Site-specific treatment responses in on-farm precision experimentation

R. G. Trevisan[1], D.S. Bullock[2], N. F. Martin[1]

[1]University of Illinois at Urbana Champaign (UIUC), Crop Sciences, W201 Turner Hall, 1102 S. Goodwin Ave, Urbana, IL 61801, USA, [2]UIUC, Agricultural, and Consumer Economics, 339 Mumford Hall, 1301 W. Gregory, Urbana, IL 61801, USA; rodrigo7@illinois.edu

Abstract

There is a large body of research on determining the impact of field variability of soil on crop yields. In contrast, site-specific information about crop responses to agronomic treatments is less frequent. On-Farm Precision Experimentation (OFPE) brings important information to understand the spatial variation of crop response to agronomic practices and thus to improve agronomic decisions. The objective of this work was to investigate the spatial variability of corn yield responses to nitrogen and seed rates using OFPE in four fields in the US Midwest. Geographically weighted regression was applied to generate local regression coefficients, which were used to delineate response zones in each field. The results showed the existence of great potential to adjust the rates of these inputs according to the response of each zone identified by the proposed method. The results of this study can be applied to reevaluate expectations on variable rate prescriptions guided largely by soil and variability.

Keywords: *Zea mays* L., nitrogen, chessboard design, geographically weighted regression, yield response functions.

Introduction

In the last decades, site-specific technologies including yield monitors, remote sensing imaging, and variable rate input application have become increasingly available to farmers. Many farmers have access to spatial information, and software tools to process this information into better decision. However, this decision-making process is usually aided by a knowledge base largely supported by agronomic trials, which involves changing management practices and subsequent monitoring of the effect on the system output (Pringle et al. 2004). Traditional agronomic experimentation is usually conducted in small plots that by design are managed to generate broad inferences from small field areas. This scenario was conceived in a time when all experimental steps were labor intensive, restricting the feasible extension of the plots. For decisions interacting with environment and management, the results obtained in these experiments may not align with farmers field's conditions.

To improve the adoption of site-specific management, it is necessary to obtain additional information about how site-specific treatments impact crop responses (Bullock and Bullock 1994). Recent advances in GPS based technologies have enabled agronomic experiments to be applied on-farm and in much larger areas than was previously possible. Such experiments generate large amounts of site-specific response data that can be used to understand the spatial variation of optimum rates. This involves approximating site-specific yield response functions, which requires methods of spatial analysis (Bullock and Lowenberg-DeBoer 2007). Site-specific prescriptions of inputs often make the

assumption that crop responses are expressed by a linear function of yield or some soil parameter, given that there is insufficient information about site-specific responses to farm management. The Data Intensive Farm Management (DIFM) project has developed a platform to carry out many on-farm experiments and better understand site-specific yield responses.

The most prominent field design applied at the DIFM is the chessboard arrangement covering the entire field. This arrangement generated an unprecedented amount of information, requiring appropriated statistical methods to extract meaningful information. Geographically weighted regression (GWR) is a statistical methodology that can be used to highlight local differences in the relationship between spatial variables. The use of this data analysis technique is particularly suitable for OFPE field arrangements since the treatments are randomly assigned to entire fields. In the GWR, the parameters of the models are allowed to vary in space and can be mapped and interpreted as a spatial variable (Fotheringham 1997; Plant 2012). It can be used as an effective method to evaluate the assumption of stationarity in situations of significant variation of regression coefficients. Therefore, the objective of this study was to apply GWR to evaluate spatial variability of corn response to nitrogen and seed rates in on-farm precision experiments.

Materials and methods

Field experiments
The datasets used in this research were generated from corn (*Zea maiz* L.) on-farm trials conducted under the DIFM project. The fields were chosen to represent typical corn production systems in the Corn Belt, with two rainfed fields from Illinois and two central pivot irrigated fields from Nebraska, in the 2017 growing season (Table 1). Each experiment had completely randomized designs (CRD) with at least four corn seeding rates (SR) and four N rate (NR). Each field had on average 40 ha and 200 experimental units. All treatment levels were implemented in the field using variable rate enabled seed planters and nitrogen applicators. Each experimental unit was 85 m long and about 18 m wide, representing two times the machinery swath width. Other farming practices were kept constant throughout the field and were conducted by the farmers in accordance with standard protocols for the region.

Table 1. Description of the four cornfields used for the on-farm precision experimentation.

Field*	Location	Irrig.	Seed Rate (seeds ha^{-1})	N Rate (kg ha^{-1})	ECS (mS m^{-1})	Elev. (m)	Yield (Mg ha^{-1})
1	Effingham – IL	No	66k – 96k	78 – 129	20 (4)	180	11.2 (1.7)
2	Moultrie – IL	No	76k – 96k	96 – 147	28 (5)	210	11.9 (0.4)
3	Hamilton – NE	Yes	69k – 89k	24 – 99	40 (7)	572	14.9 (0.6)
4	Clay – NE	Yes	69k – 91k	0 – 90	49 (6)	522	15.9 (0.8)

*Irrig: Irrigation; ECS: Soil electrical conductivity (0-30 cm); Elev: Elevation; Numbers inside parenthesis are the standard deviation.

Data collection and processing
Yield data was collected during harvest using the combine yield mapping systems. Digital elevation maps (DEM) and soil electrical conductivity (EC) were used to characterize field variability. DEM maps were created by interpolating the elevation obtained by the real-time kinematic global navigation satellite system receiver (RTK GNSS). Other

topographical indexes were derived from the DEM to represent slope, curvature, and terrain wetness index. EC data was collected previous to planting using a Veris U3 sensor system (Veris Technologies, Inc. Salina, KS). The yield point coordinates were used to sample the information from the interpolated EC and DEM maps. The length of the experimental unit was further divided to produce almost squared polygons (Figure 1a). These were used as the observational units, and all data was averaged to these polygons with approximately 400 m² (Figure 1b).
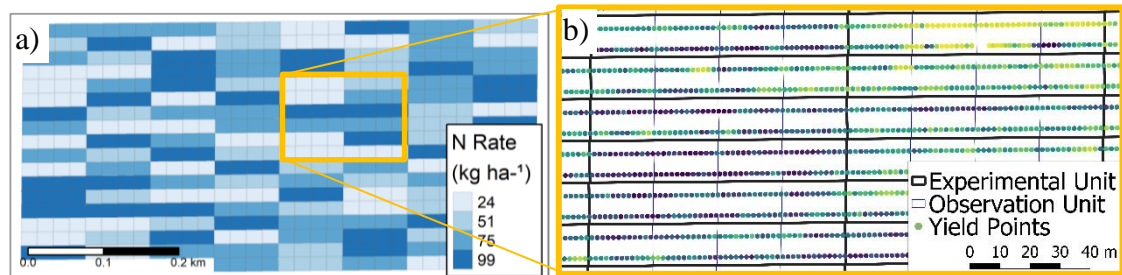

Figure 1. Field trial arrangement of the on-farm precision experiment in Field 3 at Hamilton-NE (a), and subsection of the field ilustrating yield monitor data and the observational units (b).

Methodology of Analysis
Based on ordinary least squares, GWR modifies the estimates of intercept (β_0) and regression coefficients (β_1) as a function of the geographic position (x_i, y_i) (Eq.1).

$$Y_i = \beta_0(x_i, y_i) + \beta_1(x_i, y_i) + \varepsilon_i, \qquad i = 1, \dots, n. \quad (1)$$

Following matrix notation regression coefficients are estimated as:

$$b(x_i, y_i) = (X'G(x_i, y_i)X)^{-1} X'G(x_i, y_i)Y, \quad (2)$$

Where Y is the vector whose components are the Y_i and X is a matrix containing the values of the explanatory variables X_i. In this notation, the GWR model incorporates local variation in the estimated coefficients $b(x_i, y_i)$ by incorporating a geographic weighting matrix G(x_i, y_i). This diagonal matrix incorporates the distance d_{ij} between predictors X_i at (x_i, y_i) and dependent observations Y_j at location (x_j, y_j). As d_{ij} increases, the explanatory variable is expected to decrease in influence over the response variable (Plant 2012). This decay is captured by elements g_{ij} of the matrix G (Eq. 3).

$$g_{ij}(x_i, y_i) = exp\left(-\left[\frac{d_{ij}}{h}\right]^2\right) \quad (3)$$

For this study, we used a Gaussian adaptive spatial kernel, with a bandwidth h set to include for each location several neighbors corresponding to 5% of the total number of data points. Following these specifications, GWR analysis was conducted using the R package spgwr (Bivand and Yu 2017). The regression model had yield as the dependent variable and the NR and SR effects as the independent variables, plus an intercept. Therefore, for each coordinate in the center of the Gaussian kernel, three model parameters were estimated, as well as statistical metrics such as the standard error and the R-squared.

The resulting coefficients of NR and SR were standardized between 0 and 1, by subtracting the minimum and dividing by the maximum value. These coefficients were labeled as seed rate response index (SRI) and nitrogen rate response index (NRI). As the index increases, so does the responsiveness to treatments. These indexes were submitted to unsupervised k-means clustering to create three response zones for each field to better illustrate the differences. The indexes were also compared with auxiliary variables representing the spatial variability of soil and terrain parameters, using correlation analysis.

Results and Discussion

Overall yield values were higher in the irrigated fields in Nebraska, averaging 15.4 Mg ha^{-1}, than the non-irrigated fields in Illinois, averaging 11.4 Mg ha^{-1} (Table 1). The coefficient of yield coefficient of variation (CV) was smaller than 10% for most fields, except for Field 1 in which the CV was 15%. The low CVs are probably due to the aggregation of the data in the polygons, which tends to smooth the variations, and to the high average yield observed for the fields, which denotes favorable weather conditions. Although the overall variation is not large, it is possible to observe the presence of spatial variation in the yield in the fields. The main distinct area in Field 3 (Figure 2a), is the low yielding region in the northwest of the field. Comparing the yield with the results of the GWR analysis in this area for the SR (Figure 2b), most of this area can be considered as highly responsive to the increment in seed rates, indicating a negative correlation between yield and the optimum SR. However, the east side of the field present the highest yields and also the highest responses to SR. The ability of using the results from the GWR analysis to make this type of observation, which would not be possible without mapping the regression coefficients, is one of the motivations for recommending its application in the analysis of OFPE data.

Similar results regarding the spatial variation of the response index were observed in the four fields used in this study. These results reiterate the importance of incorporating the spatial variability as one of the basis for agronomic experimentation. The presence of high responses in regions with contrasting yield levels can be seen as a good indication that yield potential or target yield alone are not good predictors of the optimum rates. Yield maps of previous years are commonly used as the deciding factor to delineate management zones or to make variable rate prescriptions. However, there is no guarantee that the optimum rate in two given regions with similar yields is in fact the same.
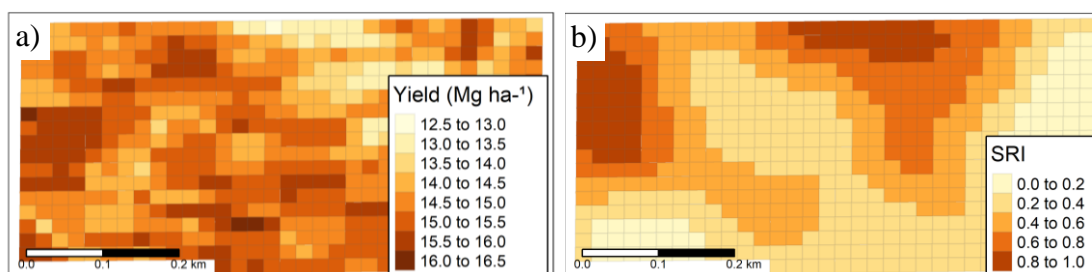

Figure 2. Corn yield (a), and the seed rate response index (b) in Field 3 at Hamilton-NE, for the 2017 crop season.

Unsupervised k-means clustering was used to generate yield response zones as an alternative to the traditional yield-based management zone delineation. There is no reason to assume the same zones are optimum for different inputs, therefore distinct response zones were created for SR and for NR in each field. The clusters were labeled as "High", "Medium", and "Low" according to the average response index in each group. The results were consistent across fields in terms of the delineation of meaningful spatial contiguous zones and the different yield response in each zone.

The two fields in Illinois were more responsive to NR than SR. In these fields, the response to NR was positive for the high and medium response zones and no response was observed in the low response zone (Figure 3). In the low response areas of this field, there seems to be no decrease in yield using the lowest N rate at 96 kg ha^{-1} in comparison with the higher rates, therefore there exists a potential to apply less N in this region of the field while maintaining the same yields. For the same fields, the response to SR was positive only in the high response zone, no response was observed in the medium response zone and negative response was observed in the low response zone. In this scenario, the average rate was already optimized for the average response of the field, but there is still potential to increase the yields using the same amount of SR, by redistributing it inside the field according to the response zones.

The two irrigated fields in Nebraska were more responsive to SR than NR. In these fields, the response to SR was positive for the high and medium response zones and no response was observed in the low response zone (Figure 4). In the high response zones the yield is increased by more than 1500 kg ha^{-1} when the SR is increased from 69k to 89k, in the medium response zone the yield increase for the same variation is still more than 1000 kg ha^{-1}, while in the low response areas of this field the increase is within the variation represented in the boxplot. Therefore, there is a great potential to adjust the seed rates according to the response of each zone. The response to NR in the same fields was positive only in the high response zone, no response was observed in the medium response zone and negative response was observed in the low response zone.
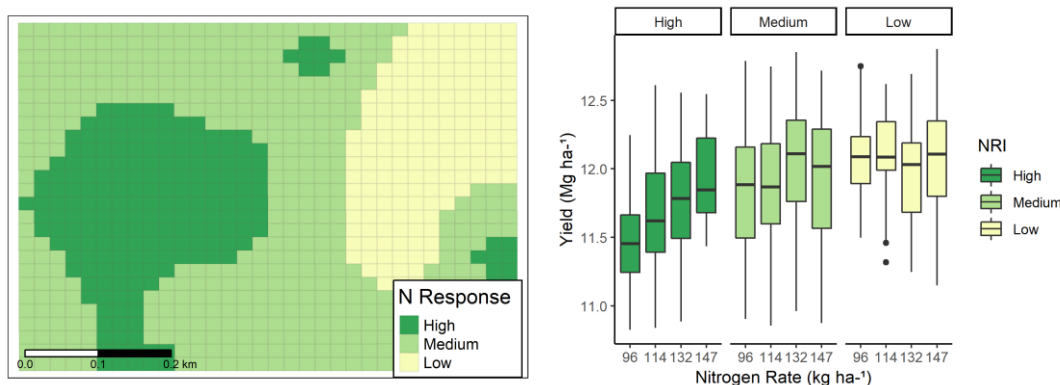

Figure 3. Zones of yield response to nitrogen (NRI) and corresponding box-plots showing the different yield response to nitrogen rate treatments, for the Field 2 at Moultrie, IL.
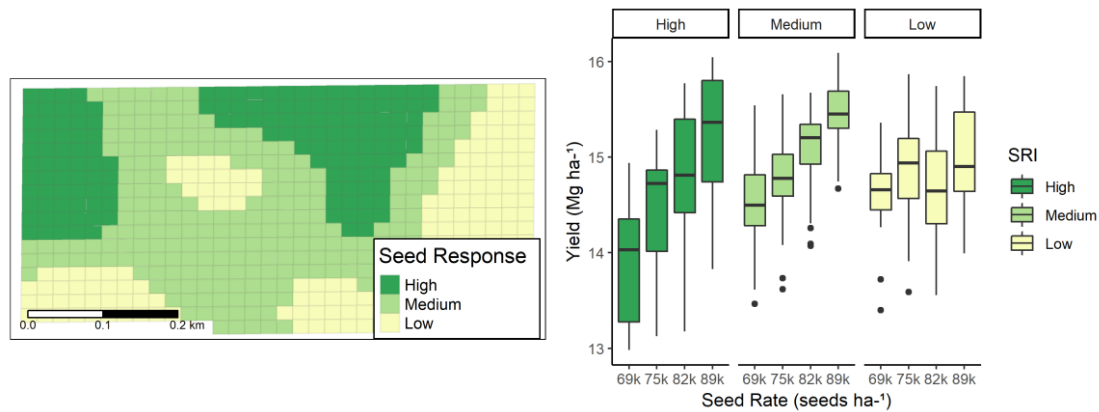
Figure 4. Zones of yield response to seed rate (SRI) and corresponding box-plots showing the different yield response to seed rate treatments, for the Field 3 at Hamilton, NE.

Correlations between the NRI and SRI and different field variables are shown in Table 2. In general, the field characteristics impacting SRI are different than the ones impacting NRI. Elevation had a dominant role impacting SRI with correlations of 0.76 and -0.60 for Field 1 and Field 3, respectively. For most of the fields the correlation between electrical conductivity and topographic characteristics were less than 0.30. These low correlations imply that there might be interactions between these field characteristics within the field, impacting these different yield responses to NRI and SRI treatments.

Depending on availability, more variables could be added to this list, including the yields and remotely sensed data in previous years. Multivariate statistical methods and other methods of data analysis can also be employed to explore the additive contribution of different factors. Future work will also focus on developing econometrics models to identify optimal treatment rates based upon input prices and site-specific yield responses to NR and SR in fields. In addition, as more fields are added to the DIFM project, relationships between site-specific responses and site-specific characteristics will be further investigated.

Table 2. Correlation coefficient between the crop response to seed and nitrogen rate and field characteristics.

Field*	Variable	ECS	ECD	Elevation	Slope	Curvature	TWI
1	SRI	0.08	0.17	-0.60	0.19	0.00	0.17
	NRI	-0.26	-0.22	-0.02	0.20	-0.04	0.08
2	SRI	0.41	0.30	-0.38	-0.36	0.26	0.11
	NRI	-0.11	-0.01	0.19	0.12	-0.11	0.02
3	SRI	-0.17	-0.20	0.18	-0.05	-0.28	-0.08
	NRI	-0.13	-0.10	0.26	-0.09	-0.38	-0.14
4	SRI	-0.30	-0.28	0.76	0.58	-0.38	0.01
	NRI	0.16	0.08	-0.44	-0.29	0.27	-0.17

*ECS: Soil electrical conductivity (0-30 cm); ECD: Soil electrical conductivity (0-90 cm); TWI: Terrain wetness index; SRI: Seed rate responsiveness index; NRI: Nitrogen rate responsiveness index.

Conclusions

The DIFM project has developed a platform to carry out many on-farm experiments and better understand site-specific yield responses. The use of GWR identified differential yield responses to seed rate and nitrogen in cornfields in different producing regions in US Midwest. Therefore, there is great potential to adjust the rates of these inputs according to the response of each zone identified by the proposed method.

The presence of high responses in regions with contrasting yield levels can be seen as a good indication that yield potential or target yield alone might not be good predictors of the optimum rates. These results should be taken into account when setting expectations on variable rate prescriptions guided largely by soil and yield variability.

Acknowledgements

References

Bivand, R., & Yu, D. (2017). spgwr: Geographically Weighted Regression. https://cran.r-project.org/package=spgwr

Bullock, D. S., & Bullock, D. G. (1994). Calculation of optimal nitrogen fertilizer rates. *Agronomy Journal*, *86*(5), 921–923.

Bullock, D. S., & Lowenberg-DeBoer, J. (2007). Using spatial analysis to study the values of variable rate technology and information. *Journal of Agricultural Economics*, *58*(3), 517–535. doi:10.1111/j.1477-9552.2007.00116.x

Fotheringham, A. S. (1997). Trends in quantitative methods I: stressing the local. *Progress in Human Geography*, *21*(1), 88–96.

Plant, R. E. (2012). *Spatial data analysis in ecology and agriculture using R*. cRc Press.

Pringle, M. J., Cook, S. E., & McBratney, A. B. (2004). Field-scale experiments for site-specific crop management. Part I: Design considerations. *Precision Agriculture*, *5*(6), 617–624. doi:10.1007/s11119-004-6346-1