths required and the corresponding decreased variation in corn yields. The 2001 growing season was considered a more typical rainfall year than the two earlier ones. In each year, the slopes of the two analyses were

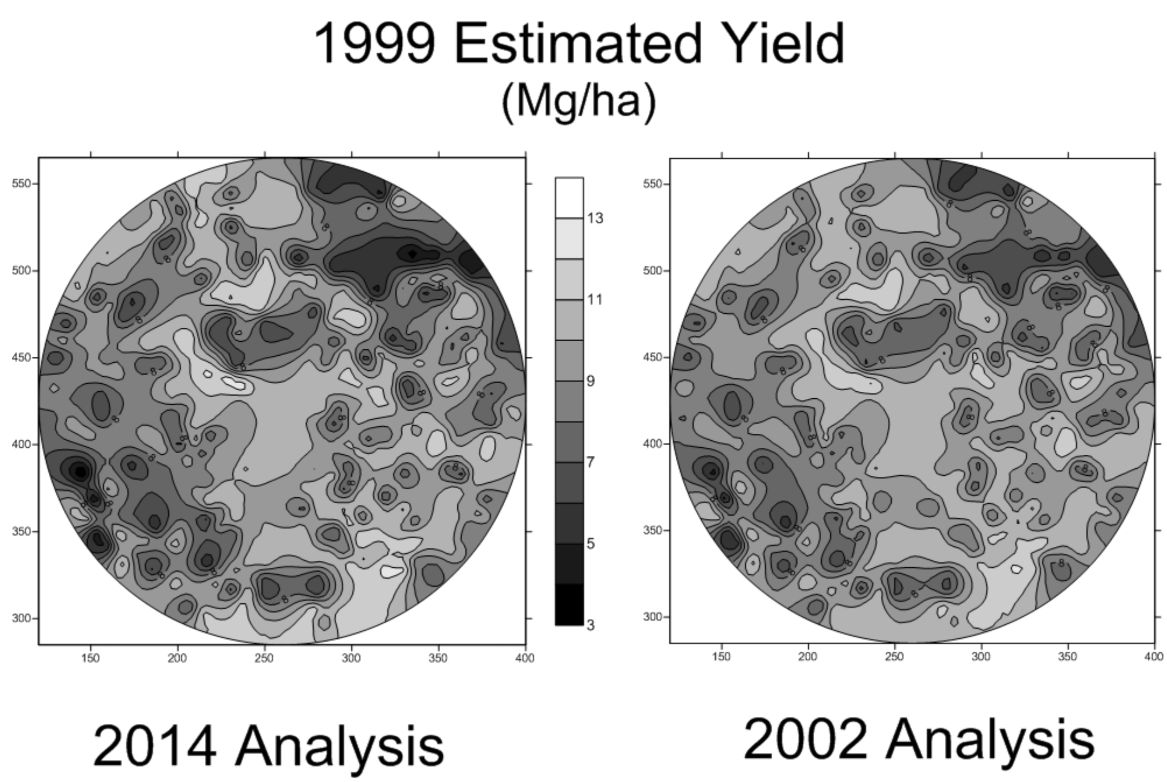

Figure 1. Estimated yield maps for the 2014 and 2002 analysis for 1999.

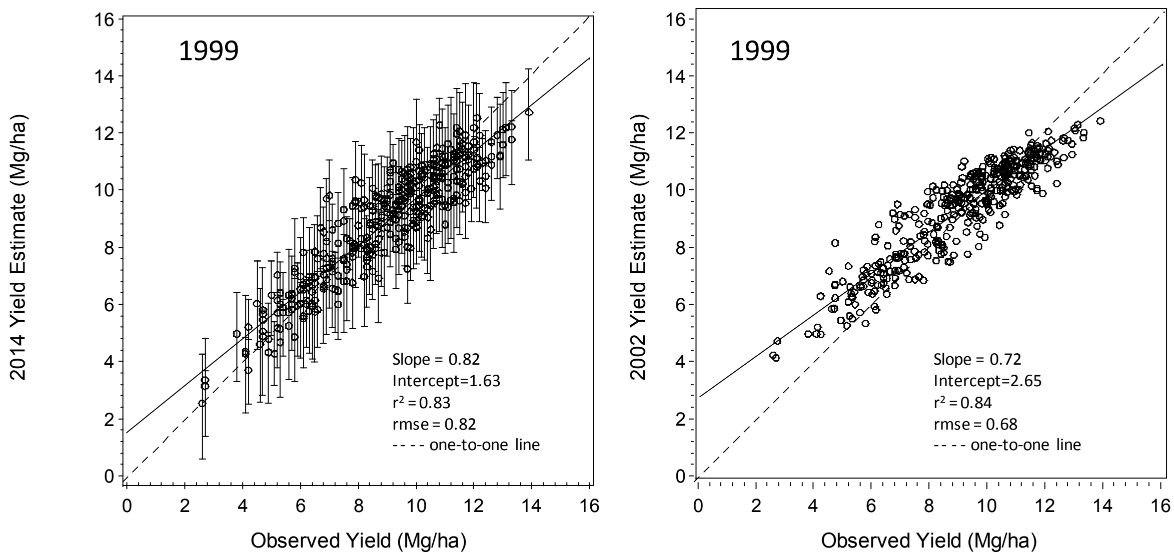

Figure 2. Regression of the 1999 estimated yields versus observed yields for the 2014 and 2002 analysis. The 2014 analysis provides the 95% credible intervals for the estimates.

significantly different. It appears that over the 3-yr study period, the 2014 analysis was able to estimate the corn yields better than the 2002 analysis. This would be a result of the two approaches: the 2002 analysis used blocks of measured yield to calculate yield response curves whereas the 2014 analysis approach estimated yield response curves using the entire sample population, taking into account the spatial dependence.

Comparison of 2014 Predicted Variables to 2002 Estimates

In our analysis, there are several quantities of interest that can all be obtained directly as output from the spatially-treatment varying coefficient model or as deterministic transformations of this output. The items of interest in this analysis are: rainfed yield (i.e., yield when irrigation equals 0), maximum yield, and irrigation that produced maximum yield. Other variables could be estimated for additional analysis but due to space constraints, only those identified above will be discussed.

Rainfed yield: The rainfed yield estimates, particularly in drought years, can provide irrigation designers a good initial estimate of the potential areas of a field where irrigation may provide the most benefit. The 1999 and 2000 corn growing seasons were generally considered drought years. In these two years and for each estimation method, estimated rainfed yields were similar (Figures 3 and 4). The 1999 mean rainfed yields for the 2014 and 2002 estimation methods were 6.75 and 6.44 Mg/ha, respectively, and for 2000 were 5.7 and 5.3 Mg/ha, respectively. In 1999, the 2014 analysis estimated rainfed yields ranging from 2.6 to 10.4 Mg/ha, and the 2002 analysis rainfed yields ranged from 3.7 to 8.2 Mg/ha. The larger ranges between the minimum and maximum rainfed yields for the two estimation methods is due to the 2002 analysis using only points that were in the same irrigation treatment which reduced the number of points used in estimating the response surface. The 2014 analysis utilized the entire data set in predicting the estimated rainfed yields and retained the influence of the extreme values. In both the 1999 and 2000 contour plots (Figures 3 and 4), there appears to be a consistent area from

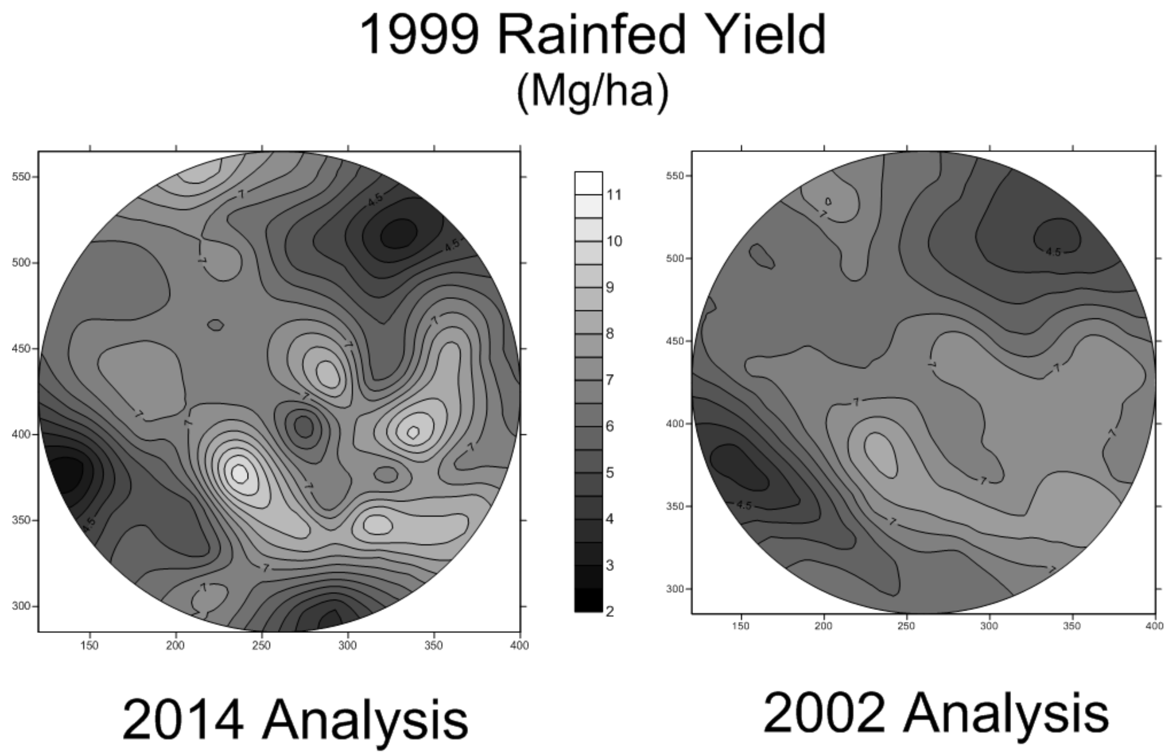

Figure 3. Rainfed estimated yield maps for the 2014 and 2002 analysis for 1999.

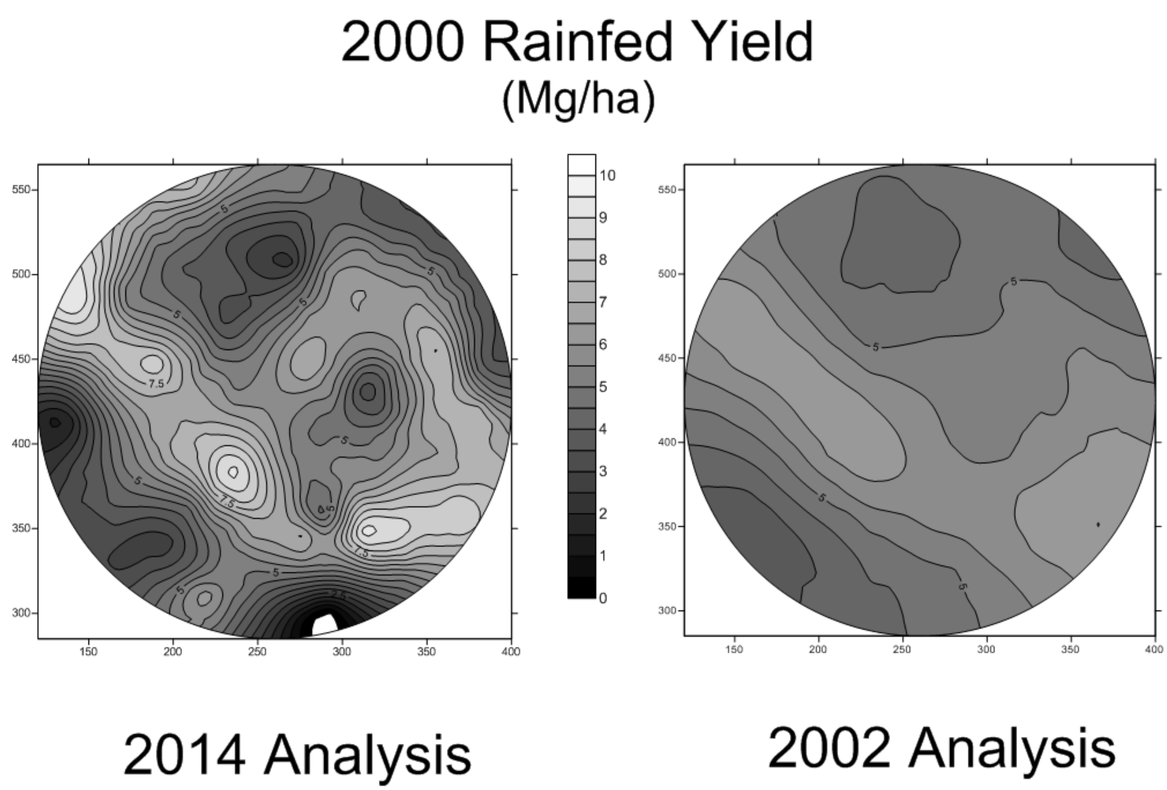

Figure 4. Rainfed estimated yield maps for the 2014 and 2002 analysis for 2000.

upper left to lower right that has relative higher rainfed yields indicating that this area would be the most productive region of the field under rainfed condition. The 2002 growing season was a more typical rainfall year and there were less defined regions within the field (data not shown).

Maximum Yield: The maximum yield estimates provide the potential spatial yields achievable under ideal conditions. The 2001 maximum calculated yields were higher than the 1999 and 2000 maximum yields (Figure 5). For the 2002 analysis, the 1999 to 2001 maximum yields ranged from 8.7 to 12.2, 8.7 to 11.6, and 10.6 to 12.8 Mg/ha, respectively with mean maximum yields 10.7, 10.4, and 11.7 Mg/ha, respectively. The 2014 analysis 1999 to 2001 maximum yields ranged from 5.8 to 12.7, 6.2 to 13.3, 9.2 to 14.4 Mg/ha, respectively, with mean maximum yields of 10.9, 10.6, and 12.0 Mg/ha, respectively.

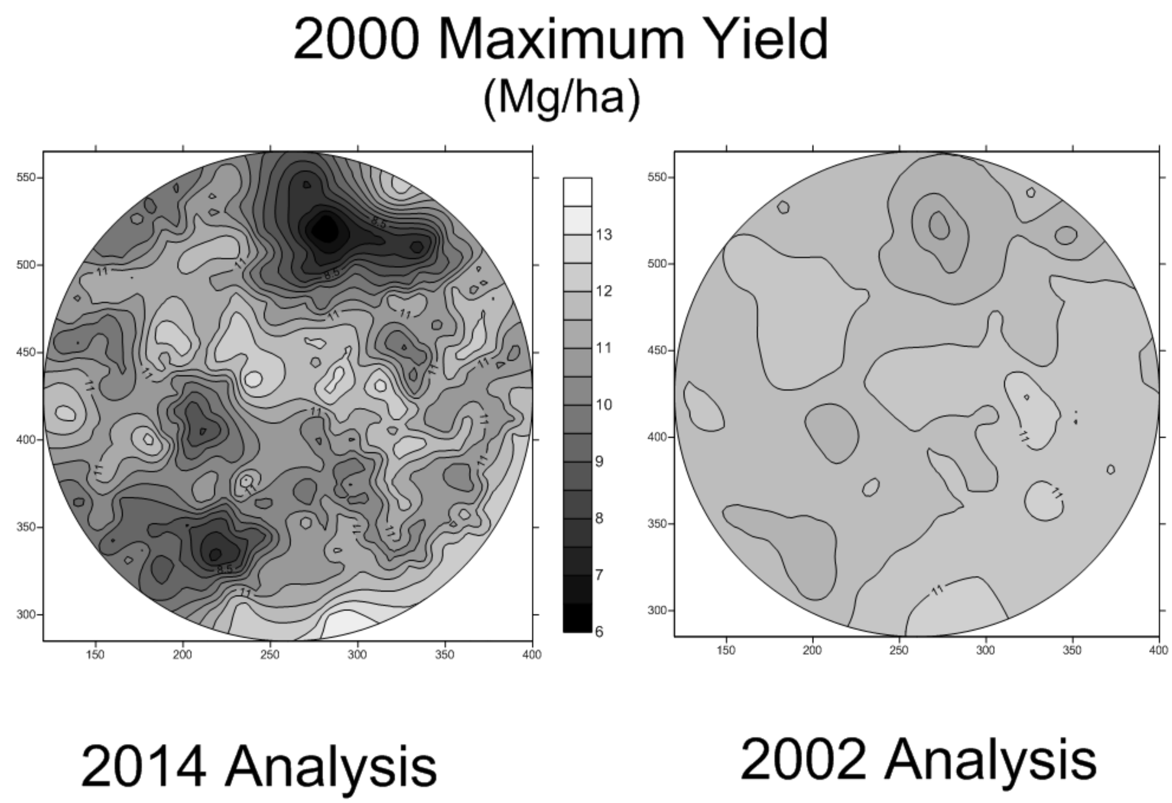

Figure 5. Maximum yield estimate yield maps for the 2014 and 2002 analysis for 2000.

Irrigation at Maximum Yield: The estimation of irrigation required for maximum yields can provide designers the appropriate design parameters for calculating maximum water application rates and irrigation system design flow rates. The irrigation depth corresponding to the calculated maximum yield using the 2014 analysis ranged from 186 mm in 2001 to 282 mm in 1999 (Figure 6) compared to the 2002 analysis which ranged from 204 in 2001 to 286 mm in 1999. The contour maps created for the irrigation depth at maximum irrigation illustrate the differences between the two estimation methods. The 2002 analysis calculated response resulted in areas of the field with little detail. The 2014 analysis utilized the entire dataset and was able to fill areas that were very flat in the 2002 analysis. In comparing the two methods using regression analysis, the slopes were 0.62, 0.39, and 0.23, for 1999-2001, respectively.

However, from 1999 to 2001, the irrigation at maximum yield regression R^2 values were 0.43, 0.19, and 0.05, respectively, indicating poor correlation between the two analyses. Clearly, the results obtained from the 2014 analysis provide more information for irrigation system design on this and other fields with similarly large variation in the soil resource.

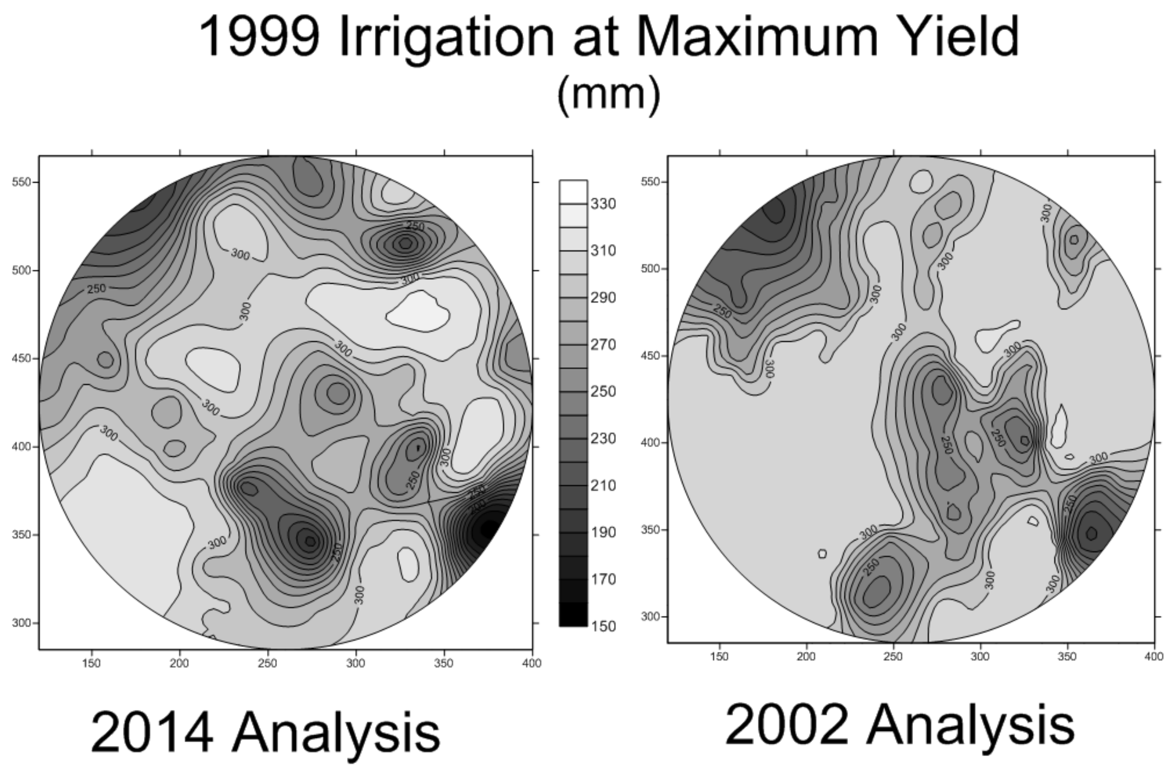

Figure 6. Irrigation at maximum yield estimate maps for the 2014 and 2002 analysis for 1999.

SUMMARY AND CONCLUSIONS

The recently developed spatially explicit analysis (2014 analysis) was used to re-analyze spatial corn yield data. The 2014 method fitted estimated yields in excellent agreement with the observed spatial corn yields. The 2014 analysis preserved more of the spatial variation in the predicted yields and response variables. Overall, the 2014 analysis predicted mean estimated yields for each response variable in relatively close agreement to the 2002 analyses and additionally provided uncertainty estimates.

Our second objective asked if the 2014 analysis would change the conclusion reached in the 2002 analysis. The 2002 analysis concluded that 1) significant differences existed in the response of corn to irrigation, both across soil map units and within soil map units; 2) differences between soil map units existed at magnitudes that would likely be important in irrigation system design and management; and 3) irrigation system managers and designers should consider the effects of unexpectedly large spatial variation in crop response.

In our re-analysis of these data, we confirm their conclusions. However, unlike the 2002 analysis, the 2014 analysis specifically accounts for spatial dependence and provides measures of uncertainty, and is therefore more rigorous and intellectually satisfactory. The 2014 analysis model coefficients are spatially varying resulting in model estimates and credible intervals obtained using all of the observations simultaneously, adding confidence to the results. In all, the 2014 analysis can provide additional insights into spatial responses of crops to irrigation and that it could be used to provide irrigation managers and designers with tools needed to make critical water management decisions.

ACKNOWLEDGEMENT

The authors thank Scott Holan, Associate Professor, Department of Statistics, University of Missouri, for his assistance with the Bayesian re-analysis and for comments on earlier versions of this manuscript.

REFERENCES

Camp, C. R., and Sadler, E. J. 1998. Site-specific crop management with a center pivot. J. of Soil & Water Cons. 53 (4):312-314.

Camp, C. R., Sadler, E. J., Evans, D. E., Usrey, L. J., and Omary, M. 1998. Modified center pivot system for precision management of water and nutrients. Applied Engineering in Agriculture, 14 (1):23-31.

Holan, S., Wang, S., Arab, A., Sadler, E. J. and Stone, K. C. 2008. Semiparametric geographically weighted response curves with application to site-specific agriculture. Journal of Agricultural, Biological and Environmental Statistics, 13(4):424–439

Sadler, E. J., Camp, C. R., Evans, D. E., and Millen, J. A. 2002. Spatial variation of corn response to irrigation. Transactions of the ASAE, 45, 1869–1881.

Sadler, E. J., Camp, C. R., Evans, D. E., and Millen, J. A. 2002. Spatial analysis of corn response to irrigation. In Robert, P. C., Rust, R. H., and Larson, W. E. (eds.) Prec. Agric.: Proc. of the 6th Int'l Conf. ASA/CSSA/SSSA, Madison, WI. (CDROM)

Sadler, E. J., Camp, C. R., Evans, D. E., and Millen, J. A. 2003. Spatial variation of N fertilizer response in the Southeastern USA Coastal Plain. Pg 603-608. In Stafford, J. V., and Werner, A. Proc. 4th European Conf. on Prec. Agric Wageningen Academic Press, Wageningen, The Netherlands. (CDROM)

Sadler, E. J., Evans, R G., Stone, K.C., and Camp, C.R. 2005. Opportunities for conservation with precision irrigation. Journal of Soil and Water Conservation, 60(6):371-379.